Chapter

In silico Techniques for Prospecting and Characterizing Monoclonal Antibodies

Tania M. Manieri, Carolina G. Magalhaes, Daniela Y. Takata, João V. Batalha-Carvalho and Ana M. Moro

Abstract

In the past few years, improvement in computational approaches provided faster and less expensive outcomes on the identification, development, and optimization of monoclonal antibodies (mAbs). *In silico* methods, such as homology modeling, to predict antibody structures, identification of epitope-paratope interactions, and molecular docking are useful to generate 3D structures of the antibody–antigen complexes. It helps identify the key residues involved in the antigen–antibody complex and enable modifications to enhance the antibody binding affinity. Recent advances in computational tools for redesigning antibodies are significant resources to improve antibody biophysical properties, such as binding affinity, solubility, stability, decreasing the timeframe and costs during antibody engineering. The immunobiological market grows continuously with new molecules, both natural and new molecular formats, such as bispecific antibodies, Fc-antibody fusion proteins, and mAb fragments, requiring novel methods for designing, screening, and analyzing. Algorithms and software set the *in silico* techniques on the innovation frontier.

Keywords: antibody structure modeling, VH-VL, computational analyses, epitope prediction, paratope prediction, molecular docking

1. Introduction

The development of new therapeutic antibodies is a multiple task challenge. The approval of OKT3 (1986), the first therapeutic mAb, opened the perspective of using this class of product in many other antibody-based therapies. Only the concept of "Magic Bullets," however, was not enough to provide for safety and efficacy, resulting in many preclinical or clinical trial failures. Soon became evident the need for the humanization of antibodies that mitigated their immunogenicity with the counterpart of decreasing their affinity in many cases. An alternative to circumvent this issue relied on back mutations. However, how to suggest such mutations?

Methodologies have emerged to optimize newly discovered antibodies, either in their affinity to the target or other properties, including *in silico* methods. The computational capacity has grown exponentially over the past few decades, providing an equally exponential advance in computational drug optimization techniques. The increased public's databases associated with bigdata works transformed the internet in the most profitable "laboratory" with free reagents (i.e., data), low-risk experiments

(i.e., *in silico assays*), and time safe results. Considering all these aspects together are a simple way to understand the webtools' strengths.

In this chapter, we present different aspects of *in silico* methodologies for prospecting and characterizing mAbs. For didactic reasons, we could start from modeling molecules by homology (Section 2), followed by the prediction of epitopes and paratopes, affinity maturation, and molecular docking in Section 3, and finally, the improvement of biophysical and biological properties in Section 4. We aim to present free tools currently available, highlighting their features and applications, allowing the readers to find the most appropriate way to solve their problems. Albeit many tools and their applicability are shown, we call attention to their sequence of use and refinement that is inherent to the particular questions to be answered. We remind you that all the tools presented here were available online and free of charge to the academy until the closing of this edition.

2. Antibody structure modeling

The ability of the antibodies to recognize a diverse set of antigens is acquired by V(D)J recombination and affinity maturation. These two mechanisms contribute to a large number of possible unique sequences of the antibodies, around 10^{11-15} [1–3].

Protein structures are strongly related to specificity and function, and their knowledge is crucial to analyze the antibody. Although many crystal structures are available in the Protein DataBank [4], this number is small (around 6700) compared to the number of possible sequences. Computational modeling is a feasible method for predicting antibodies' structures and allows us to evaluate antibodies' properties and to understand antibody–antigen interaction.

The first step in the antibody modeling is its alignment with the germline sequence and the V(D)J classification. The International Immunogenetics Information System (IMGT) [5] is the central database of germline antibody sequences. Some webtools, such as IgBlast [6] and IMGT/V QUEST [7], use this database to align and classify the annotated sequence (**Table 1**). Since differences in the variable domains are responsible for the structural and functional antibodies' diversity, most of the structure prediction methods are based on the Fv modeling (**Table 1**). Framework regions are sequences with highly conserved structures, making it easier to generate their models from template structures. CDRs from the light chain (CDRs L1–3), CDRs 1 and 2 from the heavy chain (CDRs H1–2) are relatively conserved, regarding their structures, being possible to predict their structures based on their amino acid sequence. There is a set of canonical structures that allows us to predict the conformation of each loop. Recent studies have classified non-CDR H3 loops by their type and length and identified 72 clusters [62]. CDR H3 loop is usually longer (5–26 amino acids) than the others and presents a highly diverse structure. The CDR H3 loop also influences the VL-VH orientation and, consequently, the antibody–antigen interaction [63, 64]. For these reasons, the major challenge in antibody modeling is to achieve accuracy in CDR3 loop structure prediction. Usually, the primary sequence of CDR H3 is not enough for the prediction of the loop conformation.

