Evaluation of the Ability of AlphaFold to Predict the Three-Dimensional Structures of Antibodies and Epitopes

Ksenia Polonsky,*,† Tal Pupko,† and Natalia T. Freund*

Being able to accurately predict the three-dimensional structure of an Ab can facilitate Ab characterization and epitope prediction, with important diagnostic and clinical implications. In this study, we evaluated the ability of AlphaFold to predict the structures of 222 recently published, high-resolution Fab H and L chain structures of Abs from different species directed against different Ags. We show that although the overall Ab prediction quality is in line with the results of CASP14, regions such as the complementarity-determining regions (CDRs) of the H chain, which are prone to higher variation, are predicted less accurately. Moreover, we discovered that AlphaFold mispredicts the bending angles between the variable and constant domains. To evaluate the ability of AlphaFold to model Ab–Ag interactions based only on sequence, we used AlphaFold-Multimer in combination with ZDOCK to predict the structures of 26 known Ab–Ag complexes. ZDOCK, which was applied on bound components of both the Ab and the Ag, succeeded in assembling 11 complexes, whereas AlphaFold succeeded in predicting only 2 of 26 models, with significant deviations in the docking contacts predicted in the rest of the molecules. Within the 11 complexes that were successfully predicted by ZDOCK, 9 involved short-peptide Ags (18-mer or less), whereas only 2 were complexes of Ab with a full-length protein. Docking of modeled unbound Ab and Ag was unsuccessful. In summary, our study provides important information about the abilities and limitations of using AlphaFold to predict Ab–Ag interactions and suggests areas for possible improvement. ***The Journal of Immunology*****, 2023, 211: 1578–1588.**

Antibodies are the basis of all approved vaccines and are major correlates of protection in all vertebrates (1–4). Physiologically, Abs are produced by B cells following immunization or infection (5). Importantly, these B cells have the ability to improve and affinity mature their presented Abs while also differentiating into Ab-secreting plasma cells and a specific subset of memory B cells. Subsequently, memory B cells and the Abs they produce are largely responsible for preventing reinfection and reducing the severity of the disease during secondary encounters (6). Their exquisite specificity and affinity make Abs an appealing class of drugs that are widely used in the clinic for the treatment of cancer (7); autoimmune disorders (8); and, more recently, infectious diseases (5). Over the last two decades, innovative engineering and single-cell and high-throughput cloning techniques have significantly advanced the ability to generate new Abs against various specific targets. Technical advances and optimized expression protocols now enable rapid generation of a large number of Abs, so that sometimes the Abs can be tested for their therapeutic activity in animal models within 1–2 wk of collection of the original specimen (9). Such Abs also can be used for diagnosis and as guides for vaccine design (10, 11, 12, 13).

Despite a deserved sense of achievement and accomplishment, one significant bottleneck remains within "the Ab pipeline" of every preclinical and clinical Ab, namely deciphering the mechanism of action. This usually involves investigation of the Ab structure and identification of the Ab binding site (i.e., the precise epitope on the target) (12). Precise and detailed information regarding the epitope is crucial for understanding the Ab's functions and for predicting possible escape mechanisms (14–16). If the three-dimensional (3D) structure of the Ab–Ag complex is not available, there are a variety of computational tools that can be used to dock the 3D structure of the Ab with the 3D structure of Ag. Such docking algorithms have been used with various degrees of success (17, 18) but generally require separately solved crystal structures of the Ab and the Ag. Unfortunately, although the Ag structure is usually known and resolved, the atomic coordinates of the Abs generated against it are often lacking, and their solution requires weeks and sometimes months. Thus, despite the advances in Ab isolation and sequencing and the new computational and structural epitope-mapping methods, delineating the structure of both the Ab and its epitope is still a major challenge (12, 19–21). Being able to model the structure of a new Ab based solely on the primary amino acid sequence will provide a better understanding of Ab function and Ab–Ag interactions and will greatly contribute to the use of Abs in research and clinical applications (22–24).

*Department of Clinical Microbiology and Immunology, Faculty of Medicine, Tel Aviv University, Tel Aviv, Israel; and †Shmunis School of Biomedicine and Cancer Research, George S. Wise Faculty of Life Sciences, Tel Aviv University, Tel Aviv, Israel

ORCIDs: 0000-0003-0748-9008 (K.P.); 0000-0001-9463-2575 (T.P.); 0000-0002-5386-502X (N.T.F.).

Received for publication March 3, 2023. Accepted for publication September 6, 2023.

This work was supported by the Israel Science Foundation Grants 2818/21 (to T.P.) and 3136/22 (to N.T.F.) and United States–Israel Binational Science Foundation Grant 01031771 (to N.T.F.).

K.P. planned and performed the experiments, analyzed data, prepared the figures, and wrote the manuscript together with N.T.F. and T.P. N.T.F. and T.P. planned and supervised the experiments and data analyses and wrote the manuscript together with K.P. All the authors meet the authorship criteria and gave their consent to be listed as authors on this article.

Address correspondence and reprint requests to Dr. Natalia T. Freund or Prof. Tal Pupko, Department of Clinical Microbiology and Immunology, Faculty of Medicine, Room 802, Tel Aviv University, Klachkin Street 35, Ramat Aviv, Tel Aviv-Yafo 6997801, Israel (N.T.F.) or The Shmunis School of Biomedicine and Cancer Research, The George S. Wise Faculty of Life Sciences, Britannia Building, Room 405, Tel Aviv University, Klachkin Street 35, Ramat Aviv, Tel Aviv-Yafo 6997801, Israel (T.P.). E-mail addresses: nfreund@tauex.tau.ac.il (N.T.F.) or talp@tauex.tau.ac.il (T.P.)

The online version of this article contains supplemental material.

Abbreviations used in this article: 3D, three-dimensional; CASP, Critical Assessment of Protein Structure Prediction; CDR, complementarity-determining region; FR, framework region; GDT, global distance test; PDB, Protein Data Bank; RMSD, root mean square deviation; TM, template modeling.

https://doi.org/10.4049/jimmunol.2300150

Predicting protein structure based solely on the coding of primary amino acid sequences is often referred to as "the protein folding problem." This refers to the challenge of addressing protein complexity and the distinct quaternary structure with a significant degree of flexibility in some areas and rigidity in others (16, 25). Critical Assessment of Protein Structure Prediction (CASP) is a biennial community experiment designed to determine the state of the art in modeling protein structure (26). Participants are provided with the amino acid sequences of target proteins and build models of the corresponding 3D structures. DeepMind's AlphaFold artificial intelligence system (27) has demonstrated remarkable results in CASP14 (26), achieving a median global distance test (GDT) score of 0.92, where the GDT score values range from 0 to a perfect score of 1.0. This was a significant leap compared with the previous year, when the score was 0.59 (28–30). However, although the analysis included 84 experimental models and 152 different protein prediction targets related to these models, it did not specifically focus on Ab evaluation and included only 2 prediction targets related to a single Ab model (Protein Data Bank [PDB] identifier 6VN1) (31). Therefore, although AlphaFold is an outstanding tool for predicting protein folding, its accuracy in modeling Ab structures remains unknown.

Abs share common conserved areas whose structure is easy to predict by homology. In contrast, predicting the structure of the Ab variable regions is substantially more challenging because their generation via a genetic recombination between randomly selected genome-encoded V, D, and J regions introduces a high degree of diversity. In addition, the sequence and subsequently the structural diversity are further increased by the insertion of n-p nucleotides and later random mutations into the coding sequence during the affinity maturation that follows Ag exposure (13, 14, 32, 33). This variability greatly reduces the availability of closely related examples from which structure prediction algorithms can learn, and thus these regions are predicted with lower accuracy, leading to a poor overall prediction accuracy for Ab structures (34, 35).

This study was designed to evaluate the prediction accuracy of AlphaFold as related to Abs. Specifically, we asked the following questions: (1) What is the average accuracy of AlphaFold in predicting atomic structures from primary amino acid Ab sequences? (2) Which Ab chains are predicted with better or worse accuracy? (3) What structural elements within the Ab chain are accurately inferred, and which regions are prone to mistakes? (4) What types of mistakes can be expected when AlphaFold is used to predict the 3D structures of Abs? (5) How well can AlphaFold predict specific epitopes on the surface of the corresponding targets, and how does the level of accuracy compare with that of ZDOCK, which is an alternative prediction algorithm based on docking (36)?

Materials and Methods

Data

AlphaFold prediction quality was evaluated on Abs with crystal structures available in the PDB (37). A nonredundant set of Ab structures was extracted from the SAbDab Abs database (38), with a maximum sequence similarity of 80% and a resolution cutoff of 3 Å. Only structures that were released after the AlphaFold training cutoff date of April 30, 2018, were considered. Importantly, we focused on the Fab regions of the Abs because these regions confer most of the specificity of an Ab to the corresponding Ag. H and L chains were analyzed separately, and the dataset included different species (Table I). In total, we analyzed 222 Fab chains, 95 full Fab structures, and 26 Ab–Ag complexes. Full data used in the analysis can be found at https://github.com/XseniaP/AF_evaluation/blob/master/Supplementary_Tables.xlsx.

AlphaFold

The full version of AlphaFold version 2.2.0 (https://github.com/google-deepmind/alphafold) without "templates" (no homologous structure search) was used for structure predictions (39). AlphaFold also uses the BFD database, which is currently one of the largest publicly available sets of protein sequences (26, 40, 41), with the original version of this database including more than 2.5 billion protein sequences clustered in families. DeepMind has already added many of the molecules predicted by the complete AlphaFold system to the AlphaFold Protein Structure Database (42), but not all structures of interest are present in the database as a whole, and we therefore decided to run AlphaFold version 2.2.0 locally to predict all possible structures.

Source code and program

We developed a Python script (the Script) (https://github.com/XseniaP/AF_evaluation) to perform the calculations and combine all accuracy evaluation steps into a single pipeline (Fig. 1). The script is used to call the externally developed open-source products for alignment, superimposition, metrics calculation, and visualization and then parses the results and moves the relevant information to the next step. In addition, the script provides on-demand calculations of the key metrics per domain and color coding, which are not covered by the externally developed packages.

