Microbiology Spectrum
American Society for Microbiology Press

Antibodies: Computer-Aided Prediction of Structure and Design of Function

ALEXANDER M. SEVY[1] and JENS MEILER[1]
[1]Department of Chemistry and Center for Structural Biology, Vanderbilt University, Nashville, TN 37212

ABSTRACT With the advent of high-throughput sequencing, and the increased availability of experimental structures of antibodies and antibody-antigen complexes, comes the improvement of computational approaches to predict the structure and design the function of antibodies and antibody-antigen complexes. While antibodies pose formidable challenges for protein structure prediction and design due to their large size and highly flexible loops in the complementarity-determining regions, they also offer exciting opportunities: the central importance of antibodies for human health results in a wealth of structural and sequence information that—as a knowledge base—can drive the modeling algorithms by limiting the conformational and sequence search space to likely regions of success. Further, efficient experimental platforms exist to test predicted antibody structure or designed antibody function, thereby leading to an iterative feedback loop between computation and experiment. We briefly review the history of computer-aided prediction of structure and design of function in the antibody field before we focus on recent methodological developments and the most exciting application examples.

INTRODUCTION

The central role antibodies play in our immune system makes them important targets for computation-based structural modeling. Antibodies consist of a "constant" and a "variable" region (Fig. 1). The constant region is virtually identical in all antibodies of the same isotype, while the variable region differs from one B-cell-derived antibody to the next. The variable region of an antibody is the "business end," the region that recognizes its antigen via so-called complementarity-determining regions (CDRs). Their large size (~150 kDa) and inherent variability, in particular in the CDRs, make antibodies a formidable challenge for molecular modeling. Before we begin to model antibodies, it is useful to briefly review their overall structure.

Relation of Antibody Sequence, Structure, and Function

The fundamental unit within an antibody is an immunoglobulin (IG) domain of around 70 to 110 amino acids that adopts the characteristic IG β-sandwich fold. Antibodies are homodimers of heterodimers, where each heterodimer consists of one heavy and one light chain (Fig. 1) (1), each chain having multiple IG domains. A mammalian antibody light chain consists of two IG domains, the C-terminal one called "constant" and the N-terminal one, "variable." The mammalian antibody heavy chain consists of four or five IG domains, the most N-terminal one being variable, and all others constant. The two N-terminal IG domains of the heavy chain interact with the two IG domains of a light chain to form heterodimers. These heterodimers homo-dimerize via the C-terminal IG domains of the heavy chain to form

Received: 10 September 2014, **Accepted:** 17 September 2014, **Published:** 19 December 2014
Editors: James E. Crowe, Jr., Vanderbilt University School of Medicine, Nashville, TN; Diana Boraschi, National Research Council, Pisa, Italy; and Rino Rappuoli, Novartis Vaccines, Siena, Italy
Citation: Sevy AM, Meiler J. 2014. Antibodies: computer-aided prediction of structure and design of function. *Microbiol Spectrum* 2(6):AID-0024-2014. doi:10.1128/microbiolspec.AID-0024-2014.
Correspondence: Jens Meiler, jens.meiler@vanderbilt.edu
© 2014 American Society for Microbiology. All rights reserved.

Downloaded from www.asmscience.org by
IP: 155.41.210.179
On: Thu, 31 May 2018 15:45:01

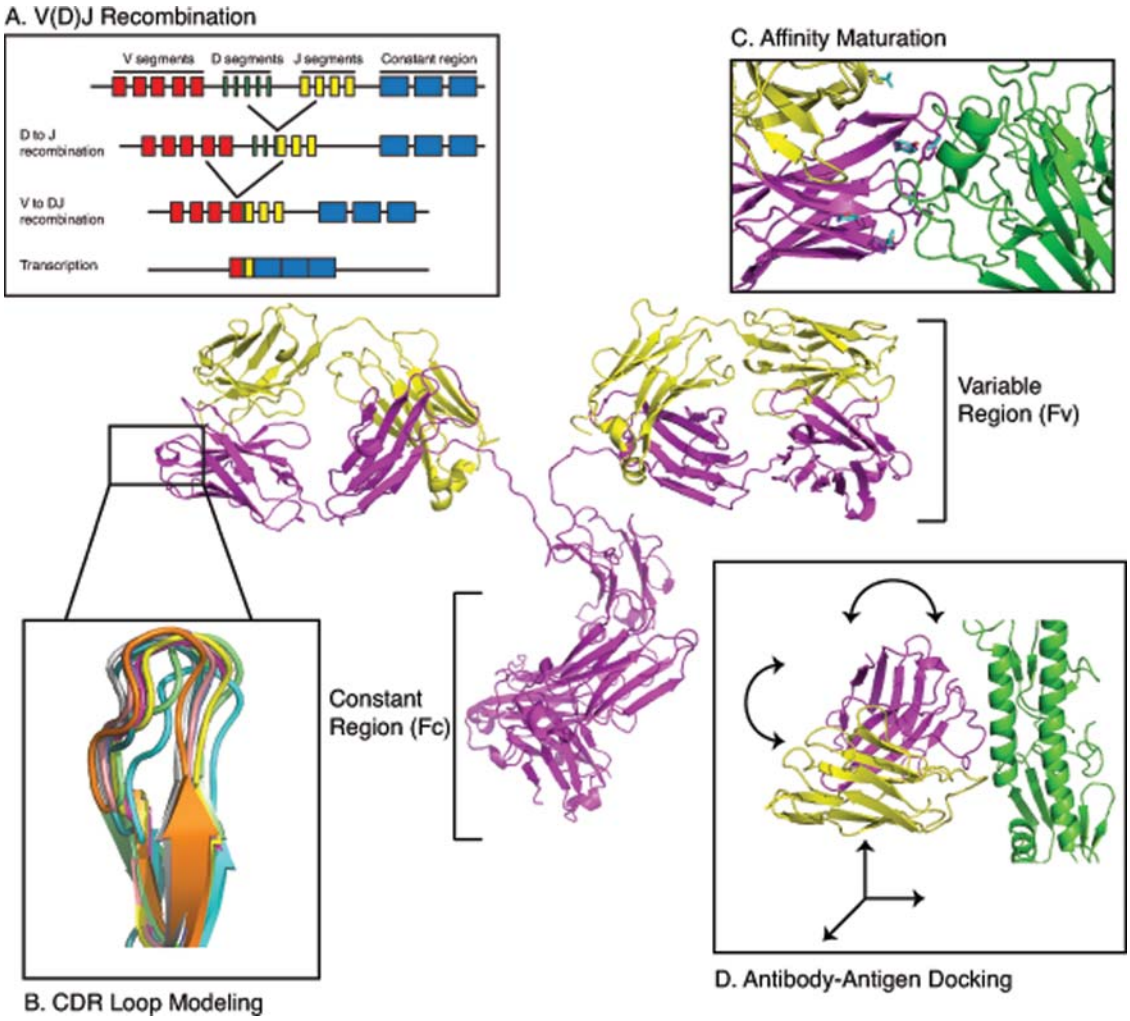


FIGURE 1 Challenges in antibody modeling. Though all antibodies share a common core structure (center panel, PDB ID 1IGT [1]; heavy chains in magenta, light chains in yellow), slight differences in variable regions and especially CDR loops can have a great effect on function. The vast sequence space generated by genetic recombination in V, D, and J genes (A) results in many different CDR loop conformations. Modeling of CDR loops from sequence information alone is a necessary computational task for accurate structure prediction (B). The ability to simulate the affinity maturation process *in silico* is another important task that can be used to generate an antibody with either increased higher affinity for its native target, or for a completely novel target (C) (matured residues shown in cyan). Accurate antibody modeling requires not only the ability to model an antibody alone, but also the ability to model its interaction with a given antigen. Computational docking techniques achieve this by sampling different positions of an antibody on its target to find the most favorable position (D). doi:10.1128/microbiolspec.AID-0024-2014.f1

the final antibody. This domain arrangement ensures that the variable domains of heavy and light chains co-localize in space to form the so-called paratope.