Some information, as the position of the key residues, seems to be necessary. CDR H3 can be divided into two regions, the torso and the head [65]. The head usually presents a standard hairpin structure [65]. The torso, which is the region closer to the framework, can be predicted comparing with similar antibodies in which crystal structures are available on databases. Some software as Rosetta [66] has a platform to predict antibody structure, which first models each CDR

Antibody structural modeling	Link	Ref.
IgBLAST	https://www.ncbi.nlm.nih.gov/igblast/	[6]
IMGT V-Quest	http://www.imgt.org/IMGTindex/V-QUEST.php	[7]
SPHINX	http://opig.stats.ox.ac.uk/webapps/newsabdab/sabpred/sphinx	[9]
AbodyBuilder	http://opig.stats.ox.ac.uk/webapps/sabdab-sabpred/Modelling.php	[10]
LYRA	http://www.cbs.dtu.dk/services/LYRA/index.php	[11]
Kotai Antibody Builder	https://sysimm.org/rep_builder/	[12]
Rosetta Suite	http://rosie.rosettacommons.org/	[13]
PEARS	http://opig.stats.ox.ac.uk/webapps/sabdab-sabpred/PEARS.php	[14]
SCWRL	http://dunbrack.fccc.edu/scwrl4/	[15]
BetaSCPWeb	http://voronoi.hanyang.ac.kr/betascpweb/	[16]
FREAD	http://opig.stats.ox.ac.uk/webapps/fread/php/	[17]
PLOP	http://www.jacobsonlab.org/plop_manual/plop_overview.htm	[18]
SCALOP	http://opig.stats.ox.ac.uk/webapps/sabdab-sabpred/SCALOP.php	[19]
Paratope prediction	**Link**	**Ref.**
Paratome	http://ofranservices.biu.ac.il/site/services/paratome/	[20]
Antibody i-Patch	http://opig.stats.ox.ac.uk/webapps/sabdab-sabpred/ABipatch.php	[21]
proABC-2	https://bianca.science.uu.nl/proabc2/	[22]
Parapred	https://github.com/eliberis/parapred	[23]
Epitope prediction	**Link**	**Ref.**
EpiPred	http://opig.stats.ox.ac.uk/webapps/sabdab-sabpred/EpiPred.php	[24]
PEASE	http://ofranservices.biu.ac.il/site/services/epitope/index.html	[25]
IEDB (database)	https://www.iedb.org/	[26]
Linear B-cells epitope	http://tools.iedb.org/bcell/	[27, 28]
DiscoTope	http://tools.iedb.org/discotope/	[29, 30]
Ellipro	http://tools.iedb.org/ellipro/	[31]
Epitome (database)	http://www.rostlab.org/services/epitome	[32]
Epitope Conservancy Analysis	http://tools.iedb.org/conservancy/	[33]
Epitope Cluster Analys	http://tools.iedb.org/cluster/	[34]
Antibody–antigen docking	**Link**	**Ref.**
ClusPro	https://cluspro.bu.edu/login.php	[35]
ZDOCK	http://zdock.umassmed.edu/	[36]
PatchDock	https://bioinfo3d.cs.tau.ac.il/PatchDock/	[37]
AutoDock	http://autodock.scripps.edu/	[38]
AutoDockFR	https://ccsb.scripps.edu/adfr/	[39]
FLIPDock	http://flipdock.scripps.edu/	[40]