Identifying structural subdomains

The variable fragment of the Fab domain of the Ab is formed by the H and L chains and is responsible for the specificity of the Ab for its Ag. As the first step and to identify problem areas, the quality of tertiary structure predictions for 222 Fab structures was analyzed, with each chain considered separately. Next, we assessed the quality of the prediction for the structural domains of the Fab (i.e., variable versus constant domains). Last, we assessed the structural subdomains within each domain (i.e., the three hypervariable loops), known as CDRs of the H and L chains (CDRH1, CDRH2, CDRH3, CDRL1, CDRL2, and CDRL3), which form the specific Ag recognition site on the surface of the Ab (43), as well as the less variable framework regions (FRs).

Kabat (44) and Chothia (45, 46) are two variable domain numbering schemes that are often used to define the location of the variable fragment regions in the sequence. Two common programs for processing immune repertoire sequencing data are IMGT (47) and NCBI IgBLAST (48), where the latter has an offline version, and Python library wrappers exist for data processing. In this study, we used the IMGT reference directory (49, 50) for Igs as a basis for generating a germline database for each species. The databases were also used as a prerequisite for the PyIR wrapper (51) to incorporate the IgBLAST-based mechanism into our Python script and identify the chain type as well as the CDR1/2/3, FR1/2/3 regions of the variable domain. If the location of the end of the CDR3 region was missing in the PyIR output, the result was verified and completed using known CDR3 motifs (43, 52–54) for each of the corresponding chain types, as presented in Supplemental Table III (the table can be found at https://github.com/XseniaP/AF_evaluation/blob/master/Supplementary_Tables.xlsx in full data file). For example, the C terminal of an H chain frequently juxtaposes a Trp-Gly-XXX-Gly motif, and the TGGGG motif often appears as an end site of the CDR3 domain in the H chain. The motifs for each of the chain types were only used to verify and complement the CDR3 region in the corresponding chains when this information was missing in the output of PyIR.

Superimposing the molecules

The predicted and experimentally determined 3D structures (the native structure) were aligned by TM-align (55), which employs the coordinates of the backbone carbon α (C_α) of a given protein structure for superimposition and distance calculations. TM-align was run in two different modes. First, metrics were calculated using sequence-independent alignment; that is, the superimposition was based on structural similarity. Such an analysis should produce the closest superimposition with the lowest distance between two molecules but may not align the same domains of the native and predicted molecules and thereby prevents the analysis of the prediction quality by domain or identification of the exact location at which the prediction accuracy deteriorates. We therefore repeated the calculations with sequence-dependent alignment, where the residue index correspondence between two structures is included as a constraint. MM-align (56), developed for sequence-independent alignment of complex protein structures, was used to compare predicted and native multimer structures, such as an entire Fab fragment and Ab–Ag structures. We also used the ChimeraX MatchMaker (57) tool, which is part of University of California, San Francisco Chimera (58). This algorithm allows us to superimpose complex protein structures by first creating a pairwise sequence alignment between the selected chains and then fitting the aligned residue pairs and calculating the distance between them.

Accuracy evaluation

The prediction accuracy was calculated from parameters selected from the official metrics used by CASP and Ab modeling assessment (39), where d_i is defined as the distance between the i^{th} pair of aligned residues in the metrics described below.

1. Root mean square deviation (RMSD) values of the subset of C_α atoms that correspond to the residues from the target crystal structure after sequence-independent structural superimposition of the two atoms. The TM-align sequence-independent mode was applied; that is, the superimposition was generated on the basis of minimum distance without any constraints on the sequence alignment. RMSD values were obtained as part of the standard TM-align output based on the following formula:

$$RMSD = \sqrt{\frac{\sum_{i=1}^{L_{ali}} d_i^2}{L_{ali}}}$$

where L_{ali} represents the number of aligned residues.

MM-align was used to calculate the RMSD values based on the differences between native and predicted complex structures.

2. RMS_CA: RMSD for the entire target structure was calculated on C_α atoms based on sequence-dependent structural superimposition, which assumes residue index correspondence between the molecules. The TM-align sequence-dependent model was applied, and the metrics were calculated by our script based on the following formula:

$$RMS\_CA = \sqrt{\frac{\sum_{i=1}^{n} d_i^2}{n}}$$

where n represents the total number of residues in the molecule.

ChimeraX MatchMaker was used to calculate RMS_CA values based on the differences between native and predicted complex structures while creating a pairwise sequence alignment.

3. GDT quantifies protein dissimilarity. Notably, GDT is less sensitive to outlier regions that might result from the poor prediction of a specific region (e.g., a CDR loop), whereas the rest of the model is reasonably accurate (59). The two GDT measures are the GDT total score (GDT_{TS}) and the GDT high accuracy score (GDT_{HA}). These measures were calculated on two structurally superimposed molecules in both sequence-dependent and sequence-independent mode. The GDT_{TS} was calculated according to the following formula:

$$GDT_{TS} = \frac{GDT_{P1} + GDT_{P2} + GDT_{P4} + GDT_{P8}}{4}$$

GDT_{HA} was calculated according to the following formula:

$$GDT_{HA} = \frac{GDT_{P0.5} + GDT_{P1} + GDT_{P2} + GDT_{P4}}{4}$$

where GDT_{Pn} denotes the percentage of residues with a distance equal to or less than n Å. We have implemented these computations with our Python program. The formulas are those used by CASP to evaluate prediction accuracy (25, 59). The models submitted to the CASP competition are ranked on the basis of GDT_{TS} score, which is used as the major assessment criterion. GDT_{TS} provides an estimate of the percentage of residues predicted under specific cutoff distances, and GDT_{HA} provides an estimate of the percentage of residues predicted with high accuracy.

4. TM-score: The template modeling (TM) score is a metric for assessing the topologic similarity between two protein structures. We used the TM-score, which is provided as a standard output of the TM-align, and is defined as (55):

$$TM - score = Max\left[\frac{1}{L_N}\sum_{i=1}^{L_{ali}}\frac{1}{1+\left(\frac{d_i}{d_0}\right)^2}\right]$$

such that L_N is the length of the native protein that other structures are aligned to; L_{ali} is the number of aligned residues; $d_0 = 1.24\sqrt[3]{L_N - 15} - 1.8$ is a scale to normalize the match difference and to rule out protein size dependence. This is based on an estimation of the average structure distance between aligned residues of the random related structures of the length L_N.

TM-score value ranges from 0 to 1, which indicates a perfect match between two structures. Scores below 0.2 indicate randomly chosen unrelated proteins.

AlphaFold-Multimer (60, 61) was used to predict structures of H and L chains bound together. Using this analysis, we could better evaluate the prediction accuracy of AlphaFold, with the following metrics:

5. Elbow angle is a measure of the orientation between the variable and the constant domains in the Fab region. The elbow angle has been shown to increase Fab flexibility and enhance the ability of the Ab to recognize different Ags (62). Ag binding causes an apparent shift of the angle, which reflects the conformational changes that occur (63). The elbow angle also plays an essential role in Ab assembly (24, 64). Hence, the elbow angle is critical for Ab structure functionality and an essential parameter in modeling and Ab engineering. The Fab elbow angle was computed using RBOW (62). The difference in elbow angle between the predicted and native structures was used to quantify the accuracy.
6. VH-VL orientation in the Fab region of Abs was measured by the ABangle computational tool using five angles (HL, HC1, LC1, HC2, and LC2) and a distance (dc) (65).
7. DockQ score is a quality measure for protein–protein docking models derived by combining F_{nat}, LRMS, and iRMS to a single score in the range [0, 1]. DockQ values can be interpreted as follows: $0.00 \leq DockQ < 0.23$ corresponds to "incorrect," $0.23 \leq DockQ < 0.49$ to "acceptable quality," $0.49 \leq DockQ < 0.80$ to "medium quality," and $DockQ \geq 0.80$ to a "high quality" docking model (66).

Analyzing VDJ mutations

We obtained the germline information for each of the Ab chains from the PyIR wrapper (51) output and counted the number of amino acid substitutions in the VDJ region compared with the germline. This number was used to define four levels: low (under 5 aa substitutions per chain), medium (under 10), high (under 20), and extensive (20 or more).

Visualizing and color coding results

A PyMol version 2.5.2 (67) molecular visualization system script was written to color code the chains. ChimeraX (58) was used to visualize superimposed multimer structures.

Docking

The Ab–Ag complex structure can be predicted by AlphaFold-Multimer. It can also be predicted by molecular docking, which takes known or predicted 3D structures of an Ab and its corresponding Ag as input and returns a single 3D structure of the Ab–Ag complex. Docking was performed using the ZDOCK 3.0.2 web server (36).