Each of these variable domains contains three CDRs that are the second (CDR1), fourth (CDR2), and sixth (CDR3) loop regions of the β-sandwich, locating the CDRs on the same end of the IG fold. The amino acid sequence within the CDRs is determined by a process called somatic recombination, where an IG domain is assembled by combining randomly chosen gene segments—V and J for the light chain and V, D, and J for the heavy chain (Fig. 1A). This process generates a large number of antibody sequences, as there are multiple copies of each gene type-somatic recombination combined with "junctional diversity" at the joints of the gene segments creates a theoretical limit of around

Downloaded from www.asmscience.org by
IP: 155.41.210.179
On: Thu, 31 May 2018 15:45:01

10^{11} unique V(D)J sequences. In particular, CDR3 of the heavy chain (HCDR3), encoded by the D-gene, is highly variable in length, structure, and dynamics (Fig. 1C). These "germline" antibodies are further modified in a process called affinity maturation (Fig. 1D). During B cell proliferation the genes encoding the variable domains experience an increased rate of point mutation. This "somatic hypermutation" causes amino acid changes in the paratopes of daughter B cells, a process that allows tighter interaction with the "epitope" region of the antigen, i.e., affinity maturation.

Motivations for Antibody Modeling

The large number of theoretically possible antibodies and the large number of antibodies actually present in humans prohibit a comprehensive experimental characterization of antibody structure and dynamics. While great progress has been made in antibody structure determination via crystallization (currently around 2,000 depositions in the Protein Data Bank [PDB] contain the phrase "antibody"), the number of experimental structures available in the PDB will always be small compared to the total immune repertoire, leaving room for structure prediction of important antibodies with unknown structure. As antibody structures in the PDB have increased exponentially in recent years (Fig. 2), computational biologists have gained a greater understanding of the molecular determinants of proper loop folding and antigen binding, ultimately allowing high-throughput, accurate structural modeling on a scale infeasible for experimental methods alone. Understanding the structural determinants of antibody-antigen interaction (i.e., how the paratope engages the epitope) is critical for understanding antibody function and processes such as affinity maturation. As the number of cocrystal structures of antibody and antigen will always be small, computational docking algorithms can provide models for antibody-antigen complexes that are not experimentally determined (Fig. 1D). Note that computational structure prediction usually does not replace the experiment but complements experimental data. For examples, starting from an experimental structure of antibody and antigen, their interaction could be modeled; starting from one antibody-antigen complex, models for affinity-matured antibody-antigen complexes can be constructed; or modeling can add atomic detail not present in lower-resolution electron density maps obtained, for example, through electron microscopy (EM). Computational design of antibody-antigen complexes can be used, for example, to study the process of affinity maturation *in silico* or identify antibodies with novel sequences not present or not yet observed in nature (Fig. 1B). Computational antibody engineering is also applied to humanize or stabilize therapeutic antibodies.

Challenges for Antibody Modeling

Challenges when predicting the structures of antibodies via comparative modeling include how changes in IG sequence change the relative orientation of the two variable IG domains in the complex and therefore influence the paratope structure. Obviously, modeling the conformation of CDR loops is a substantial challenge, in particular for CDR3, which has an average length of around 15 amino acids in humans and can be as long as 30 residues or more—well beyond the loop lengths tested in a typical loop construction benchmark (4, 8, and 12 amino acids). The plasticity of the paratope with the often flexible CDRs presents a formidable challenge to antibody/antigen docking simulations, as they require flexibility of both paratope and epitope, creating a huge conformational search space. When engineering antibodies, these challenges are multiplied by the enormous number of possible antibody sequences and the resulting gigantic size of the sequence space that needs to be sampled, in addition to the conformational space. These challenges result in another formidable motivation for modeling antibodies: the benchmarking of new computational techniques. The challenges related to modeling antibodies combined with the availability of many experimental structures make antibody structure prediction and functional design an important playground to test new algorithms.

COMPARATIVE MODELING OF ANTIBODIES

Given the large number of theoretically possible and actually existing antibodies, experimental structure determination will remain reserved to a small fraction of particularly important antibodies. Therefore, computational construction of a structural model is of central importance. A particular focus of comparative modeling techniques is accurate modeling of the CDRs.

Canonical Structure of CDRs

The concept of canonical structures of light CDR (LCDR) and heavy CDR (HCDR) loops can be traced back to seminal work by Chothia, Thornton, Lesk, and others in the 1980s and 1990s describing the conformational space sampled by all CDR loops with known structure and linking these to conserved residues among these sequences (2–4). Formally, the canonical structure hypothesis states that the CDR loops of antibodies typically

Downloaded from www.asmscience.org by
IP: 155.41.210.179
On: Thu, 31 May 2018 15:45:01

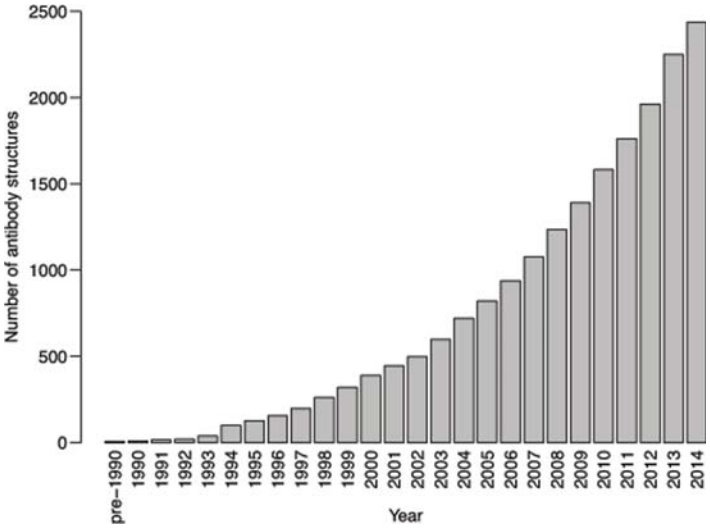


FIGURE 2 Exponential growth of structurally determined antibodies. Total antibody structures in the Protein Data Bank (PDB) are shown by year. The increase in structures enables more accurate computational approaches to antibody modeling and engineering. doi:10.1128/microbiolspec.AID-0024-2014.f2

adopt one of a discrete set of conformations and that these conformations can be inferred from the amino acid sequence. This concept has been pursued to discretize the conformational space of CDRs and predict a loop conformation based on its primary sequence. The set of canonical conformations has been relatively well defined for all LCDR loops and for the heavy HCDR1 and 2 loops. However, the HCDR3 loop is by far the most variable in sequence and length, drastically increasing the conformational space it can sample. As the number of experimentally determined antibody structures has exploded since the initial reports of canonical structures, attempts to group CDR loops and create a definitive set of loop conformations have continued to add new clusters to the known set. While it is likely that this trend will continue for some time, the fact that the number of clusters is still orders of magnitude lower than the known antibody repertoire validates the original canonical structure hypothesis by Chothia et al. for CDRs other than HCDR3 (2).

Studies of canonical non-HCDR3 loop conformations have focused on clustering known structures and deriving common characteristics in primary sequence to enable *a priori* prediction of a loop conformation based on sequence alone. The number of canonical loop conformations has increased along with experimental structures available for analysis; initial studies using only 17 structures identified 18 non-HCDR3 clusters (2), whereas more recent studies using ~1,200 structures have increased the number of non-HCDR3 clusters to 72 (5). Clusters are identified by their loop type and length, with the majority of non-HCDR3 loops falling between 8 and 13 residues in length. Though the number of non-HCDR3 loop clusters has increased with each subsequent analysis, the overall clustering pattern still maintains a high degree of uniformity and predictability. Figure 3 shows median loop structures of loops deriving from the largest cluster for each of the CDR L1, L2, L3, H1, and H2 loops (5). Studies have consistently reported 85 to 90% accuracy in predicting the structure of non-HCDR3 loops based on their gene source and primary sequence, lending credibility to the use of cluster analysis in non-HCDR3 loop modeling.