Antibody structural modeling	Link	Ref.
Swarmdock	https://bmm.crick.ac.uk/~svc-bmm-swarmdock/	[41]
SnugDock	http://rosie.graylab.jhu.edu/snug_dock	[42]
HADDOCK	https://wenmr.science.uu.nl/haddock2.4/	[43]
FRODOCK	http://frodock.chaconlab.org/	[44]
Biophysical properties of mAbs	**Link**	**Ref.**
DeepDDG	http://protein.org.cn/ddg.html	[45]
ScooP	http://babylone.ulb.ac.be/SCooP/	[46]
Mupro1.0	http://sysbio.rnet.missouri.edu/multicom_toolbox/index.html	[47, 48]
Ease MM	https://sparks-lab.org/server/ease-mm/	[49]
Strum	https://zhanglab.ccmb.med.umich.edu/STRUM/	[50]
MCSM	http://biosig.unimelb.edu.au/mcsm/stability	[51]
SDM	http://marid.bioc.cam.ac.uk/sdm2/	[52]
TAP	http://opig.stats.ox.ac.uk/webapps/newsabdab/sabpred/tap	[53]
AgreeRATE-Disc	https://www.iitm.ac.in/bioinfo/aggrerate-disc/	[54]
ISMBLab-PPI	http://ismblab.genomics.sinica.edu.tw/predict.php?pred=PPI	[55]
Prediction of glycosylation spots	**Link**	**Ref.**
NetNGlyc 1.0	https://services.healthtech.dtu.dk/service.php?NetNGlyc-1.0	[56]
N-GlyDE	http://bioapp.iis.sinica.edu.tw/N-GlyDE/	[57]
GlycoMod	https://web.expasy.org/glycomod/	[58]
GlyConnect	https://glyconnect.expasy.org/home	—
GlycoSiteAlign	https://glycoproteome.expasy.org/glycositealign/	[59]
NetCGlyc	http://www.cbs.dtu.dk/services/NetCGlyc/	[60]
NetOGlyc	http://www.cbs.dtu.dk/services/NetOGlyc/	[61]

Table 1.
Free onlinetools and databases. Adapted from Norman et al., 2019 [8].

and framework based on very similar antibodies and then generates the CDR H3 conformations by assembling small peptides fragments. Software as SPHINX [9], uses *ab initio* modeling algorithm to predict CDR H3 conformation. Some software that performs the antibody structure modeling, and also the CDR H3 modeling, are listed in **Table 1**.

Antibody modeling is an essential step for most of the procedures discussed below, and the researcher must proceed according to the necessary refinement.

3. Antibody-antigen complex: methods of paratope and epitope prediction; molecular docking

In the past few years, improvement in computational approaches provided faster and less expensive outcomes on the identification, development, and optimization of monoclonal antibodies (mAbs). One of the leading goals of the rational

development of antibodies is the identification of epitope-paratope interactions. 3D structures of the antibody–antigen complex using X-ray crystallography are the gold standard to reach the binding site information; however, these experimental methods can be money and time-consuming, and scarce to obtain. Thus, computational methods mean a rapid alternative across antibody discovery.

In silico methods, such as homology modeling, molecular docking, and interface prediction can be used to generate 3D models of the antibody–antigen complexes and to predict critical residues involved in antigen binding. Once the antibody–antigen contact residues are known, it can be computationally mutated to screen for residues that could increase antibody specificity and affinity against the target, if desired. Computational techniques to perform such a process fall into those that predict the paratope, the epitope, or the entire antibody–antigen complex (**Figure 1**).

3.1 Paratope prediction

Paratope represents the antibody amino acid residues in direct contact with the antigen. Since antigen-binding typically involves residues in the CDRs, about 80% of the amino acids constituting the paratope are in the CDRs [67]. However, only a third of the CDR residues participate in antigen binding [68]. Besides the residues in the CDRs, some framework regions are also involved in antigen-binding [67], relevant to identify the paratope residues precisely.