Results

Overall prediction accuracy

All evaluation steps were developed into single script (see *Materials and Methods*) and combined into a single pipeline (Fig. 1). AlphaFold predictions of the published structures of 222 Ab chains (summerized in Table I) generated scores that were generally similar to the official CASP14 results for protein prediction (27–29, 68): The mean and median RMSD scores were 2.38 Å and 2.13 Å, respectively with 43.69% of the chains scoring lower than 2Å and 70.27% of the sample scoring lower than 3 Å (Table II). The average GDT_{TS} and GDT_{HA} values for all the predictions in our set were 0.72 and 0.50, respectively, where a GDT_{TS} of 0.72 (out of 1.0) reflects a

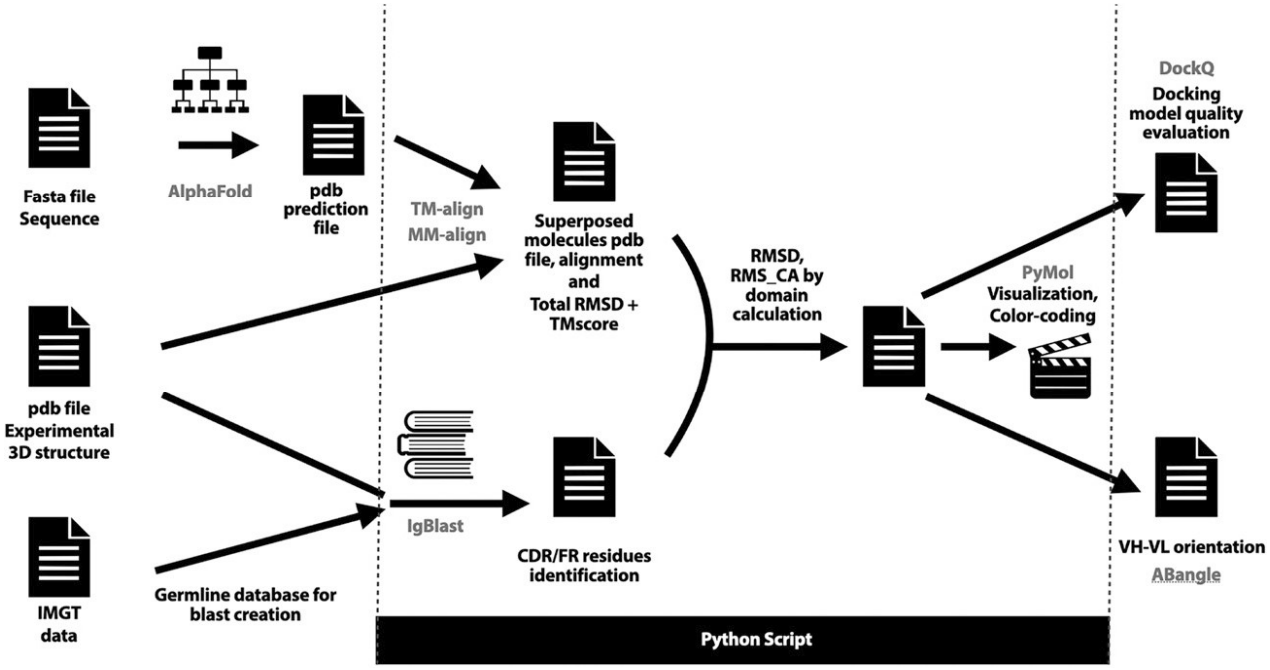

FIGURE 1. Study visualization chart. The analysis workflow starts with a target sequence in Fasta format, a PDB file of the native structure, and a PDB file of the structure predicted by AlphaFold version 2.2.0, which corresponds to the input sequence. IMGT reference directory data are predownloaded, and a germline database is created accordingly. We then use TM-align and MM-align to superimpose the native and predicted structures, which provides overall RMSD and TM score measures. We also run PyIR wrapper of IgBLAST to detect regions in the sequence corresponding to the CDRs/FRs of the antibody. Thereafter, additional metrics per region are calculated to identify the regions with the lowest prediction quality. Elbow angle is calculated using RBOW software, and VH-VL orientation is obtained using the ABangle computational tool. Docking model quality was evaluated by using DockQ software. We also produce PyMol scripts to color code the molecule residues based on their prediction quality reflected by RMSD, and we used ChimeraX to compare the predicted and native multimer structures.

good approximation to the native 3D structures. The GDT_{HA} score, which penalizes larger deviations from the native structure, suggests that, on average, 50% of the positions within a Fab can be predicted with very high accuracy. The average TM score for all individual H and L chains in our analysis was 0.83, which represents not only the same fold but also actually a very close match between the predicted and native structures.

Prediction accuracy for difference chain types and for different species

The mean RMSD values were higher in the H chains than in the L chain: 2.68Å compared with 2.09 Å, respectively (two-tailed *t* test *p* value 1.98×10^{-5}; Table III). The highest observed values excluding outliers according to the interquartile range method were 4.96 Å for H chains versus 4.81 Å for L chains, whereas the lowest values excluding outliers according to the interquartile range method were 1.05 Å for H chains versus 0.72 Å for L chains (Fig. 2A). λ L chains had a higher average RMSD value (two-tailed *p* value 2.12×10^{-5}) than κ L chains (2.74 versus 1.79 Å respectively; Table III, Fig. 2B). Similar results of 0.72 and 0.77 for λ and κ L chains, respectively, were obtained when the GDT_{TS} score was used to quantify prediction accuracy (Supplemental Fig. 1). Comparing Fab chains from different species revealed a slight, albeit statistically significant, increase in RMSD in *Macaca mulatta*, with a mean RMSD value of 3.43Å compared with other species for which the mean RMSD was in the range 1.61–2.52 Å (ANOVA one-way test *p* value 4.13×10^{-4}; Supplemental Fig. 2). The lower predictability of *M. mulatta* Fabs may stem from the underrepresentation of *M. mulatta* structures in the training set of the AlphaFold version 2.2.0 model as well as the low representation in our set.

The number of somatic hypermutations within the VDJ region has little effect on the AlphaFold prediction accuracy

As the next step, we tested the hypothesis that prediction accuracy is affected by the number of VDJ mutations. To this end, we grouped the analyzed Ab chains by the number of VDJ amino acid substitutions from the germline into low (<5), medium (between 5 and 9), high (between 10 and 19), and extensive (20 or more). The resultant mean RMSD values for each group were 2.00, 2.47, 2.42, and 2.56 Å, respectively. Although the differences among the groups are statistically insignificant ($p > 0.05$ by ANOVA), the value for the low group is significantly lower than that for all the others ($p < 0.05$

Table I. Distribution of the 222 analyzed Fab chains

	Chains Distribution
Species	Human = 140, others = 82 (including *Macaca mulatta* = 8, mouse = 61, rabbit = 11, rat = 2)
Chain type	L chain = 112 (κ = 77, λ = 35), H chain = 110
Year published	2020–2023
Resolution	1.3 – 3 Å

High-resolution Ig Fab fragment chains were chosen for the analysis to ensure that the distance between the predicted model and the reference molecule is not affected by resolution. All the structures analyzed were published after AlphaFold version 2.2.0 training cutoff date of April 30, 2018.

Table II. The 222 predicted chains' distribution by RMSD value against corresponding original molecules

RMSD Value	Number of Chains/Predictions	% of Total
<2 Å	97	43.69
<3 Å	156	70.27
<4 Å	205	92.34
Total	222	100

Table III. Predicted chains' distribution by RMSD value against corresponding native molecules according to chain type and L chain class

	H Chains		L Chains		L κ Chains		L λ Chains	
RMSD	Number of Predictions	% of Total	Number of Predictions	% of Total	Number of Predictions	% of Total	Number of Predictions	% of Total
<2 Å	36	32.73	61	54.46	52	67.53	9	25.71
<3	65	59.09	91	81.25	68	88.31	23	65.71
<4 Å	96	82.27	109	97.32	77	100	32	91.43
Total	110	100	112	100	77	100	35	100
Mean value	2.68		2.09		1.79		2.74	

by one-tailed *t* test), even though the difference is represented by a small value (below 0.5Å).

Difficulty in predicting the correct angles within the Ab

The chains with a lower prediction quality and higher RMSD were examined more closely. As a case study, we selected the H chain fragment from the human anti-HIV-1 neutralizing Ab BG24 Fab (PDB identifier 7UCE) (69). Discrete AlphaFold predictions of the H chain constant and variable domains were highly accurate: RMS_CA of 0.48 Å and 0.42 Å for individually predicted constant and variable domains, respectively (Fig. 3A, 3B). However, a consideration of the entire H chain resulted in an RMS_CA value of 4.93 Å, indicating a mediocre prediction (Fig. 3C). Superimposing the individual domain predictions onto the prediction of the H chain (Fig. 3D) suggests that the lack of fit is not caused by individual domain predictions, but rather reflects a difficulty in predicting the angle between the variable and constant domains within the chain. This caused large deviations in one of the domains, leading to the high RMS_CA value.

L and H chains exist as a heteromeric complex, and their correct assembly and interaction are critical for molecular function. It is therefore important to predict the entire Fab accurately. AlphaFold-Multimer predictions for 95 entire Fabs produced RMS_CA scores ranging from 0.67 Å to 4.65 Å, with a median value of 2.14 Å. Of these Fabs, 48.42% had scores less than or equal to 2 Å (Table IV). The prediction accuracy was not strongly correlated with the elbow angle itself (Fig. 4A). However, there was a significant correlation of $R^2 = 0.78$ between the total RMS_CA of the Fabs and the absolute value of the elbow angle deviations between the predicted and native structures (Fig. 4B). This suggests that the errors in angle estimation are responsible for the poor prediction of the Fab structure. Furthermore, our analyses indicate that AlphaFold tends to overestimate the elbow angle, with average values of 161.0 and 172.6 degrees, respectively, for the native and predicted Fabs (two-tailed *t* test *p* value 0.0035). Moreover, the range of predicted elbow angles was more restricted, with a SDs of 30.83 and 22.32 degrees for the elbow angles of native and predicted structures, respectively. Fabs for which the predicted elbow angle deviated insignificantly from the native elbow angle (difference under 20 degrees) demonstrated high prediction quality, with a mean RMS_CA of 1.45 Å (Fig. 4C). The analysis of VH-VL orientation angles did not detect significant differences between the native and predicted structures. In addition, 10 selected variable domains (Fv regions) were all predicted with a very high accuracy and average RMSD of 0.47 Å and 1.13 Å between pruned atom pairs and across all pairs, respectively (Supplemental Table I). This analysis reconfirmed our conclusions about high prediction accuracy of VH-VL orientation angles. Overall, the results demonstrate that the difficulty in predicting the angle between the variable and constant domains in the Fab region is a general issue with AlphaFold.