Predicting the Conformation of HCDR3

Prediction of HCDR3 loop conformations presents a considerably larger challenge, as they are much more structurally diverse and tend to be longer, ranging usually from 5 to 26 residues, with an average of 16 residues (6). In rare cases they can be substantially longer. HCDR3 loops are generally divided into "torso" and "head" regions for clustering purposes (7), and loops are characterized by either a bulged or extended beta-sheet conformation in the torso region. Since HCDR3 loops cannot typically be placed into a conformational cluster based on sequence alone, recent work has focused on developing a set of rules to predict certain aspects of the conformation based on key residue positions (8). The more stable bulged conformation tends to be preferred by HCDR3 loops, stabilized by a hydrogen bond between a conserved tryptophan and a backbone carbonyl. One rule dictates that an aspartate residue two positions upstream of the conserved tryptophan is sufficient to displace this hydrogen bond and results in a shift to the extended conformation. Another states that the position of a basic residue opposite this

Downloaded from www.asmscience.org by
IP: 155.41.210.179
On: Thu, 31 May 2018 15:45:01

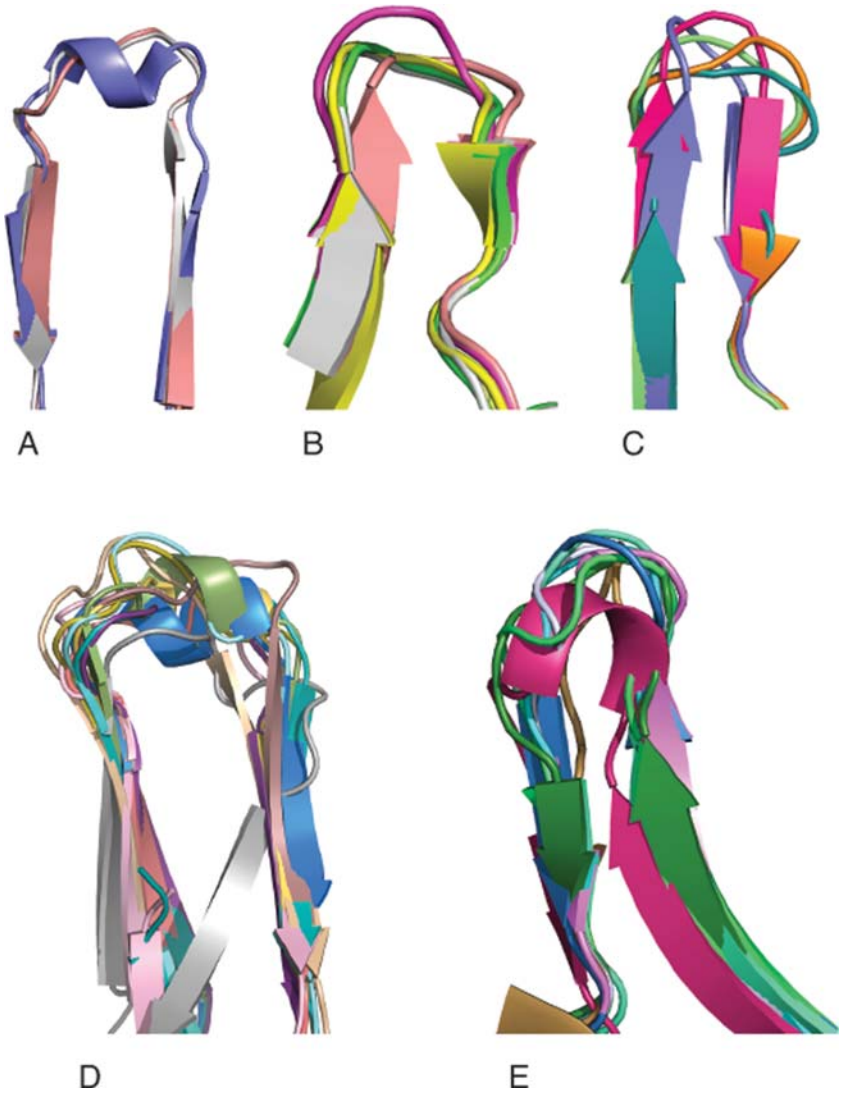


FIGURE 3 Canonical CDR loop conformations. Pictured above are median loop structures representing the largest cluster of (A) CDR L1, (B) L2, (C) L3, (D) H1, and (E) H2. Light and heavy chain loop variability varies widely between the CDR loops, with heavy chain loops tending to be more variable. doi:10.1128/microbiolspec.AID-0024-2014.f3

aspartate dictates the formation of a bulged or extended conformation. Taken together, these rules are able to correctly predict torso conformation with ~85% accuracy, comparable to that of non-HCDR3 loops.

Though the torso conformation of HCDR3 loops can be predicted by sequence analysis, the antigen-binding head region of the loop remains intractable to clustering, leading to considerable efforts to use *de novo* modeling to predict HCDR3 conformation. Currently, software packages such as Rosetta (9) and Prime (10) have been adapted to predict low-energy loop conformations, either based on peptide fragments gleaned from the PDB or generated *de novo*. These protocols have achieved varying degrees of success: HCDR3 loops shorter than 12 to 14 residues can be consistently predicted within a reasonable margin of error (~2 Å), while longer loops are less predictable and tend to have higher deviation. As the average HCDR3 loop is ~16 residues in length, current algorithms for *de novo* modeling remain insufficient to address a large proportion of important antibodies. However, the recent advances in predicting the conformation of the HCDR3 base via clustering and improvements in *de novo* loop construction promise to enable more reliable HCDR3 prediction in the near future.

Programs, Platforms, and Servers Dedicated to Antibody Modeling

Because of the large number of antibody-based therapeutics, there is substantial interest in determining the structure of antibodies in a high-throughput, accurate

Downloaded from www.asmscience.org by
IP: 155.41.210.179
On: Thu, 31 May 2018 15:45:01

manner. Experimental methods such as X-ray crystallography and nuclear magnetic resonance (NMR) spectroscopy, while highly informative, are labor-, time-, and resource-intensive. In addition, complete antibodies tend to be too large for NMR spectroscopy. X-ray crystallography, in the absence of the antigen, often struggles to determine the conformation of a long HCDR3 loop in the biologically relevant conformation needed to engage the antigen. Sometimes the HCDR3 loop can be locked into a different conformation by crystal packing or its coordinates cannot be determined because of flexibility. Therefore, many groups have worked on automated protocols to computationally model antibody variable region structures to predict the conformational space including the conformation needed for binding. This has resulted in a number of publicly accessible antibody-modeling servers, which use a combination of comparative modeling, *de novo* structure prediction, and energy minimization to generate an ensemble of potential conformations. Recently, organized efforts such as the Antibody Modeling Assessment (AMA) have focused on comparing these modeling servers and determining the accuracy of antibody-modeling techniques relative to one another (11, 12).

Antibody-modeling servers typically rely on comparative modeling to model framework regions and CDR loops, with the notable exception of HCDR3. Framework regions are well conserved between antibodies, and a suitable template can usually be found among antibodies with experimentally determined structures in the PDB. One concern with comparative modeling is the issue of generating chimeras, combining heavy and light chain frameworks and CDR loops from different templates to use the template with the highest sequence homology. The relative orientation of the heavy and light chain V domains has a significant impact on the antigen-binding properties of an antibody (13). Although the framework regions are well conserved, combining heavy and light chain frameworks from different templates results in just one and possibly incorrect relative orientation of the heavy and light chain V domains, introducing error into the predicted antibody structure. In addition, although LCDR loops and HCDR1 and 2 have canonical conformations, grafting these loops onto a disparate framework can result in errors in the relative placement of these loops and their interactions. To address these problems, modeling servers have introduced several different solutions. Some servers such as PIGS allow the user a great deal of input regarding the manner in which chimeras are built, allowing the user to tune these parameters for each model (14). The MOE and WAM modeling servers build an ensemble of chimeras built from different templates and use force field energy minimization algorithms to relieve clashes and determine the most likely conformation (15, 16). RosettaAntibody takes a similar approach but uses a knowledge-based potential rather than molecular mechanics force field to relieve clashes in framework and loop placements (17).

Another major challenge is modeling the noncanonical HCDR3 loop. The sequence and structural variability of this loop make it difficult to model it simply by homology in most cases. The PIGS server attempts to model HCDR3 in the same manner as it models the other loops, by sequence homology to known HCDR3 structures (14). This approach can be effective in cases where a similar HCDR3 exists in the PDB, and as more structures are added, this likelihood increases. The MOE server also grafts HCDR3 loops from a template, followed by a more complicated protocol of HCDR3 clustering, force field energy minimization to build an ensemble of structures (16). However, the variability of the HCDR3 loop is such that a more sophisticated modeling technique is necessary in many cases. Other notable approaches for HCDR3 loop modeling involve either fragment-based or *de novo* modeling. Rosetta Antibody uses a fragment-based approach, pulling short peptide fragments from the PDB and using robotics-based algorithms in conjunction with knowledge-based potentials to close the HCDR3 loop in an energetically favorable conformation (17). WAM uses a similar approach, modeling based on PDB-derived peptides based on different parameters of the HCDR3, such as length and predicted kinked conformation. However, the server follows fragment assembly with a force field-based minimization to achieve a local energy minimum (15).