Several computational methods exist to predict paratopes (**Table 1**). The *online* tool Paratome [20] indicates the antigen-binding regions given the amino acid sequence or 3D structure. It identifies structural elements consensus, which is commonly involved in antigen binding between antibodies by aligning a set of all
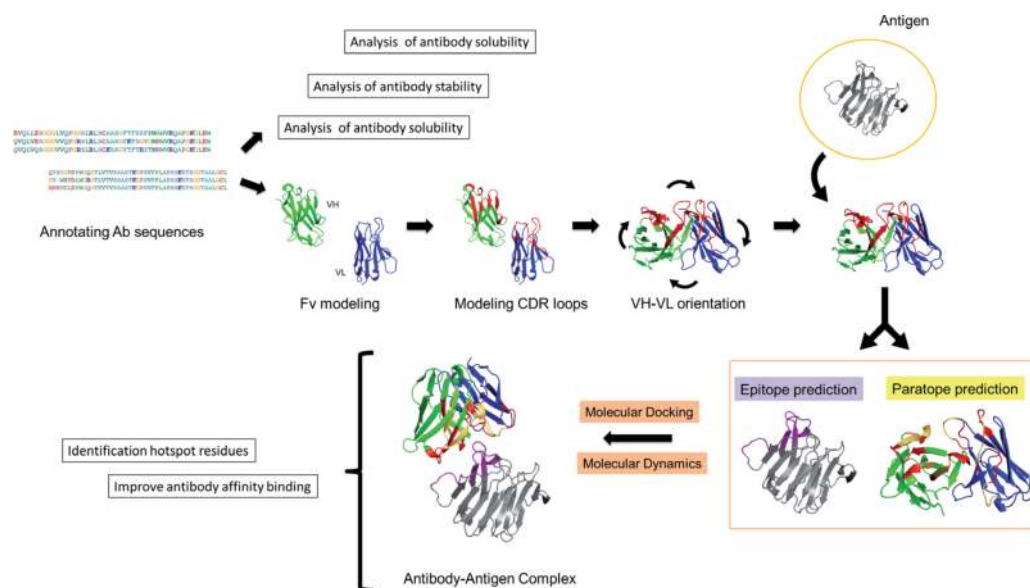
Figure 1.
Schematic representation of the computational approaches flows used in the antibody design and characterization. Methods to in silico *antibody analyses can be summarized into annotating the sequence, modeling the framework, modeling the CDR loops, and optimizing the VH-VL orientation, followed by predicting the paratope, the epitope, and the antibody–antigen contact residues. Mostly, the antibody design is performed, aiming the identification of hot spots residues suitable for mutagenesis to improve the antibody binding affinity to the target and antibody biophysical properties, such as stability, solubility. From the annotated sequence, it is possible to carry out the aggregation and solubility analyses and predict the glycosylation sites. These analyses can serve as verification steps throughout the process of antibody design and characterization.*

known antigen–antibody complexes on PDB. Paratome can also identify positions in the framework region that might contribute to antigen recognition [20, 67]. However, this tool does not provide information on the specific residues directly involved in the binding, relevant to antibody engineering experiments, such as *in silico* affinity maturation [69].

Statistical approaches such as Antibody i-Patch [21] utilizes structural information of both the antibody and antigen to generate its paratope prediction. This software assigns a score to each residue indicating how likely they are to be in contact with a given antigen. The higher score implies that the residue is more likely to form part of the paratope, information useful in guiding mutations in the artificial affinity maturation process. The tool considers the structure of both antibody and antigen. It should return more bespoke results, generating more accurate antigen-binding residue predictions.

Recently, machine learning approaches overcame Paratome and Antibody i-Patch. The PRediction Of AntiBody Contacts (proABC) [69] is a random forest algorithm, based on a machine-learning method, which uses the antibody sequences (eliminating the need for a 3D structure), the hypervariable loop canonical forms and lengths [70], and the germline family [71] as features to predict which residues of an antibody are involved in recognizing its cognate antigen. The prediction includes the nature of their contacts, distinguishing between hydrogen bonds, hydrophobic, and other non-bonded interactions. The proABC-2 [22] is an update of the original random-forest antibody paratope predictor, which uses the same set of features, but based on a deep learning framework, thus generating improved predictions and, as a consequence, increasing the success rate and quality of the docked models.

The Parapred [23] was the first algorithm based on modern deep learning for paratope prediction. This method only requires the amino acid sequence of a CDR and four adjacent residues as its input, without a full sequence, homology model, crystal structure, or antigen information. Its predictions improve the speed and accuracy of a rigid docking algorithm. The AG-Fast-Parapred [72] is an outperform of Parapred, which for the first time, provides antigen information in an in-depth paratope predictor.

Computational predictors of paratopes can provide valuable information to guide the modeling of antibody–antigen complexes. They will enable the accurate identification of residues that are the most important in determining the antibody's activity, leaving other residue positions as potential mutation sites, open to exploring other molecular characteristics by engineering.