The accuracy of predicting different structural elements within Fabs

Our analysis of the accuracy of predicting the various domains within an Ab indicate that CDR3 in the H chain and CDR1 and CDR3 in the L chain are predicted with lower accuracy: H chain mean RMSD values for CDR1/2/3 and FR1/2/3 were 2.50Å, 2.24Å, 3.60Å, 1.99Å, 1.81Å, and 1.74Å, respectively (Fig. 5A), whereas the L chain mean RMSD values for CDR1/2/3 and FR1/2/3 were 2.40Å, 1.58Å, 2.43Å, 1.94Å, 1.72Å, and 1.71Å, respectively (Fig. 5B). When considering the secondary structures within the Fabs, helices were more difficult to predict than sheets in Ab H chains: the mean RMSD values for H chain helices and sheets were 3.04 Å and 2.18 Å, respectively, with a two-tailed *t* test *p* value of 2.3×10^{-6} (Fig. 5C). The corresponding mean RMSD values in the L chain were 1.87 Å and 1.76 Å, respectively, with no statistically significant difference (Fig. 5D). These results were reconfirmed when the analysis was repeated by comparing sequence-dependent structural superposition: the RMS_CA value for CDR3 in H chains was 4.2Å with other regions ranging between 1.53 and 2.35 Å (Fig. 5E), and the RMS_CA values for CDR1and CDR3 in L chains were 2.80 Å and 2.79 Å, respectively, with other regions ranging between 1.80 Å and 2.12 Å (Fig. 5F).

Docking outperforms AlphaFold-Multimer in predicting accurate Ab–Ag complexes

As the next step, we wished to map the Ab binding site by predicting 26 different Ab–Ag complexes. In 11 of 26 complexes, the Ag was

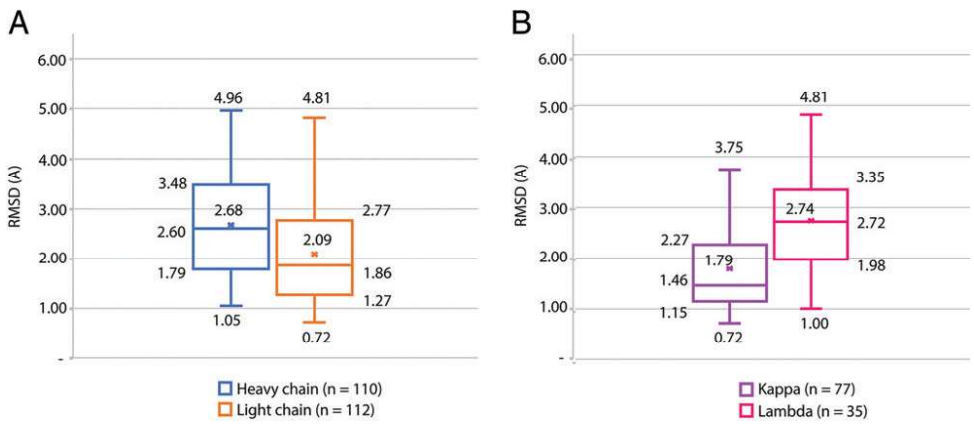

FIGURE 2. Prediction quality evaluation by chain type and L chain class. Boxplot of RMSD (Å) data presenting the distance between superimposed native and predicted molecules for (**A**) L in orange, ($n = 112$) and H in blue ($n = 110$) chains and (**B**) κ in purple ($n = 77$) and λ in magenta ($n = 35$) chains. The bottom and the top whiskers represent the minimum and the maximum value in the dataset excluding any outliers, the bottom and the top of the box are the 25th (Q1) and 75th (Q3) percentiles, the line inside the box is the 50th percentile (Q2) or the median value in the dataset, the x inside the box is a mean value of the dataset.

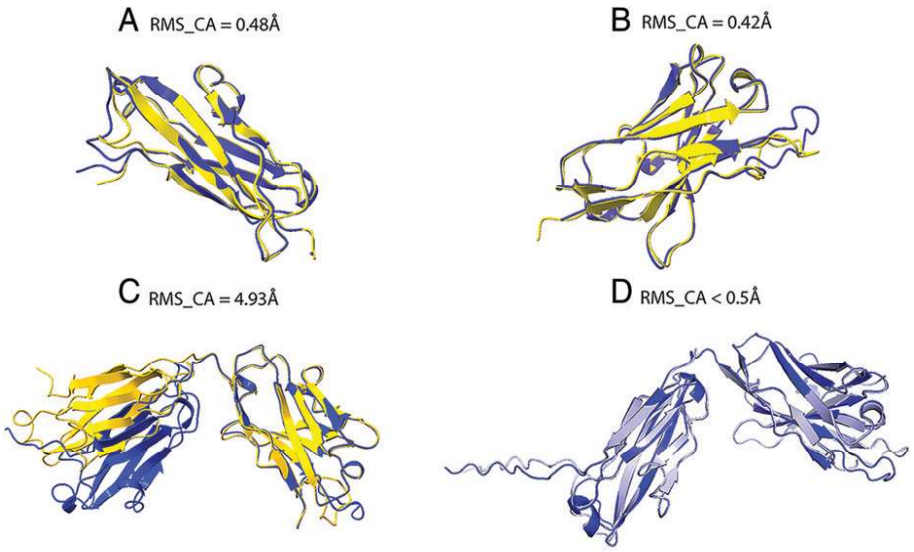

FIGURE 3. Structural analysis of the Fab fragment from the anti-HIV-1 neutralizing Ab BG24 H chain (PDB identifier 7UCE) superimposed on the AlphaFold prediction according to the sequence-dependent alignment. Cartoon diagrams of the H chain of the native 3D structure (in yellow) superimposed onto the same domain individually predicted by AlphaFold version 2.2.0 (in blue). (**A**) Constant domain. (**B**) Variable domain. (**C**) The complete H chain (variable and constant domains) presented as a cartoon diagram. (**D**) Cartoon diagrams of the variable and constant domains predicted separately (in light gray) but superimposed over the complete predicted H chain 3D structure (in blue).

a peptide 11–18 residues long. We tested four alternative options: (1) We assumed that the structures of both the Ab and the Ag are known and used the ZDOCK docking program to predict the Ab–Ag complex from the individual bound-extracted components; (2) we predicted the structure of the Ab alone using AlphaFold-Multimer first, and then we predicted the docking, using ZDOCK, with the atomic coordinates of the native structure of the Ag; (3) we predicted the structure of the Ag and the Ab using AlphaFold and AlphaFold-Multimer, respectively, and then docked them using ZDOCK; and (4) we skipped the docking altogether and asked AlphaFold-Multimer to predict the entire Ab–Ag complex from the protein sequences of the L chain, the H chain, and the Ag. We define docking prediction as medium quality if the DockQ score is in the range [0.49, 0.8] and as high quality if the DockQ score is equal to or greater than 0.80. The results indicated that when the native structures of all components were used as input for ZDOCK (alternative 1), 11 of the 26 complexes were docked with medium or high quality (DockQ score ≥0.49). Interestingly, in 9 of these 11 complexes, the Ag was a short peptide, whereas only 2 complexes that were predicted with medium or high quality were those of an Ab bound to a full-length protein. Notably, docking was based on the atomic coordinates of the individual Ab and Ag that were extracted from the Ab–Ag bound structure. This would not be the "regular" scenario, because, in most of the cases, one would predict an Ab–Ag complex by docking the atomic coordinates of unbound Ab and Ag. In contrast, 2 of the 26 complexes predicted solely from the protein sequences by AlphaFold-Multimer (alternative 4) were of medium to high quality, with a DockQ score ≥0.49 (Supplemental Table II). Significantly, none of the complexes was predicted with high accuracy when the predicted structures of both the Ab and the Ag (alternative 3) or the predicted structure of Ab and the native Ag structure (alternative 2) were used as input for ZDOCK. Overall, as could be expected, using the atomic coordinates of the individual Ab and Ag that were extracted from the Ab–Ag bound structures as input provided a more accurate docking prediction than docking with predicted structures or than predicting the entire complex using AlphaFold-Multimer. Most of the successfully docked complexes involved short-peptide Ags rather than full-length proteins. Of note, although both complexes that AlphaFold predicted with medium to high quality also scored high for individually predicted Ag or Ab, this does not always guarantee an accurate prediction of the complex by AlphaFold-Multimer (Fig. 6).

Table IV. Elbow angle and RMS_CA (Å) between 20 native and predicted sequence-dependently superimposed structures of predicted and corresponding native Fab molecules (H and L chain combined)

PDB Identifier	L Chain Class	Native Elbow Angle (degrees)	Predicted Elbow Angle (degrees)	Absolute Difference between Native and Predicted Fab (degrees)	RMS_CA (Å) between Native and Predicted Fab
6WAS	λ	226.6	215.3	11.3	2.8
7C88	κ	137.2	144.8	7.6	1.19
7L77	κ	136.6	162.1	25.5	3.05
7LYV	κ	135.8	185.6	49.8	4.09
7MF7	λ	187.3	201.3	14.0	2.05
7MU4	λ	232.0	195.6	36.4	3.62
7N3F	λ	194.2	195.0	0.8	1.78
7N4J	λ	226.4	209.0	17.4	2.54
7OX1	λ	135.5	207.2	71.7	3.38
7Q6C	κ	150.2	141.9	8.3	1.38
7O4Y	κ	135.0	174.2	39.2	4.04
7U0E	λ	207.0	190.0	17.0	2.74
7UCE	λ	198.0	189.0	9.0	1.61
7Z0Y	λ	222.9	190.2	32.7	3.66
8CZ5	κ	173.1	175.0	1.9	0.94
8D6Z	κ	147.1	144.1	3.0	0.96
8D47	κ	169.0	170.0	1.0	1.19
8DCC	λ	133.6	201.9	68.3	3.57
8DFG	κ	175.1	174.5	0.6	1.56
8DTX	κ	169.0	174.4	5.4	1.07