Since the introduction of these antibody-modeling servers, a standing question has been which method is most effective—in user friendliness, structural accuracy, and transparency of results. To this end, the Antibody Modeling Assessment has held a biannual blinded study to give groups the chance to model the structures of unpublished antigen-binding fragment (Fab) X-ray structures to assess the state of the field. The assessments in 2011 and 2014 involved blind prediction of 9 and 11 Fab fragments, respectively, and analyzed the results by a number of parameters, including total root mean square deviation (RMSD), loop and framework-specific RMSD, and overall structural integrity. Overall RMSD values for all servers are generally within 1.0 to 1.5 Å, with results varying between framework and CDR loops. Framework regions were generally predicted most accurately, with RMSD values consistently within 1.0 Å.

Downloaded from www.asmscience.org by
IP: 155.41.210.179
On: Thu, 31 May 2018 15:45:01

The highest deviation was seen for HCDR3 loops, with RMSD values ranging from 0.5 to 8.0 Å and predictive ability depending heavily on the details of individual Fab fragments. Overall, the servers are comparable in their modeling accuracy, with variation dependent on particulars of the Fab fragment and metric used to analyze accuracy (11, 12, 18).

ANTIBODY DOCKING AND EPITOPE MAPPING

Another significant challenge in modeling antibody-antigen complexes is docking the antibody onto its epitope on the surface of the antigen. Though general protein-protein docking has been successful in many cases, antibody-antigen docking represents a special challenge among protein-protein docking cases. Although useful in other protein-protein docking problems, shape complementarity is not a good determinant of correct antibody placement, as epitopes and paratopes are typically flat. Binding affinity is instead determined by hydrophobic "hot spot" and electrostatic interactions (19, 20). Short, aggregation-prone regions are common in many antibodies, consisting of mainly aromatic residues and concentrated in HCDR2 loops, and can contribute significantly to buried surface area in antibody-antigen interactions (21). In addition, the model for the antibody used as a starting point for docking might have to be altered when the epitope is engaged: if a single conformation is used, conformation of CDR loops or the relative orientation of the V domains might have to be changed when engaging the epitope. If an ensemble of states is used, the best conformation needs to be selected and possibly refined further for optimal binding (22, 23). Additionally, if the location of the epitope is unknown, a global search needs to be performed. This adds another degree of complexity by requiring the docking algorithm to sample epitopes across the entire surface of the antigen. Several approaches have been implemented to address these challenges, such as algorithms to remodel antibody loop conformations in the presence of the antigen and application of experimental data to inform the energy evaluation of docked models.

Antibody Docking Algorithms

One docking algorithm that has had success in the realm of antibody-antigen docking is the RosettaDock algorithm, originally described by Gray et al. (24). The algorithm uses alternating rounds of low-resolution rigid body perturbations and high-resolution side-chain and backbone minimization to generate a model of the docked complex. The RosettaDock protocol relies on random perturbation of the complex and creates large numbers ($\sim 10^5$) of models to capture a global energy minimum. The original protocol was able to recover native conformations with an RMSD on the order of 5 to 10 Å for local searches, with a higher error for global searches. Encouragingly, the antibody-antigen docked complexes showed a strong energy funnel, with low energy structures corresponding to a low RMSD to the native structure. This funnel validates the creation of a large number of complexes, with confidence that those with the lowest scores are most likely to recover the native conformation.

Since its original publication, the RosettaDock protocol has been benchmarked thoroughly with large numbers of both native and homology-modeled structures. Advances in comparative antibody-modeling capabilities beg the question of how well these comparative models can be docking into their native antigens to produce a native-like structure. In general, antibody-modeling servers provide similar results in terms of success in docking comparative models. However, since scores fail to correlate perfectly with RMSD values, with the lowest RMSD model frequently not being the best scoring and vice versa, the best results have been obtained with ensemble docking models, using multiple models from the lower end of the energy landscape as inputs for docking. Using such an ensemble approach has been shown to flatten the energy funnel and increase the proportion of models deemed high and medium quality (9).

Additional docking protocols have been published based on the core algorithm in RosettaDock but tailored specifically to antibody-antigen complexes. In particular, the SnugDock algorithm has been shown to increase docking accuracy for antibody-antigen complexes (25). This algorithm uses the same approach of alternating low- and high-resolution perturbation and minimization steps. However, it adds additional perturbation moves that are specific to antibody-binding motifs, such as CDR loop remodeling and reorientation of the angle between the V domains. This addresses the issue of antibody CDR loops adopting alternate conformations in the bound and free states, and the algorithm is designed to accommodate slight errors in comparative modeling that may disrupt adoption of the native conformation during docking. Although SnugDock benchmarking shows similar results to RosettaDock when using the lowest energy models for docking, ensemble docking using SnugDock shows a marked improvement of the energy funnel and increase of high- and medium-quality structures. This result is consistent with a conformational

Downloaded from www.asmscience.org by
IP: 155.41.210.179
On: Thu, 31 May 2018 15:45:01

selection paradigm for the initial antibody-antigen interaction with a subsequent induced fit.

Incorporation of Experimental Data To Guide Antibody Docking and Epitope Mapping

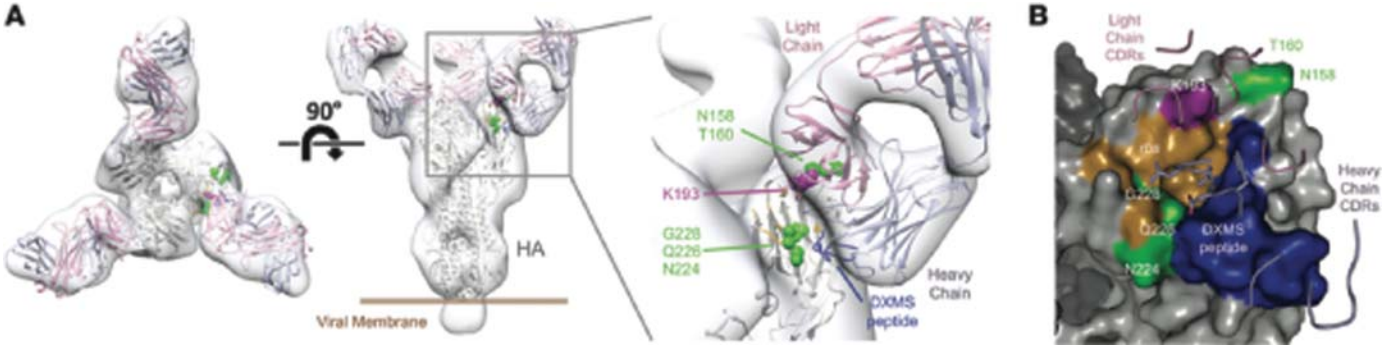
The ability to identify a precise binding epitope of an antibody with its corresponding antigen is of obvious use for any antibody being studied. The knowledge of a binding epitope at the individual residue level allows more detailed analysis of the nature of an antibody's interaction with its antigen, including prediction of the mechanism of neutralization of an antibody, escape mutations that may evade binding, specificity of an antibody in binding-related antigens, etc. However, epitope mapping is an area that has relied on low- and medium-throughput experimental techniques, which can make it difficult to map the epitopes of large numbers of antibodies.

Experimental methods such as EM, NMR spectroscopy chemical shift mapping, competitive enzyme-linked immunosorbent assays, site-directed mutagenesis, force spectroscopy, and hydrogen-deuterium exchange have made it possible to obtain structural information on the antibody-antigen complex, including epitope and paratope, without the need for a crystal structure (26–31). Each technique faces its own challenges that make high-throughput epitope mapping difficult (32), but these data have been used successfully to create reliable computational models. A comprehensive example of integrating experimental data into computational modeling is shown in Fig. 4: data from EM, hydrogen-deuterium exchange, and site-directed mutagenesis were used to create a high-resolution model of an antibody-antigen complex for an influenza hemagglutinin directed antibody (33).

Computational methods have also made great strides in recent years to enable the large-scale, rapid mapping of binding epitope. As with experimental methods, computational approaches come with their own disadvantages and cannot be used reliably in isolation. However, various hybrid methods that use different types of experimental data to reduce the complexity of computational searches have shown great promise in providing a feasible approach to the epitope mapping problem. A promising approach to increase the accuracy of docking predictions is the incorporation of experimental data to supplement *in silico* epitope prediction. RosettaDock, in particular, has been benchmarked on both local and global searches and predictably shows better results when perturbations are kept to a minimum and the antibody is placed in the vicinity of the epitope to begin with. Such approaches reduce the conformational space that needs to be sampled.