3.2 Epitope prediction

The epitope is the antigen region in contact with the antibody in an antibody–antigen complex. Accurate identification of an epitope is a substantial step in characterizing the function of an antibody, helps predict possible cross-reactivity, and understand antibody mechanisms of action. The gold standard to determine the antibody epitope and the paratope is the 3D structure of the antibody–antigen complex by X-ray crystallography. Adding to these methods, peptide array, peptide ELISAs, phage display, expressed fragments, partial proteolysis, mass spectrometry, and mutagenesis analyses are also experimental methods applied to identify antibody epitopes. However, those assays can be expensive, time-consuming, and their outcome is uncertain [73].

Computational methods serve as an alternative to identify antibody epitopes (**Table 1**) [74]. Methods for computational B-cell epitope prediction can be categorized into the sequence and structure-based methods; the former focus on

identifying contiguous stretches of primary amino acid sequence to predict linear epitopes. In contrast, the latter takes into account their 3D structure to predict conformational epitopes. The first *in silico* B-cell epitope prediction methods focused on amino acid properties within a sequence, such as hydrophobicity, hydrophilicity, or antigenicity [75–79]. They aim to identify propensities and patterns of a set of residues on the antigen capable of binding to an antibody [29]. However, in many epitope prediction methods, it lacks information on the cognate antibody resulting in limited practical use since the epitopes predicted are generical. Therefore, antibody-specific epitope prediction methods later replaced these approaches.

The first antibody-specific epitope prediction method was suggested in 2007 by Rapberger et al. [80]. Some approaches, such as ASEP [81], BEPAR [82], ABEpar [83], and PEASE [25], are antibody-specific epitope prediction methods that do not require antibody structure. The PEASE (Predicting Antibody-Specific Epitopes) method is based on a machine-learning model and utilizes the sequence of the antibody, in the absence of structural information. It evaluates a pair score for all combinations of residues from the antibody CDR and residues from the surface-exposed region of antigen. The predictions are provided both at the residue level and as patches on the antigen structure using antibody–antigen contact preferences and other properties computed from the antibody sequence and antigen structure or sequence. The EpiPred [24] is an antibody-specific epitope prediction method that identifies the epitope region on the antigen combining conformational matching of the antibody–antigen structures and a specific antibody–antigen score. Patches on the antigen structure are ranked according to how likely they are to be the epitope. This method aims to generate epitope predictions specific for a given antibody to facilitate docking.

The most recent approaches, such as MabTope [84] and the method suggested by Jespersen et al. [85], are docking-based prediction methods of the epitope. The MabTope methodology integrates both a docking-based prediction method and experiment steps. MabTope involves three phases; in the first, docking the antibody on its target to generate possible conformations of the antigen–antibody complex (docking poses); secondly, ranking these docking poses with the design of the peptides predicted to be part of the epitope; and last, experimental validation procedures based on these peptides. The method suggested by Jespersen et al. combines geometric and physicochemical features correlated in paratope-epitope interactions with statistical and machine learning algorithms. This method can identify the cognate antigen target for a given antibody, besides the antibody target for a given antigen.

Several B-cell epitope databases were developed over the last decades, compiling validated information of the experimentally annotated B-cell epitopes. The Immune Epitope Database (IEDB) [26] is a multifaceted database that includes epitope sequence and structure, source antigen, the organism from which the epitope is derived, and details of the experiments describing recognition of the epitope. IEDB provides tools to predict linear B-cell epitopes based on sequence characteristics of the antigen [27, 28], and also to predict B-cell epitopes from protein structure, using methods based on solvent-accessible surfaces, such as DiscoTope [29, 30] and ElliPro [31]. The database Epitome [32] compiles a collection of antibody–antigen complex structures, describes the residues (on antigen and antibody CDRs) involved in the interactions, and provides information concerning specific structural characteristics of the binding regions.

The epitope information from the B-cell epitope databases can evaluate existing epitope prediction methods and develop new and better algorithms for prediction. The identification or prediction of epitopes might be useful as an information for more sophisticated computational antibody design methods, such as antibody–antigen docking.

3.3 Antibody–antigen docking

The paratope and epitope prediction methods can offer useful information on antibody–antigen recognition by identifying a subset of residues involved in antigen–antibody interface formation. However, they do not provide information about the specific pairwise relations between the residues on the antibody and the antigen. This issue can be dealt with antibody–antigen docking, a specialized application of the broader field of molecular docking [86].