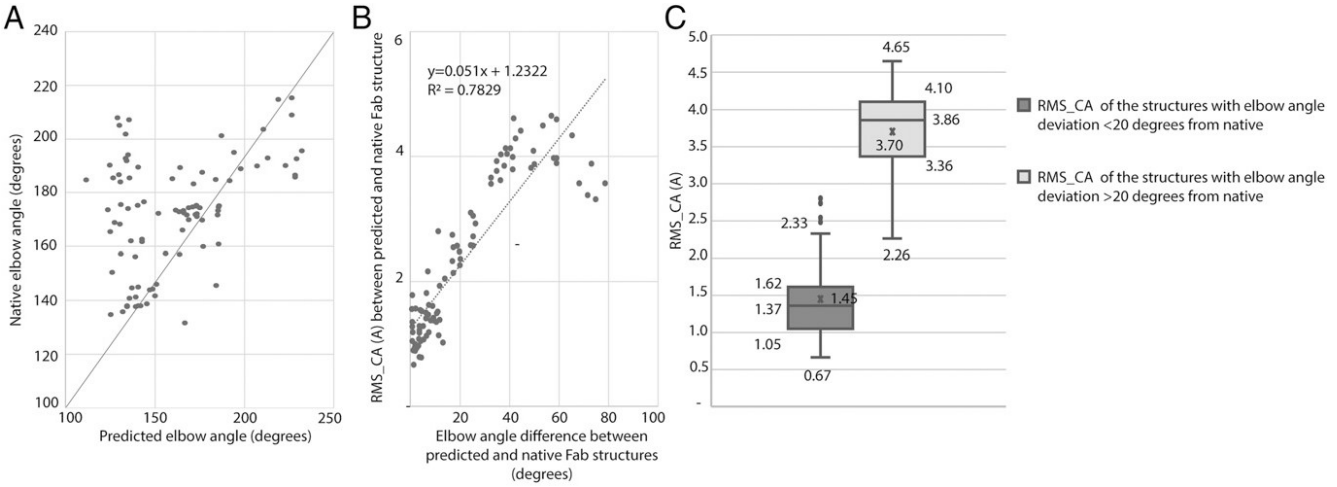

FIGURE 4. Evaluation of the elbow angle deviation correlation with prediction quality. Elbow angle deviations in correlation with prediction quality for $n = 95$ Fab structures. (**A**) Scatterplot of elbow angle (degrees) data in native versus predicted Fab structures. (**B**) Absolute elbow angle difference in degrees between predicted and native Fabs in correlation with prediction quality (RMS_CA). (**C**) Boxplot of RMS_CA (Å) data displaying the distance between superimposed native and predicted molecules for the Fab structures for two groups of Fabs, with elbow angle difference between native and predicted of <20 degrees and ≥20 degrees. The bottom and the top whiskers represent the minimum and the maximum value in the dataset excluding any outliers, the bottom and the top of the box are the 25th (Q1) and 75th (Q3) percentiles, the line inside the box is the 50th percentile (Q2) or the median value in the dataset, the x inside the box is a mean value of the dataset.

Prediction speed and technical requirements

One of the challenges faced while working with AlphaFold version 2.2.0 was a technical graphics processing unit requirement. Given that the graphics processing unit requirements are met, it could take up to 30 min to predict a 216-residue-long molecule. To overcome these shortcomings, we tested several alternative approaches. One approach was to use ColabFold (70), a web server that permits AlphaFold to run faster online (e.g., it took us up to 10 min to run the prediction based on a 216-residue sequence, as opposed to 30 min using the full version of AlphaFold version 2.2.0). In addition to ColabFold online, we also tested ABodyBuilder2 (ImmuneBuilder) (71), which is another fast

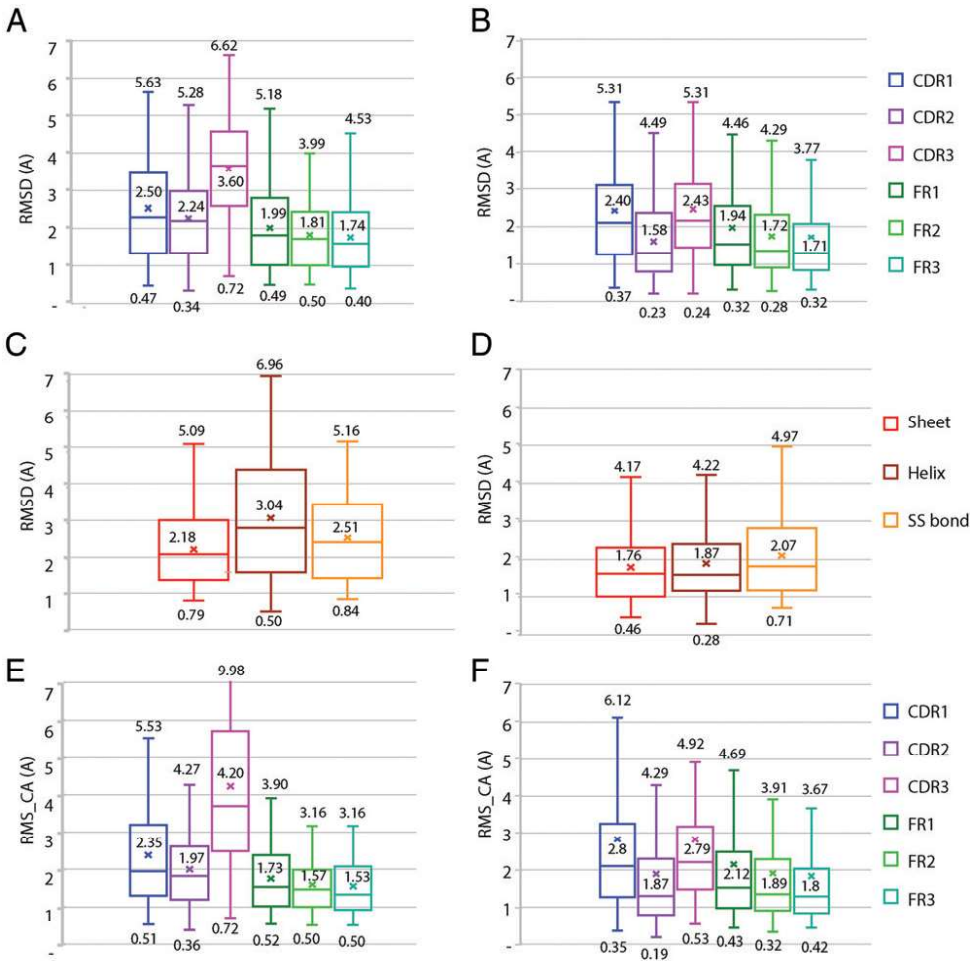

FIGURE 5. Prediction quality evaluation by subregion and by secondary structure. Boxplot of data displaying the distance between superimposed native and predicted molecules by subregion (CDR1/2/3, FR1/2/3) and by secondary structure: RMSD (Å) by CDRs and FRs in Fab (**A**) H chain ($n = 110$) and (**B**) L chain ($n = 112$); RMSD (Å) by secondary structure in Fab (**C**) H chain ($n = 110$) and (**D**) L chain ($n = 112$); RMS_CA (Å) by CDRs and FRs in Fab (**E**) H chain ($n = 110$) and (**F**) L chain ($n = 112$). All superimpositions were sequence independent in (A)–(D) and sequence dependent in (E) and (F). The bottom and the top whiskers represent the minimum and the maximum value in the dataset excluding any outliers, the bottom and the top of the box are the 25th (Q1) and 75th (Q3) percentiles, the line inside the box is the 50th percentile (Q2) or the median value in the dataset, the x inside the box is a mean value of the dataset.

Table V. Twenty chains' prediction quality comparison between ESMFold and AlphaFold: RMSD (Å) between superimposed predicted and native molecule

PDB Identifier and Chain	ESMFold Prediction Time, s	ESMFold RMSD, Å	ESMFold TM Score	AlphaFold, RMSD, Å	AlphaFold TM Score
7DOH_H	5	1.87	0.90	3.83	0.71
7DOH_L	10	3.14	0.77	0.87	0.98
7L77_H	6	3.75	0.73	2.68	0.82
7L77_L	8	0.91	0.98	2.01	0.89
7O31_H	10	3.17	0.80	1.55	0.95
7O31_L	10	4.58	0.63	1.90	0.90
7OX3_A	7	2.33	0.86	2.34	0.86
7OX3_B	15	0.83	0.98	3.08	0.78
7RM3_A	9	2.41	0.86	1.50	0.94
7RM3_B	12	1.29	0.95	1.85	0.91
8CZ5_H	12	3.49	0.72	4.04	0.68
8CZ5_L	6	2.73	0.81	1.31	0.95
8D47_H	7	1.88	0.88	3.01	0.83
8D47_L	6	1.57	0.93	1.27	0.95
8DFH_H	7	4.77	0.61	2.02	0.58
8DFH_L	6	2.33	0.52	1.55	0.93
8DTX_H	6	1.52	0.94	3.28	0.75
8DTX_L	6	2.49	0.84	1.20	0.97
8C3V_H	21	2.24	0.89	1.54	0.94
8C3V_L	5	0.89	0.97	2.88	0.79

deep learning model for Ab Fv region prediction. ABodyBuilder2 required ~30 s to predict a single Ab Fv region, and the prediction accuracy for 10 tested molecules was similar to that of ColabFold online for the pruned atom pairs and somewhat better across all atom pairs (Supplemental Table I). In October 2022, Meta AI released an alternative, faster, and more accessible solution as part of its sequence-to-structure predictor ESMFold and the ESM Metagenomic Atlas database. This claimed to be 60-fold faster than state-of-the-art predictions while maintaining resolution and accuracy (72). The Atlas serves as an open database of 617 million predicted protein structures and allows a rapid prediction to be obtained from a sequence. When we applied this option to 222 Ab chains with a mean amino acid sequence length of 221 residues, the running time was even faster, with an average of 9.3 s required to predict a single structure (Table V). Moreover, more than 85% of the tested Ab sequences were predicted in under 13 s (Table VI). When we evaluated the prediction quality of these test molecules compared with AlphaFold predictions, the mean value of RMSD was 2.53 Å for ESMFold and 2.38 Å for AlphaFold. These differences in accuracy are not statistically significant (Fig. 7). For both algorithms, most of the predictions fall into the RMSD range of 1–2 Å, whereas the AlphaFold results had a higher number of predictions in the range under 2 Å (Fig. 7B). Overall, we can conclude that the prediction quality of ESMFold is very similar to that of AlphaFold and that the shorter time required by ESMFold provides an accessibility advantage.

Discussion

Understanding the Ab structure and subsequently the Ab–Ag interactions provides a basis for the rational design of vaccines and drugs. The gold standard for structure determination remains X-ray crystallography, which provides high-resolution atomic coordinates of the protein backbone and side chains, as well as the atomic interactions between the different protein groups. However, this method requires a very high degree of expertise, tends to be expensive in terms of time and money, and is not always applicable. For these reasons, a variety of methods have been developed to predict the Ab structure from the coding sequence alone.