In addition to reducing sampling space, limited experimental data can also be used to improve scoring. One major issue with docking using current methods is that a near-native conformation is frequently present in the large set of models, but the distinguishing power of the energy function is not sufficient to identify the near-native conformation without any *a priori* knowledge of the structure. Low-resolution epitope mapping using experimental methods can provide an extra distinguishing feature to eliminate incorrect models and identify those that are most likely to adopt the native conformation. Simonelli et al. have validated this hypothesis by using NMR chemical shift mapping to identify an epitope before the start of computational docking (26). Models that do not agree with the experimental results can be discarded, and lower scoring models are more likely to represent the native conformation.

FIGURE 4 Integrating experimental data to aid computational modeling. (A) Low-resolution cryo-EM maps, (A and B) combined with hydrogen-deuterium exchange (DXMS) data and site-directed mutagenesis, were used to generate a docked model of a potent anti-influenza antibody. Reprinted from *Journal of Clinical Investigation* (33) with permission from the publisher. doi:10.1128/microbiolspec.AID-0024-2014.f4



Downloaded from www.asmscience.org by
IP: 155.41.210.179
On: Thu, 31 May 2018 15:45:01

Docking algorithms have been adapted to incorporate experimental data, which increases the accuracy of the final docked complex and the residues involved in the antibody-antigen interface. There are many different steps during the docking protocol in which experimental data can be incorporated. Mutagenesis data and cryo-EM density maps have been used to constrain the initial placement of the antibody on the surface of the antigen prior to docking, allowing for a more limited search and resulting in more native-like contacts (33–35). In addition, complex scores can be weighted such that residues experimentally shown to interact are encouraged to do so during the docking protocol (34, 35). This approach drives the formation of models that contain the correct residue pairwise interactions across the interface, while allowing for the creation of new interacting pairs within this framework. Final models can then be clustered and analyzed visually for adherence to experimental data, eliminating those that do not contain key residues at the interface or have topologies inconsistent with EM data.

Another source of experimental data that has been used to improve computational epitope prediction is hydrogen-deuterium exchange (36). This technique involves the dissociation of amide hydrogens on a protein backbone and replacement with deuterium from a solvent. The extent of exchange can then be quantified using NMR or mass spectrometry. Antibody binding to its antigen causes a decrease in solvent accessibility of peptides at the epitope and a concomitant decrease in deuterium incorporation. Epitopes of many different antibodies have been mapped using this technique, with higher throughput than other methods such as X-ray crystallography or mutagenesis (27, 28, 37). Several studies have taken advantage of this phenomenon to improve docking predictions by constraining the initial placement of the antibody and rewarding residue interactions that agree with hydrogen-deuterium exchange data (33, 38).

Other approaches to epitope mapping have moved away from docking and instead use bioinformatics strategies to identify interacting regions. One such approach uses binding of randomly created peptide libraries to determine motifs that are important for antibody recognition (39). The algorithm then identifies pairs of residues within peptide fragments that covary with the binding by the target antibody, allowing for identification of key residues both proximal and distal to the epitope. This method has been validated using antibodies targeting viral protein gp120 from HIV and is beneficial as a complement to docking. Another similar approach uses neutralization data of an antibody with different viral strains varying by sequence. This algorithm identifies mutations that tend to correlate with a decrease in neutralization and proposes these as the binding epitope. In addition, the algorithm can incorporate structural data to eliminate purported epitope residues on the basis of solvent-accessible surface area (40, 41). This method can be beneficial in cases where a large body of neutralization data already exists in the literature for an antibody, as extra experimental data do not have to be collected solely for use by the algorithm.

ANTIBODY DESIGN

The computational design of antibodies is not only the most stringent test of our understanding of the rules that govern antibody structure and interaction, but it also has exciting applications in designing an antibody optimized for a given epitope (affinity maturation) or an antibody that recognizes multiple similar target epitopes (broad neutralization). Through this approach the relation between the sequence, structure, and activity of antibodies can be better understood, as the sequence and structural space can be explored in a more comprehensive manner than possible by analysis of naturally occurring antibodies only. Recently, an important proof-of-principle experiment for computer-aided epitope-focused vaccine design was reported (42). For this paradigm to reach its full potential, knowledge of the optimal antibodies to engage an epitope and the relation between sequence, structure, and activity inferred from computational design must be integrated. Besides the obvious application for the development of better therapeutic antibodies, computational antibody engineering also has the potential for formulation and humanization of therapeutic antibodies (43–46).

Broad Neutralization versus Affinity Maturation

Affinity maturation is a process by which the variable region of an antibody undergoes somatic hypermutation to introduce point mutations in the framework and CDR loops to select for a variant with increased binding affinity for its target. Along with V(D)J recombination and junctional diversity, it is a fundamental reason why a finite set of antibody sequences is able to recognize a virtually infinite array of antigens (47). Experimental methods can be used to re-create maturation in an *ex vivo* context to create an antibody with higher affinity for a given target. For example, phage display combined with random mutagenesis can screen for mutations in a high-throughput manner to create new, tighter-binding

Downloaded from www.asmscience.org by
IP: 155.41.210.179
On: Thu, 31 May 2018 15:45:01

antibodies (48). This form of "directed evolution" has been successfull in the maturation of many types of antibodies (48). Additionally, computational algorithms have been developed that mimic the process of directed evolution to produce similar results. Computational methods have been successful in both studying the nature of affinity maturation in germline and matured antibodies and in further refining matured antibodies to increase affinity even more.

Affinity maturation can be simulated *in silico* to analyze compromises in an antibody sequence that lead to a decrease in polyspecificity with a concomitant increase in specificity for a single partner. For example, a general computational method for the design of antibody CDR loops for targeted epitope binding was introduced by Pantazes and Maranas (49). Computational methods such as multistate design are capable of determining the protein sequence optimal for binding an arbitrary number of binding partners (50, 51). This technique has been applied to explore the changes in antibody sequence and conformation responsible for the shift from a polyspecific, germline antibody to one with higher affinity for a single target. In complementary work, Babor and Kortemme and Willis et al. used multistate design to show that antibody germline sequences are optimal for conformational flexibility of both CDR loops and framework residues, allowing the binding of multiple targets, whereas affinity-matured antibodies have decreased flexibility (52, 53). These authors have also identified the key residues responsible for either mono- or polyspecificity for several commonly seen germline genes. These studies validate the biological relevance of design algorithms, since sequences can be both computationally matured and reverted to germline by using different sets of antigens as inputs.

In silico Affinity Maturation

The understanding of the nature of affinity maturation, as well as the ability to selectively modify the specificity of antibody sequences, has led to advances in antibody engineering that enable maturing antibodies *in silico* to create new sequences with higher affinity. In one case, Clark et al. were able to use computational design to mature an antibody and generate candidate sequences with higher predicted affinity (54). Using a combination of side chain repacking and electrostatic optimization, a triple mutant was created with 10 times higher affinity. A comparable increase in affinity was achieved by Lippow et al. by redesigning an antilysozyme antibody along with the therapeutic antibody cetuximab (55). The design protocol was also able to predict mutations in bevacizumab that had been previously shown to increase affinity. The designed mutations primarily affected the electrostatic nature of the binding interface, either by removing a poorly satisfied polar residue at the interface or by adding a polar residue at the solvent-facing periphery of the interface (55). A similar approach has been taken to increase the species cross-reactivity of an antibody, rather than increasing affinity for a previously targeted antigen. By analyzing sequence differences between two serine protease orthologs, Farady et al. created novel antibody designs by restricting the search space to positions that contact points of difference between orthologs (56). In this manner they were able to target positions that would be most likely to establish new contacts across the binding interface to enable interaction at a reasonable affinity. This method was able to create antibody mutants with increases in affinity of over two orders of magnitude.

One significant limitation of most computational design protocols is that they require a high-resolution crystal structure of the antibody-antigen complex, or alternatively high-resolution structures of each component separately. However, several antibody designs have been made for complexes that do not have a solved structure available, using a combination of comparative modeling, protein-protein docking, and design. Barderas et al. used experimental epitope mapping data to dock a comparative model of an antigastrin antibody onto the surface of its target (57). They then used the docked models to estimate regions of antibody-antigen interaction and created mutants using both phage display and *in silico* affinity maturation to mutagenize antibody residues in contact with the antigen and produce designs with high predicted affinity. In several cases the *in silico* suggested mutations matched the mutations seen by directed evolution, and overall the designs were able to increase affinity to nanomolar levels. Another case used docking of an antidengue antibody with an NMR-mapped epitope to identify and rationally design mutations in the antibody CDR loops (58). The authors used this information to create several types of antibody mutations, including those that abolish binding, those that increase affinity for a single target, and those that increase the breadth of binding to multiple serotypes.

Eliciting Neutralizing Antibodies through Antigen Design

Although much focus has been directed at engineering antibodies with desired properties, recent work has targeted the opposite side of the problem: engineering an

Downloaded from www.asmscience.org by
IP: 155.41.210.179
On: Thu, 31 May 2018 15:45:01

antigen that can elicit a desired antibody in an effective and reproducible manner. This comes with the ultimate goal of the rational design of antigens to be used in vaccination that can elicit antibodies targeting a precise, conserved, and neutralizing epitope. Giles et al. attempted to create a broadly reactive antigen by identifying divergent sequences between different clades of influenza hemagglutinin and clustering the consensus sequences. In this way they created a compromise antigen, incorporating elements of various clades designed to elicit antibodies binding all of them (59). This strategy was validated with the finding that the broadly reactive antigen was capable of eliciting greater antibody breadth than a polyvalent virus-like particle displaying the native antigens (60). Wu et al. pursued a similar goal, instead using HIV surface protein gp120 as the antigen. This group used computational design to engineer a modified gp120 that maintained the structure of the neutralizing portion of the molecule while eliminating other antigenic regions that may elicit nonneutralizing antibodies (61). Since the antigenic region of gp120 they targeted was the CD4 binding site, a highly conserved region among divergent HIV strains, they used this engineered antigen as a vaccine to elicit broadly neutralizing antibodies and successfully identified two novel antibodies with a high level of breadth of reactivity (61).

Epitope Grafting To Elicit Neutralizing Antibodies

A complementary task to antigen design is so-called epitope grafting—removing an epitope from its native antigen and grafting it onto a protein structure that can present it to immune receptors in a way that maximizes the immune response. This design task encompasses several unique challenges, the most difficult of which is the selection of an appropriate peptide onto which the epitope can be placed, known as the scaffold. This scaffold must maintain the native conformation of the desired epitope and have minimal immunogenicity in its off-epitope regions, all while maintaining favorable biophysical properties. Advances in scaffold selection and design have made this problem tractable and have shown promise in immunization.

Early work in this field focused on placing a neutralizing linear epitope on a stable scaffold that could maintain the native conformation while enhancing presentation of this epitope. HIV has been widely used as a test case, since there are well-characterized broadly neutralizing antibodies with high-resolution crystal structures available (62). Two independent groups were able to graft the linear epitopes of two broadly neutralizing anti-HIV antibodies, 4E10 and 2F5, onto scaffolds to enhance their presentation and affinity for the desired antibodies (63, 65). These groups used similar approaches of searching the PDB for scaffolds with a region of backbone conformation with high structural similarity to the epitope and using Rosetta Design to place epitope side chains onto the backbone scaffold and minimize the energy of side chain packing. This protocol resulted in several designed peptides with nanomolar affinity for their respective antibody targets and the ability to elicit broadly neutralizing sera upon immunization (63–65). These earlier studies provided an important proof-of-principle for vaccine-based antigen design.

One major limitation of initial reports on epitope grafting is the nature of the scaffold selection process. Since it relies on placing epitope side chains on a backbone with a high degree of similarity to the epitope backbone, success is highly dependent on the presence of a suitable scaffold among structurally determined proteins. In cases where there is no template with structural similarity, this procedure is ineffective. To address this concern, Azoitei et al. published a backbone grafting method wherein they select a structurally similar template and, rather than transfer individual epitope side chains to a template, remove the epitope-mimetic region en masse and transplant the entire epitope backbone region from the native antigen (66). This protocol was benchmarked and shown to produce peptides with higher affinity than the previously described side-chain grafting method, with the added advantage that backbone grafting can be applied to templates with lower structural similarity.

Another challenge in the pursuit of a generalizable epitope grafting method is that initial reports focused on linear, continuous epitopes, even though many neutralizing epitopes are discontinuous. Azoitei et al. developed a more aggressive scaffold search and design method to identify scaffolds that can accommodate multiple discontinuous epitope peptides while maintaining their conformations and relative orientations and minimizing steric clashes (67). The result of this protocol was an epitope scaffold that accurately recapitulates the conformation of the epitope and interactions necessary for antibody-antigen interactions. McLellan et al. pursued a similar strategy to create a discontinuous epitope-presenting scaffold with a respiratory syncytial virus (RSV) epitope (68). In addition to the selection of an appropriate scaffold, there has also been work in redesigning a scaffold to reduce unwanted, nonepitope immunogenicity. Correia et al. showed that epitope scaffolds can be further optimized

Downloaded from www.asmscience.org by
IP: 155.41.210.179
On: Thu, 31 May 2018 15:45:01

by flexible backbone remodeling and resurfacing to enhance thermostability, increase binding affinity, and reduce immune reactivity (64).

To circumvent the problem of scaffold selection, Correia et al. recently described a method to build a *de novo* scaffold for optimal epitope presentation of an RSV epitope (42). They selected a three-helix bundle as a template topology and used extensive rounds of sequence optimization and minimization to create an optimized conformation for presentation. These scaffolds were purified and shown to be thermodynamically stable, with affinities in the picomolar range (42). The purified scaffolds were then used to immunize macaques and were shown to induce neutralizing titers comparable to natural RSV infection, a significant landmark for the epitope-grafting methodology.

CONCLUSION

Computational approaches to predict the structure of antibodies and antibody-antigen complexes are of critical importance, as the large number of naturally occurring antibodies restricts experimental characterization to the most important cases. While computational methods are already sufficiently accurate to be useful, it remains a focus of future research to develop better sampling methods and more accurate energy functions. The existing limitations in modeling antibodies make computational methods most useful when applied in a tight feedback loop with experimental data, a situation that is expected to continue for the foreseeable future. The increasing throughput in methods to collect limited experimental data on antibodies and antibody-antigen complexes will further increase the need for computational methods to add atomic detail not present in these datasets. The rapidly increasing availability of high-resolution crystal structures of antibodies and antibody-antigen complexes is expected to improve computational prediction algorithms, for example, through increasing the accuracy of knowledge-based potentials and through further completing the conformational clusters of CDR loop conformations. Ultimately, the reliable computational design of antibodies that recognize a target epitope is a long-term goal for computational structural biology. Given the even larger number of theoretically possible antibody sequences, computational prioritization of the ones to characterize experimentally is imperative. Computational design of the tightest or most broadly neutralizing antibodies is not only important for the development of optimal therapeutic antibodies or the development of vaccines. Antibody design is also the most stringent test of our understanding of the rules that govern antibody structure and interaction.

ACKNOWLEDGMENT

Conflict of interest: We disclose no conflicts.

REFERENCES

1. Harris LJ, Larson SB, Hasel KW, McPherson A. 1997. Refined structure of an intact IgG2a monoclonal antibody. *Biochemistry* **36:**1581–1597.

2. Chothia C, Lesk AM. 1987. Canonical structures for the hypervariable regions of immunoglobulins. *J Mol Biol* **196:**901–917.

3. Martin AC, Thornton JM. 1996. Structural families in loops of homologous proteins: automatic classification, modelling and application to antibodies. *J Mol Biol* **263:**800–815.

4. Al-Lazikani B, Lesk AM, Chothia C. 1997. Standard conformations for the canonical structures of immunoglobulins. *J Mol Biol* **273:**927–948.

5. North B, Lehmann A, Dunbrack RL, Jr. 2011. A new clustering of antibody CDR loop conformations. *J Mol Biol* **406:**228–256.

6. Briney BS, Willis JR, Crowe JE, Jr. 2012. Location and length distribution of somatic hypermutation-associated DNA insertions and deletions reveals regions of antibody structural plasticity. *Genes Immun* **13:**523–529.

7. Morea V, Tramontano A, Rustici M, Chothia C, Lesk AM. 1998. Conformations of the third hypervariable region in the VH domain of immunoglobulins. *J Mol Biol* **275:**269–294.