Table VI. ESMFold 222 Ab chains prediction time distribution

Prediction Time (s)	Number of Predictions	% of Total
5–7	95	42.8
8–10	50	22.5
11–13	44	19.8
14–16	30	13.5
17+	3	1.4
Total	222	100

Computational methods for protein structure prediction from the amino acid sequence have been available since CASP1 was launched in 1994. Earlier computational methods were mostly homology based, but, although they made it possible to predict the 3D structures of simple small proteins such as PDB identifier 2JZ5:A, with a known template with a very high sequence similarity, they were not successful in predicting the structures of larger and more conformationally challenging molecules such as Abs (73, 74). In addition, although more advanced protein prediction methods demonstrated a considerable degree of improvement, they were still inaccurate when analyzing large protein molecules and were therefore generally not applicable for deciphering complex molecular interactions and even less applicable for drug or vaccine design (75–77). The advent of artificial intelligence and development of new machine learning methods provided a significant leap in the ability to predict protein structures. In 2020, AlphaFold demonstrated an improved ability to predict protein structures with a median GDT_{TS} score of 0.92. Although the analysis included many complex and structurally challenging molecules, the prediction of Ab structures and interactions has not been thoroughly evaluated. Although a number of AlphaFold protein predictions have been reported over the past few years (78–80), these did not address the specific issues of Abs. In addition, most of these studies lacked local quality assessment of the predictions and did not try to estimate the reliability of the different subdomains of the predicted Ab or identify the areas that were the most difficult to predict.

Here, we evaluated the ability of AlphaFold to predict Ab 3D structures from a coding linear sequence by employing an algorithmic pipeline that compares the 3D structure of the predicted Ab chain molecule with the native X-ray crystallography structure. Importantly, the system scores the structural subdomains as well as the whole molecule. This methodology enabled a rapid comparison of 222 predicted structures with the corresponding published PDB atomic structures. The results allowed us to identify the problematic areas in the

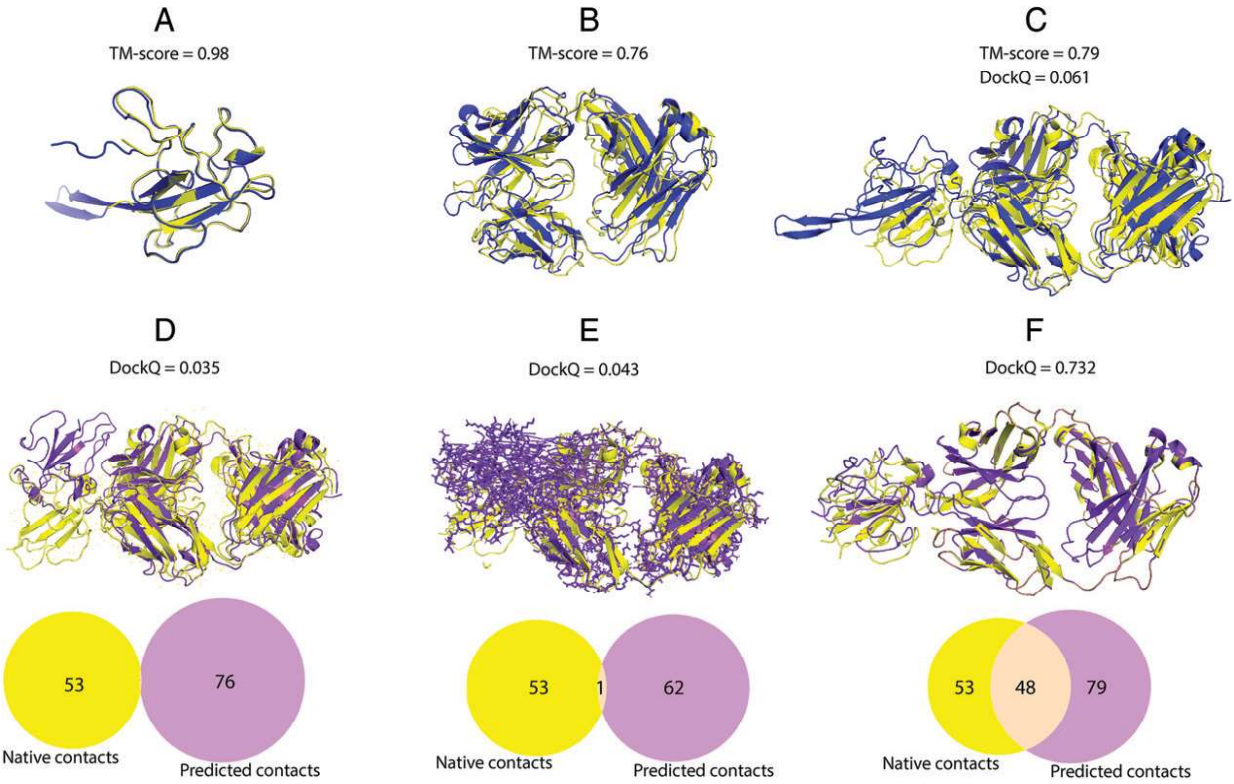

FIGURE 6. Ab–Ag multimer complex prediction and docking comparison example (PDB identifier 7N3D). (**A**) Native (yellow) Ag 3D structure superimposed on the AlphaFold prediction (blue). (**B**) Native (yellow) Ab Fab 3D structure superimposed on the AlphaFold-Multimer prediction (blue). (**C**) Ab–Ag complex native (yellow) 3D structure superimposed on the AlphaFold-Multimer prediction (blue). (**D**) Ag native structure docked by ZDOCK with Ab Fab structure predicted by AlphaFold (violet) superimposed on the native complex (yellow). (**E**) Ag and Ab Fab structures predicted by AlphaFold docked by ZDOCK (violet) superimposed on the native complex (yellow). (**F**) Ag native structure docked by ZDOCK with the Ab Fab native structure (violet) superimposed on the native complex (yellow). Venn diagrams in (D–F) show the number of docking contacts in native and predicted structures and their overlap, which represents the contacts predicted accurately.

predicted molecules, which can then be further analyzed and used to refine the models. We also predicted the 222 Ab chains by ESMFold released by Meta AI and compared the performance with the AlphaFold results.

Our study clearly demonstrates that low prediction quality often stems from an erroneous prediction of the elbow angle, with no major issue in VH-VL orientation prediction. Interestingly, there is some evidence in the literature that Ab elbow angles are influenced by the L chain class and that Fabs with λ chains have a wider range of angles (62). Moreover, various structural studies have reported that elbow angles are often altered by conformational changes that occur upon binding and give rise to significant differences between the elbow angles of the bound versus unbound Fabs (62, 81, 82). Notably, the nature of the Ag can affect the elbow bending angles within a given Fab, such as the elbow angle of anti-HIV-1 V3 Fab 2219 changes from 210.4 degrees upon binding to UG1033 peptide (PDB identifier 2B1A) to 229.4 degrees upon binding to MN peptide (PDB identifier 2B0S). Clearly, reliable prediction of the elbow angle in bound and unbound Abs is a challenge that computational tools still have to address.

Docking by both ZDOCK and AlphaFold-Multimer predicted the structure of Ab–Ag complexes relatively poorly, with ZDOCK

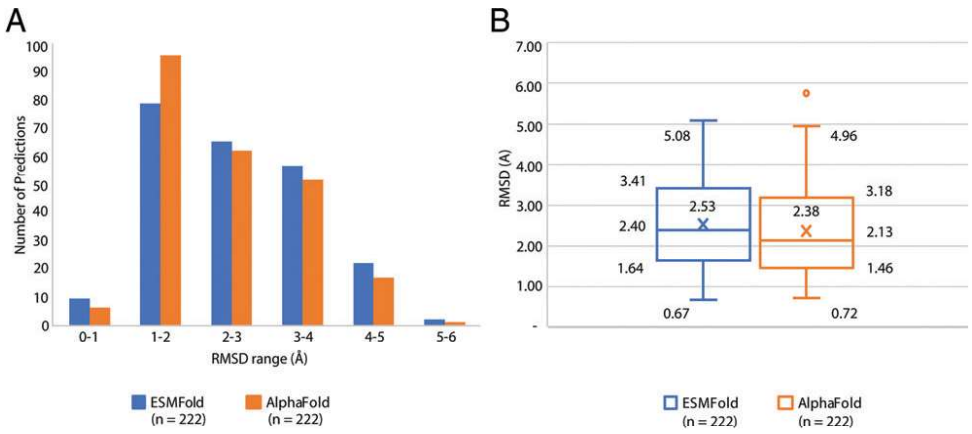

FIGURE 7. Distribution of 222 predictions by ESMFold and AlphaFold arranged by prediction quality, RMSD (Å). (**A**) Column chart of the prediction distribution by RMSD range (Å) for ESMFold in blue and for AlphaFold in orange. (**B**) Boxplot of RMSD (Å) data for 222 different predictions by ESMFold and AlphaFold. The bottom and the top whiskers represent the minimum and the maximum value in the dataset excluding any outliers, the bottom and the top of the box are the 25th (Q1) and 75th (Q3) percentiles, the line inside the box is the 50th percentile (Q2) or the median value in the dataset, the x inside the box is a mean value of the dataset.

succeeding in a higher number of complexes. We note several limitations in our comparison. First, the input for the two alternatives is different, where the input for the viable docking option is the native structures of both the Ag and the Ab, whereas the AlphaFold prediction is based solely on the amino acid sequences. Second, only 26 complexes were used for the comparison. It is important to note that all the complexes analyzed were released after the AlphaFold training cutoff date to avoid bias. Nevertheless, our study of 26 different complexes offers a sneak peek at the current level of prediction accuracy expected of Ab–Ag complex structures.

The main conclusion of our study is that despite significant advancement in the ability to predict protein structures, the current methods are not very accurate at modeling the 3D structure of Abs. This is predominantly due to low accuracy in prediction of the variable domains of the H chains, difficulty in predicting the CDRs (in both H and L chains), and inability to model the elbow angle between the constant and variable domains correctly. Any methodology for predicting Ab structures should specifically address these hurdles. Future improvements in methodology should also improve computationally based epitope mapping algorithms.