8. Kuroda D, Shirai H, Kobori M, Nakamura H. 2008. Structural classification of CDR-H3 revisited: a lesson in antibody modeling. *Proteins* **73:**608–620.

9. Sivasubramanian A, Sircar A, Chaudhury S, Gray JJ. 2009. Toward high-resolution homology modeling of antibody Fv regions and application to antibody-antigen docking. *Proteins* **74:**497–514.

10. Zhu K, Day T. 2013. *Ab initio* structure prediction of the antibody hypervariable H3 loop. *Proteins* **81:**1081–1089.

11. Teplyakov A, Luo J, Obmolova G, Malia TJ, Sweet R, Stanfield RL, Kodangattil S, Almagro JC, Gilliland GL. 2014. Second antibody modeling assessment. II. Structures and models. *Proteins* **82:**1563–1582. doi:10.1002/prot.24554.

12. Almagro JC, Beavers MP, Hernandez-Guzman F, Maier J, Shaulsky J, Butenhof K, Labute P, Thorsteinson N, Kelly K, Teplyakov A, Luo J, Sweet R, Gilliland GL. 2011. Antibody modeling assessment. *Proteins* **79:**3050–3066.

13. Narayanan A, Sellers BD, Jacobson MP. 2009. Energy-based analysis and prediction of the orientation between light- and heavy-chain antibody variable domains. *J Mol Biol* **388:**941–953.

14. Marcatili P, Rosi A, Tramontano A. 2008. PIGS: automatic prediction of antibody structures. *Bioinformatics* **24:**1953–1954.

15. Whitelegg NR, Rees AR. 2000. WAM: an improved algorithm for modelling antibodies on the WEB. *Protein Eng* **13:**819–824.

16. Maier JKY, Labute P. 2014. Assessment of fully automated antibody homology modeling protocols in MOE. *Proteins* **82:**1599–1610.

17. Sircar A, Kim ET, Gray JJ. 2009. RosettaAntibody: antibody variable region homology modeling server. *Nucleic Acids Res* **37:**W474–W479.

18. Almagro JC, Teplyakov A, Luo J, Sweet RW, Kodangattil S, Hernandez-Guzman F, Gilliland GL. 2014. Second antibody modeling assessment (AMA-II). *Proteins* **82:**1553–1562.

19. Sundberg EJ, Urrutia M, Braden BC, Isern J, Tsuchiya D, Fields BA, Malchiodi EL, Tormo J, Schwarz FP, Mariuzza RA. 2000. Estimation of the hydrophobic effect in an antigen-antibody protein-protein interface. *Biochemistry* **39:**15375–15387.

20. Moreira IS, Fernandes PA, Ramos MJ. 2007. Hot spot computational identification: application to the complex formed between the hen egg

Downloaded from www.asmscience.org by
IP: 155.41.210.179
On: Thu, 31 May 2018 15:45:01

white lysozyme (HEL) and the antibody HyHEL-10. *Int J Quantum Chem* **107:**299–310.

21. Wang X, Singh SK, Kumar S. 2010. Potential aggregation-prone regions in complementarity-determining regions of antibodies and their contribution towards antigen recognition: a computational analysis. *Pharm Res* **27:**1512–1529.

22. Vajda S. 2005. Classification of protein complexes based on docking difficulty. *Proteins* **60:**176–180.

23. Brenke R, Hall DR, Chuang GY, Comeau SR, Bohnuud T, Beglov D, Schueler-Furman O, Vajda S, Kozakov D. 2012. Application of asymmetric statistical potentials to antibody-protein docking. *Bioinformatics* **28:**2608–2614.

24. Gray JJ, Moughon S, Wang C, Schueler-Furman O, Kuhlman B, Rohl CA, Baker D. 2003. Protein-protein docking with simultaneous optimization of rigid-body displacement and side-chain conformations. *J Mol Biol* **331:**281–299.

25. Sircar A, Gray JJ. 2010. SnugDock: paratope structural optimization during antibody-antigen docking compensates for errors in antibody homology models. *PLoS Comput Biol* **6:**e1000644. doi:10.1371/journal.pcbi.1000644.

26. Simonelli L, Beltramello M, Yudina Z, Macagno A, Calzolai L, Varani L. 2010. Rapid structural characterization of human antibody-antigen complexes through experimentally validated computational docking. *J Mol Biol* **396:**1491–1507.

27. Sevy AM, Healey JF, Deng W, Spiegel PC, Meeks SL, Li R. 2013. Epitope mapping of inhibitory antibodies targeting the C2 domain of coagulation factor VIII by hydrogen-deuterium exchange mass spectrometry. *J Thromb Haemost* **11:**2128–2136.

28. Coales SJ, Tuske SJ, Tornasso JC, Hamuro Y. 2009. Epitope mapping by amide hydrogen/deuterium exchange coupled with immobilization of antibody, on-line proteolysis, liquid chromatography and mass spectrometry. *Rapid Commun Mass Spectrom* **23:**639–647.

29. Meeks SL, Healey JF, Parker ET, Barrow RT, Lollar P. 2007. Anti-human factor VIIIC2 domain antibodies in hemophilia A mice recognize a functionally complex continuous spectrum of epitopes dominated by inhibitors of factor VIII activation. *Blood* **110:**4234–4242.

30. Chaves RC, Teulon JM, Odorico M, Parot P, Chen SWW, Pellequer JL. 2013. Conformational dynamics of individual antibodies using computational docking and AFM. *J Mol Recognit* **26:**596–604.

31. Gogolinska A, Nowak W. 2013. Molecular basis of lateral force spectroscopy nano-diagnostics: computational unbinding of autism related chemokine MCP-1 from IgG antibody. *J Mol Model* **19:**4773–4780.

32. Sharon J, Rynkiewicz MJ, Lu ZH, Yang CY. 2014. Discovery of protective B-cell epitopes for development of antimicrobial vaccines and antibody therapeutics. *Immunology* **142:**1–23.

33. Thornburg NJ, Nannemann DP, Blum DL, Belser JA, Tumpey TM, Deshpande S, Fritz GA, Sapparapu G, Krause JC, Lee JH, Ward AB, Lee DE, Li S, Winarski KL, Spiller BW, Meiler J, Crowe JE, Jr. 2013. Human antibodies that neutralize respiratory droplet transmissible H5N1 influenza viruses. *J Clin Invest* **123:**4405–4409.

34. McKinney BA, Kallewaard NL, Crowe JE, Jr, Meiler J. 2007. Using the natural evolution of a rotavirus-specific human monoclonal antibody to predict the complex topography of a viral antigenic site. *Immunome Res* **3:**8.

35. Schneider S, Zacharias M. 2012. Atomic resolution model of the antibody Fc interaction with the complement C1q component. *Mol Immunol* **51:**66–72.

36. Percy AJ, Rey M, Burns KM, Schriemer DC. 2012. Probing protein interactions with hydrogen/deuterium exchange and mass spectrometry: a review. *Anal Chim Acta* **721:**7–21.

37. Baerga-Ortiz A, Hughes CA, Mandell JG, Komives EA. 2002. Epitope mapping of a monoclonal antibody against human thrombin by R/D-exchange mass spectrometry reveals selection of a diverse sequence in a highly conserved protein. *Protein Sci* **11:**1300–1308.

38. Aiyegbo MS, Sapparapu G, Spiller BW, Eli IM, Williams DR, Kim R, Lee DE, Liu T, Li S, Woods VL, Jr, Nannemann DP, Meiler J, Stewart PL, Crowe JE, Jr. 2013. Human rotavirus VP6-specific antibodies mediate intracellular neutralization by binding to a quaternary structure in the transcriptional pore. *PloS One* **8:**e61101. doi:10.1371/journal.pone.0061101.

39. Bublil EM, Yeger-Azuz S, Gershoni JM. 2006. Computational prediction of the cross-reactive neutralizing epitope corresponding to the monoclonal antibody b12 specific for HIV-1 gp120. *FASEB J* **20:**1762–1774.

40. Chuang GY, Acharya P, Schmidt SD, Yang Y, Louder MK, Zhou T, Kwon YD, Pancera M, Bailer RT, Doria-Rose NA, Nussenzweig MC, Mascola JR, Kwong PD, Georgiev IS. 2013. Residue-level prediction of HIV-1 antibody epitopes based on neutralization of diverse viral strains. *J Virol* **87:**10047–10058.