Acknowledgments

We thank members of the Freund laboratory and the Pupko laboratory for help and discussions. K.P. is grateful to Michael Polonsky for technical advice and support throughout the research process. We thank the Edmond J. Safra Center for Bioinformatics at Tel Aviv University Fellowship to K.P.

Disclosures

The authors have no financial conflicts of interest.

References

1. Plotkin, S. A. 2010. Correlates of protection induced by vaccination. *Clin. Vaccine Immunol.* 17: 1055–1065.
2. Trombetta, C. M., and E. Montomoli. 2016. Influenza immunology evaluation and correlates of protection: a focus on vaccines. *Expert Rev. Vaccines* 15: 967–976.
3. Pantaleo, G., and R. A. Koup. 2004. Correlates of immune protection in HIV-1 infection: what we know, what we don't know, what we should know. *Nat. Med.* 10: 806–810.
4. Wei, J., K. B. Pouwels, N. Stoesser, P. C. Matthews, I. Diamond, R. Studley, E. Rourke, D. Cook, J. I. Bell, J. N. Newton, et al.; COVID-19 Infection Survey team. 2022. Antibody responses and correlates of protection in the general population after two doses of the ChAdOx1 or BNT162b2 vaccines. *Nat. Med.* 28: 1072–1082.
5. Forthal, D. N. 2014. Functions of antibodies. *Microbiol. Spectr.* 2: AID-0019-2014.
6. Victora, G. D., and M. C. Nussenzweig. 2022. Germinal centers. *Annu. Rev. Immunol.* 40: 413–442.
7. Pento, J. T. 2017. Monoclonal antibodies for the treatment of cancer. *Anticancer Res.* 37: 5935–5939.
8. Hafeez, U., H. K. Gan, and A. M. Scott. 2018. Monoclonal antibodies as immunomodulatory therapy against cancer and autoimmune diseases. *Curr. Opin. Pharmacol.* 41: 114–121.
9. Rogers, T. F., F. Zhao, D. Huang, N. Beutler, A. Burns, W. T. He, O. Limbo, C. Smith, G. Song, J. Woehl, et al. 2020. Isolation of potent SARS-CoV-2 neutralizing antibodies and protection from disease in a small animal model. *Science* 369: 956–963.
10. Freund, N. T., J. F. Scheid, H. Mouquet, and M. C. Nussenzweig. 2015. Amplification of highly mutated human Ig lambda light chains from an HIV-1 infected patient. *J. Immunol. Methods* 418: 61–65.
11. Wardemann, H., S. Yurasov, A. Schaefer, J. W. Young, E. Meffre, and M. C. Nussenzweig. 2003. Predominant autoantibody production by early human B cell precursors. *Science* 301: 1374–1377.
12. Gershoni, J. M., A. Roitburd-Berman, D. D. Siman-Tov, N. Tarnovitski Freund, and Y. Weiss. 2007. Epitope mapping: the first step in developing epitope-based vaccines. *BioDrugs* 21: 145–156.
13. Steichen, J. M., D. W. Kulp, T. Tokatlian, A. Escolano, P. Dosenovic, R. L. Stanfield, L. E. McCoy, G. Ozorowski, X. Hu, O. Kalyuzhniy, et al. 2016. HIV vaccine design to target germline precursors of glycan-dependent broadly neutralizing antibodies. *Immunity* 45: 483–496.
14. Barnes, C. O., H. B. Gristick, N. T. Freund, A. Escolano, A. Y. Lyubimov, H. Hartweger, A. P. West, Jr., A. E. Cohen, M. C. Nussenzweig, and P. J. Bjorkman. 2018. Structural characterization of a highly-potent V3-glycan broadly neutralizing antibody bound to natively-glycosylated HIV-1 envelope. *Nat. Commun.* 9: 1251.
15. Watson, A., H. Li, B. Ma, R. Weiss, D. Bendayan, L. Abramovitz, N. Ben-Shalom, M. Mor, E. Pinko, M. Bar Oz, et al. 2021. Human antibodies targeting a *Mycobacterium* transporter protein mediate protection against tuberculosis. *Nat. Commun.* 12: 602.
16. Li, R., M. Mor, B. Ma, A. E. Clark, J. Alter, M. Werbner, J. C. Lee, S. L. Leibel, A. F. Carlin, M. Dessau, et al. 2022. Conformational flexibility in neutralization of SARS-CoV-2 by naturally elicited anti-SARS-CoV-2 antibodies. *Commun. Biol.* 5: 789.
17. Pagadala, N. S., K. Syed, and J. Tuszynski. 2017. Software for molecular docking: a review. *Biophys. Rev.* 9: 91–102.
18. Agrawal, P., H. Singh, H. K. Srivastava, S. Singh, G. Kishore, and G. P. S. Raghava. 2019. Benchmarking of different molecular docking methods for protein-peptide docking. *BMC Bioinformatics* 19(Suppl 13): 426.
19. Sanchez-Trincado, J. L., M. Gomez-Perosanz, and P. A. Reche. 2017. Fundamentals and methods for T- and B-cell epitope prediction. *J. Immunol. Res.* 2017: 2680160.
20. Ehlers, A. M., M. A. Blankestijn, A. C. Knulst, M. Klinge, and H. G. Otten. 2019. Can alternative epitope mapping approaches increase the impact of B-cell epitopes in food allergy diagnostics? *Clin. Exp. Allergy* 49: 17–26.
21. Backert, L., and O. Kohlbacher. 2015. Immunoinformatics and epitope prediction in the age of genomic medicine. *Genome Med.* 7: 119.
22. Krawczyk, K., J. Dunbar, and C. M. Deane. 2017. Computational tools for aiding rational antibody design. *Methods Mol. Biol.* 1529: 399–416.
23. Guest, J. D., T. Vreven, J. Zhou, I. Moal, J. R. Jeliazkov, J. J. Gray, Z. Weng, and B. G. Pierce. 2021. An expanded benchmark for antibody-antigen docking and affinity prediction reveals insights into antibody recognition determinants. *Structure* 29: 606–621.e5.
24. Norman, R. A., F. Ambrosetti, A. M. J. J. Bonvin, L. J. Colwell, S. Kelm, S. Kumar, and K. Krawczyk. 2020. Computational approaches to therapeutic antibody design: established methods and emerging trends. *Brief. Bioinform.* 21: 1549–1567.
25. Chiu, M. L., D. R. Goulet, A. Teplyakov, and G. L. Gilliland. 2019. Antibody structure and function: the basis for engineering therapeutics. *Antibodies (Basel)* 8: 55.
26. Kryshtafovych, A., T. Schwede, M. Topf, K. Fidelis, and J. Moult. 2021. Critical assessment of methods of protein structure prediction (CASP)—round XIV. *Proteins* 89: 1607–1617.
27. Jumper, J., R. Evans, A. Pritzel, T. Green, M. Figurnov, O. Ronneberger, K. Tunyasuvunakool, R. Bates, A. Žídek, A. Potapenko, et al. 2021. Highly accurate protein structure prediction with AlphaFold. *Nature* 596: 583–589.
28. The AlphaFold team. 2020. AlphaFold: a solution to a 50-year-old grand challenge in biology. Available at: https://www.deepmind.com/blog/alphafold-a-solution-to-a-50-year-old-grand-challenge-in-biology. Accessed: January 15, 2022.
29. Jumper, J., R. Evans, A. Pritzel, T. Green, M. Figurnov, O. Ronneberger, K. Tunyasuvunakool, R. Bates, A. Žídek, A. Potapenko, et al. 2021. Applying and improving AlphaFold at CASP14. *Proteins* 89: 1711–1721.
30. Simpkin, A. J., F. Sánchez Rodríguez, S. Mesdaghi, A. Kryshtafovych, and D. J. Rigden. 2021. Evaluation of model refinement in CASP14. *Proteins* 89: 1852–1869.
31. Kinch, L. N., R. D. Schaeffer, A. Kryshtafovych, and N. V. Grishin. 2021. Target classification in the 14th round of the critical assessment of protein structure prediction (CASP14). *Proteins* 89: 1618–1632.
32. Scheid, J. F., H. Mouquet, B. Ueberheide, R. Diskin, F. Klein, T. Y. K. Oliveira, J. Pietzsch, D. Fenyo, A. Abadir, K. Velinzon, et al. 2011. Sequence and structural convergence of broad and potent HIV antibodies that mimic CD4 binding. *Science* 333: 1633–1637.
33. Eisen, H. N., and G. W. Siskind. 1964. Variations in affinities of antibodies during the immune response. *Biochemistry* 3: 996–1008.
34. Peters, A., and U. Storb. 1996. Somatic hypermutation of immunoglobulin genes is linked to transcription initiation. *Immunity* 4: 57–65.
35. Xu, J. L., and M. M. Davis. 2000. Diversity in the CDR3 region of V_H is sufficient for most antibody specificities. *Immunity* 13: 37–45.
36. Pierce, B. G., K. Wiehe, H. Hwang, B. H. Kim, T. Vreven, and Z. Weng. 2014. ZDOCK server: interactive docking prediction of protein-protein complexes and symmetric multimers. *Bioinformatics* 30: 1771–1773.
37. Berman, H. M., T. Battistuz, T. N. Bhat, W. F. Bluhm, P. E. Bourne, K. Burkhardt, Z. Feng, G. L. Gilliland, L. Iype, S. Jain, et al. 2002. The Protein Data Bank. *Acta Crystallogr. D Biol. Crystallogr.* 58: 899–907.
38. Dunbar, J., K. Krawczyk, J. Leem, T. Baker, A. Fuchs, G. Georges, J. Shi, and C. M. Deane. 2014. SAbDab: the structural antibody database. *Nucleic Acids Res.* 42(D1): D1140–D1146.
39. Rubeira CO. 2021. AlphaFold 2 is here: what's behind the structure prediction miracle. Oxford Protein Informatics Group. Available at: https://www.blopig.com/blog/2021/07/alphafold-2-is-here-whats-behind-the-structure-prediction-miracle/. Accessed: January 15, 2022.
40. Steinegger, M., M. Mirdita, and J. Söding. 2019. Protein-level assembly increases protein sequence recovery from metagenomic samples manyfold. *Nat. Methods* 16: 603–606.
41. Steinegger, M., and J. Söding. 2018. Clustering huge protein sequence sets in linear time. *Nat. Commun.* 9: 2542.
42. Varadi, M., S. Anyango, M. Deshpande, S. Nair, C. Natassia, G. Yordanova, D. Yuan, O. Stroe, G. Wood, A. Laydon, et al. 2022. AlphaFold Protein Structure Database: massively expanding the structural coverage of protein-sequence space with high-accuracy models. *Nucleic Acids Res.* 50(D1): D439–D444.
43. Dondelinger, M., P. Filée, E. Sauvage, B. Quinting, S. Muyldermans, M. Galleni, and M. S. Vandevenne. 2018. Understanding the significance and implications of antibody numbering and antigen-binding surface/residue definition. *Front. Immunol.* 9: 2278.
44. Kabat, E. A., and T. T. Wu. 1971. Attempts to locate complementarity-determining residues in the variable positions of light and heavy chains. *Ann. N. Y. Acad. Sci.* 190: 382–393.