41. Chuang GY, Liou D, Kwong PD, Georgiev IS. 2014. NEP: web server for epitope prediction based on antibody neutralization of viral strains with diverse sequences. *Nucleic Acids Res* **42:**W64–W71.

42. Correia BE, Bates JT, Loomis RJ, Baneyx G, Carrico C, Jardine JG, Rupert P, Correnti C, Kalyuzhniy O, Vittal V, Connell MJ, Stevens E, Schroeter A, Chen M, Macpherson S, Serra AM, Adachi Y, Holmes MA, Li Y, Klevit RE, Graham BS, Wyatt RT, Baker D, Strong RK, Crowe JE, Jr, Johnson PR, Schief WR. 2014. Proof of principle for epitope-focused vaccine design. *Nature* **507:**201–206.

43. Hinton PR, Johlfs MG, Xiong JM, Hanestad K, Ong KC, Bullock C, Keller S, Tang MT, Tso JY, Vasquez M, Tsurushita N. 2004. Engineered human IgG antibodies with longer serum half-lives in primates. *J Biol Chem* **279:**6213–6216.

44. Igawa T, Tsunoda H, Tachibana T, Maeda A, Mimoto F, Moriyama C, Nanami M, Sekimori Y, Nabuchi Y, Aso Y, Hattori K. 2010. Reduced elimination of IgG antibodies by engineering the variable region. *Protein Eng Des Sel* **23:**385–392.

45. Kanduc D, Lucchese A, Mittelman A. 2001. Individuation of monoclonal anti-HPV16 E7 antibody linear peptide epitope by computational biology. *Peptides* **22:**1981–1985.

46. Tsurushita N, Hinton PR, Kumar S. 2005. Design of humanized antibodies: from anti-Tac to Zenapax. *Methods* **36:**69–83.

47. Teng G, Papavasiliou FN. 2007. Immunoglobulin somatic hypermutation. *Annu Rev Genet* **41:**107–120.

48. Filpula D. 2007. Antibody engineering and modification technologies. *Biomol Eng* **24:**201–215.

49. Pantazes RJ, Maranas CD. 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.

50. Leaver-Fay A, Jacak R, Stranges PB, Kuhlman B. 2011. A generic program for multistate protein design. *PloS One* **6:**e20937. doi:10.1371/journal.pone.0020937.

51. Havranek JJ, Harbury PB. 2003. Automated design of specificity in molecular recognition. *Nat Struct Biol* **10:**45–52.

52. Willis JR, Briney BS, DeLuca SL, Crowe JE, Jr, Meiler J. 2013. Human germline antibody gene segments encode polyspecific antibodies. *PLoS Comput Biol* **9:**e1003045. doi:10.1371/journal.pcbi.1003045.

53. Babor M, Kortemme T. 2009. Multi-constraint computational design suggests that native sequences of germline antibody H3 loops are nearly optimal for conformational flexibility. *Proteins* **75:**846–858.

54. Clark LA, Boriack-Sjodin PA, Eldredge J, Fitch C, Friedman B, Hanf KJ, Jarpe M, Liparoto SF, Li Y, Lugovskoy A, Miller S, Rushe M, Sherman W, Simon K, Van Vlijmen H. 2006. Affinity enhancement of an *in vivo* matured therapeutic antibody using structure-based computational design. *Protein Sci* **15:**949–960.

55. Lippow SM, Wittrup KD, Tidor B. 2007. Computational design of antibody-affinity improvement beyond *in vivo* maturation. *Nat Biotechnol* **25:**1171–1176.

Downloaded from www.asmscience.org by
IP: 155.41.210.179
On: Thu, 31 May 2018 15:45:01

56. Farady CJ, Sellers BD, Jacobson MP, Craik CS. 2009. Improving the species cross-reactivity of an antibody using computational design. *Bioorg Med Chem Lett* **19:**3744–3747.

57. Barderas R, Desmet J, Timmerman P, Meloen R, Casal JI. 2008. Affinity maturation of antibodies assisted by *in silico* modeling. *Proc Natl Acad Sci USA* **105:**9029–9034.

58. Simonelli L, Pedotti M, Beltramello M, Livoti E, Calzolai L, Sallusto F, Lanzavecchia A, Varani L. 2013. Rational engineering of a human anti-dengue antibody through experimentally validated computational docking. *PloS One* **8:**e55561. doi:10.1371/journal.pone.0055561.

59. Giles BM, Ross TM. 2011. A computationally optimized broadly reactive antigen (COBRA) based H5N1 VLP vaccine elicits broadly reactive antibodies in mice and ferrets. *Vaccine* **29:**3043–3054.

60. Giles BM, Bissel SJ, DeAlmeida DR, Wiley CA, Ross TM. 2012. Antibody breadth and protective efficacy are increased by vaccination with computationally optimized hemagglutinin but not with polyvalent hemagglutinin-based H5N1 virus-like particle vaccines. *Clin Vaccine Immunol* **19:**128–139.

61. Wu X, Yang ZY, Li Y, Hogerkorp CM, Schief WR, Seaman MS, Zhou T, Schmidt SD, Wu L, Xu L, Longo NS, McKee K, O'Dell S, Louder MK, Wycuff DL, Feng Y, Nason M, Doria-Rose N, Connors M, Kwong PD, Roederer M, Wyatt RT, Nabel GJ, Mascola JR. 2010. Rational design of envelope identifies broadly neutralizing human monoclonal antibodies to HIV-1. *Science* **329:**856–861.

62. Stamatatos L, Morris L, Burton DR, Mascola JR. 2009. Neutralizing antibodies generated during natural HIV-1 infection: good news for an HIV-1 vaccine? *Nat Med* **15:**866–870.

63. Correia BE, Ban YE, Holmes MA, Xu H, Ellingson K, Kraft Z, Carrico C, Boni E, Sather DN, Zenobia C, Burke KY, Bradley-Hewitt T, Bruhn-Johannsen JF, Kalyuzhniy O, Baker D, Strong RK, Stamatatos L, Schief WR. 2010. Computational design of epitope-scaffolds allows induction of antibodies specific for a poorly immunogenic HIV vaccine epitope. *Structure* **18:**1116–1126.

64. Correia BE, Ban YE, Friend DJ, Ellingson K, Xu H, Boni E, Bradley-Hewitt T, Bruhn-Johannsen JF, Stamatatos L, Strong RK, Schief WR. 2010. Computational protein design using flexible backbone remodeling and resurfacing: case studies in structure-based antigen design. *J Mol Biol* [Epub ahead of print.] doi:10.1016/j.jmb.2010.09.061.

65. Ofek G, Guenaga FJ, Schief WR, Skinner J, Baker D, Wyatt R, Kwong PD. 2010. Elicitation of structure-specific antibodies by epitope scaffolds. *Proc Natl Acad Sci USA* **107:**17880–17887.

66. Azoitei ML, Ban YE, Julien JP, Bryson S, Schroeter A, Kalyuzhniy O, Porter JR, Adachi Y, Baker D, Pai EF, Schief WR. 2012. Computational design of high-affinity epitope scaffolds by backbone grafting of a linear epitope. *J Mol Biol* **415:**175–192.

67. Azoitei ML, Correia BE, Ban YE, Carrico C, Kalyuzhniy O, Chen L, Schroeter A, Huang PS, McLellan JS, Kwong PD, Baker D, Strong RK, Schief WR. 2011. Computation-guided backbone grafting of a discontinuous motif onto a protein scaffold. *Science* **334:**373–376.

68. McLellan JS, Pancera M, Carrico C, Gorman J, Julien JP, Khayat R, Louder R, Pejchal R, Sastry M, Dai K, O'Dell S, Patel N, Shahzad-ul-Hussan S, Yang Y, Zhang B, Zhou T, Zhu J, Boyington JC, Chuang GY, Diwanji D, Georgiev I, Kwon YD, Lee D, Louder MK, Moquin S, Schmidt SD, Yang ZY, Bonsignori M, Crump JA, Kapiga SH, Sam NE, Haynes BF, Burton DR, Koff WC, Walker LM, Phogat S, Wyatt R, Orwenyo J, Wang LX, Arthos J, Bewley CA, Mascola JR, Nabel GJ, Schief WR, Ward AB, Wilson IA, Kwong PD. 2011. Structure of HIV-1 gp120 V1/V2 domain with broadly neutralizing antibody PG9. *Nature* **480:**336–343.

Downloaded from www.asmscience.org by
IP: 155.41.210.179
On: Thu, 31 May 2018 15:45:01