45. Chothia, C., and A. M. Lesk. 1987. Canonical structures for the hypervariable regions of immunoglobulins. *J. Mol. Biol.* 196: 901–917.
46. Chothia, C., A. M. Lesk, A. Tramontano, M. Levitt, S. J. Smith-Gill, G. Air, S. Sheriff, E. A. Padlan, D. Davies, W. R. Tulip, et al. 1989. Conformations of immunoglobulin hypervariable regions. *Nature* 342: 877–883.
47. Ehrenmann, F., and M. P. Lefranc. 2012. IMGT/DomainGapAlign: the IMGT tool for the analysis of IG, TR, MH, IgSF, and MhSF domain amino acid polymorphism. *Methods Mol. Biol.* 882: 605–633.
48. Ye, J., N. Ma, T. L. Madden, and J. M. Ostell. 2013. IgBLAST: an immunoglobulin variable domain sequence analysis tool. *Nucleic Acids Res.* 41(W1): W34–W40.
49. Lefranc, M.-P., and G. Lefranc. 2001. *The Immunoglobulin FactsBook*, 1st Ed. Academic Press, San Diego, CA.
50. Brochet, X., M. P. Lefranc, and V. Giudicelli. 2008. IMGT/V-QUEST: the highly customized and integrated system for IG and TR standardized V-J and V-D-J sequence analysis. *Nucleic Acids Res.* 36(Web Server): W503–W508.
51. Soto, C., J. A. Finn, J. R. Willis, S. B. Day, R. S. Sinkovits, T. Jones, S. Schmitz, J. Meiler, A. Branchizio, and J. E. Crowe, Jr. 2020. PyIR: a scalable wrapper for processing billions of immunoglobulin and T cell receptor sequences using IgBLAST. *BMC Bioinformatics* 21: 314.
52. Strohl, W. R., and L. M. Strohl. 2012. Antibody structure–function relationships. In *Therapeutic Antibody Engineering.* Elsevier Science, Philadelphia, p. 37–595.
53. Pantazes, R. J., and C. D. Maranas. 2010. OptCDR: a general computational method for the design of antibody complementarity determining regions for targeted epitope binding. *Protein Eng. Des. Sel.* 23: 849–858.
54. Ofran, Y., A. Schlessinger, and B. Rost. 2008. Automated identification of complementarity determining regions (CDRs) reveals peculiar characteristics of CDRs and B cell epitopes. *J. Immunol.* 181: 6230–6235.
55. Zhang, Y., and J. Skolnick. 2005. TM-align: a protein structure alignment algorithm based on the TM-score. *Nucleic Acids Res.* 33: 2302–2309.
56. Mukherjee, S., and Y. Zhang. 2009. MM-align: a quick algorithm for aligning multiple-chain protein complex structures using iterative dynamic programming. *Nucleic Acids Res.* 37: e83.
57. Meng, E. C., E. F. Pettersen, G. S. Couch, C. C. Huang, and T. E. Ferrin. 2006. Tools for integrated sequence-structure analysis with UCSF Chimera. *BMC Bioinformatics* 7: 339.
58. Pettersen, E. F., T. D. Goddard, C. C. Huang, G. S. Couch, D. M. Greenblatt, E. C. Meng, and T. E. Ferrin. 2004. UCSF Chimera—a visualization system for exploratory research and analysis. *J. Comput. Chem.* 25: 1605–1612.
59. Zemla, A. 2003. LGA: a method for finding 3D similarities in protein structures. *Nucleic Acids Res.* 31: 3370–3374.
60. Evans, R., M. O'Neill, A. Pritzel, N. Antropova, A. Senior, T. Green, A. Žídek, R. Bates, S. Blackwell, J. Yim, et al. 2022. Protein complex prediction with AlphaFold-Multimer. *bioRxiv* 2021.10.04.463034.
61. Tunyasuvunakool, K., J. Adler, Z. Wu, T. Green, M. Zielinski, A. Žídek, A. Bridgland, A. Cowie, C. Meyer, A. Laydon, et al. 2021. Highly accurate protein structure prediction for the human proteome. *Nature* 596: 590–596.
62. Stanfield, R. L., A. Zemla, I. A. Wilson, and B. Rupp. 2006. Antibody elbow angles are influenced by their light chain class. *J. Mol. Biol.* 357: 1566–1574.
63. Barozet, A., M. Bianciotto, T. Siméon, H. Minoux, and J. Cortés. 2018. Conformational changes in antibody Fab fragments upon binding and their consequences on the performance of docking algorithms. *Immunol. Lett.* 200: 5–15.
64. Sandin, S., L. G. Öfverstedt, A. C. Wikström, O. Wrange, and U. Skoglund. 2004. Structure and flexibility of individual immunoglobulin G molecules in solution. *Structure* 12: 409–415.
65. Dunbar, J., A. Fuchs, J. Shi, and C. M. Deane. 2013. ABangle: characterising the VH-VL orientation in antibodies. *Protein Eng. Des. Sel.* 26: 611–620.
66. Basu, S., and B. Wallner. 2016. DockQ: a quality measure for protein-protein docking models. *PLoS One* 11: e0161879.
67. Schrödinger. 2021. The PyMOL Molecular Graphics System. Version 2.6.
68. Callaway, E. 2020. 'It will change everything': DeepMind's AI makes gigantic leap in solving protein structures. *Nature* 588: 203–204.
69. Barnes, C. O., T. Schoofs, P. N. P. Gnanapragasam, J. Golijanin, K. E. Huey-Tubman, H. Gruell, P. Schommers, N. Suh-Toma, Y. E. Lee, J. C. Cetrulo Lorenzi, et al. 2022. A naturally arising broad and potent CD4-binding site antibody with low somatic mutation. *Sci. Adv.* 8: eabp8155.
70. Mirdita, M., K. Schütze, Y. Moriwaki, L. Heo, S. Ovchinnikov, and M. Steinegger. 2022. ColabFold: making protein folding accessible to all. *Nat. Methods* 19: 679–682.
71. Abanades, B., W. K. Wong, F. Boyles, G. Georges, A. Bujotzek, and C. M. Deane. 2023. ImmuneBuilder: deep-learning models for predicting the structures of immune proteins. *Commun. Biol.* 6: 575.
72. Lin, Z., H. Akin, R. Rao, B. Hie, Z. Zhu, W. Lu, N. Smetanin, R. Verkuil, O. Kabeli, Y. Shmueli, et al. 2023. Evolutionary-scale prediction of atomic-level protein structure with a language model. *Science* 379: 1123–1130.
73. Wu, S., A. Szilagyi, and Y. Zhang. 2011. Improving protein structure prediction using multiple sequence-based contact predictions. *Structure* 19: 1182–1191.
74. Yang, J., R. Yan, A. Roy, D. Xu, J. Poisson, and Y. Zhang. 2015. The I-TASSER Suite: protein structure and function prediction. *Nat. Methods* 12: 7–8.
75. Kuhlman, B., and P. Bradley. 2019. Advances in protein structure prediction and design. *Nat. Rev. Mol. Cell Biol.* 20: 681–697.
76. Pearce, R., and Y. Zhang. 2021. Deep learning techniques have significantly impacted protein structure prediction and protein design. *Curr. Opin. Struct. Biol.* 68: 194–207.
77. Yang, J., I. Anishchenko, H. Park, Z. Peng, S. Ovchinnikov, and D. Baker. 2020. Improved protein structure prediction using predicted interresidue orientations. *Proc. Natl. Acad. Sci. USA* 117: 1496–1503.
78. Yin, R., B. Y. Feng, A. Varshney, and B. G. Pierce. 2022. Benchmarking AlphaFold for protein complex modeling reveals accuracy determinants. *Protein Sci.* 31: e4379.
79. Saldaño, T., N. Escobedo, J. Marchetti, D. J. Zea, J. Mac Donagh, A. J. Velez Rueda, E. Gonik, A. García Melani, J. Novomisky Nechcoff, M. N. Salas, et al. 2022. Impact of protein conformational diversity on AlphaFold predictions. *Bioinformatics* 38: 2742–2748.
80. Johansson-Åkhe, I., and B. Wallner. 2022. Improving peptide-protein docking with AlphaFold-Multimer using forced sampling. *Front Bioinform* 2: 959160.
81. Ponomarenko, N., S. D. Chatziefthimiou, I. Kurkova, Y. Mokrushina, Y. Mokrushina, A. Stepanova, I. Smirnov, M. Avakyan, T. Bobik, A. Mamedov, et al. 2014. Role of κ→λ light-chain constant-domain switch in the structure and functionality of A17 reactibody. *Acta Crystallogr. D Biol. Crystallogr.* 70: 708–719.
82. Sela-Culang, I., S. Alon, and Y. Ofran. 2012. A systematic comparison of free and bound antibodies reveals binding-related conformational changes. *J. Immunol.* 189: 4890–4899.