Molecular docking tools (**Table 1**) allow predicting the best binding interface of two interacting proteins. Different docking algorithms have been developed over the years to predict the 3D structure of biological complexes, and they typically involve two steps: sampling and scoring. In the sampling step, the conformational space surveys for thousands of possible complex conformations ('decoys'); in the scoring, the decoys are ranked using scoring functions, which sort the decoys to identify or predict the models that are closer to the native conformation (lowest energy structure). The sampling strategy applied during the simulation is used to classify the docking methods. The global docking algorithms do not consider any previous information about the binding interfaces and perform an exhaustive search of the interaction space. The local or integrative docking approaches, on the other hand, use the available experimental data or predicted information about the binding interface to drive the sampling during the docking [87].

There are three types of docking: rigid-body docking, partial flexible docking, and flexible docking [88, 89]. Most protein–protein docking algorithms perform rigid-body docking, which means that both binding partners are kept inflexible, as rigid molecules, hindering the exploration of conformational degrees of freedom during the binding. These methods are based on the fast Fourier transform search algorithm [90] and usually are applied when the structures are complementary [89]. Examples of used rigid-body docking software are ClusPro [35], ZDOCK [36], and PatchDock [37]. ClusPro is an antibody specific docking, unlike ZDOCK and PatchDock. In partial flexible docking, the antibody remains rigid, while the antigen is flexible [89]. One of the docking tools that applied this concept is AutoDock [38]. AutoDockFR [39] also allows partial flexibility of the antibody. However, removing the conformational limitations can improve the binding site identification, since, in most situations, protein flexibility is a crucial factor to be considered [91]. Therefore, flexible docking involves both interacting molecules as flexible structures. FLIPDock [40], Swarmdock [41], SnugDock [42], and HADDOCK [92–94] are examples of these approaches. SnugDock and HADDOCK allow some flexibility alongside chains and the backbone during a refinement stage. Snugdock is the first antibody specific docking to apply flexibility to the target antibody resulting in flexible binding interfaces, which can compensate for the errors caused by homology modeling [42].

The docking approaches depend on the 3D structures of the components. For antibodies, modeling methods can generate reasonably accurate structures [10, 95, 96]. Since these methods cannot compete with the reliability of crystallography-derived structures, the performances of docking methods are continuously evaluated by the Critical Assessment of Predicted Interactions (CAPRI) experiment [97, 98].

Although there are many successful cases in predicting the protein–protein complexes, docking of antibody–antigen complexes is still challenging [99–101] due to the inherent properties of their interfaces [102, 103]. As the improvement of predicting antibody–antigen interaction methods, we expect that the results of paratope prediction, epitope prediction, and antibody–antigen docking methods would offer a valuable, fast and economical alternative to obtain reliable information about which to base rational antibody design decisions (**Figure 1**).

3.4 *In silico* affinity maturation

Recent advances in computational prediction of the 3D structure of an antibody–antigen complex stimulated the development of *in silico* methods for redesigning antibodies to improve their biophysical properties, such as binding affinity. These computational methods can screen a large number of variants in a virtual library, in a short timeframe and a cost-effective manner, and select the one most optimized, based on a better understanding of antibody–antigen interactions and structural analysis through different algorithms.

The availability of crystal structures of antibody–antigen complexes is an essential factor in achieving computational antibody affinity maturation. However, when the crystal structures of the complexes are not available, as seen above, many modeling software can predict the 3D structure of the antibody–antigen complex [104]. When we use molecular docking for this purpose, it is possible to identify residues involved in intermolecular interactions and select candidate residues that can be mutated to improve antibody affinity [99, 102–105].

The prediction of binding affinities usually utilizes energy functions, such as physics-based force fields or knowledge-based statistical potentials derived from the structural database, to estimate changes in the free energy of an antibody–antigen complex with a focus on getting the global minimum energy conformation [106]. Some algorithms and methods identify the lowest energy function of two-body interactions through changes made in the amino acid sequence or the rotameric state of an amino acid [107, 108]. Computational tools, such as molecular dynamics, simulate the dynamic behavior of antibody structures, and provide alternative candidates that can be evaluated by further experimental assessments [89, 109]. Also, some tools can identify hotspot residues on protein interfaces, for which mutation to alanine strongly attenuates binding, and calculate the values for the change in the binding energy of the protein complex upon mutation [110–112]. These platforms are useful to study the effect of a particular amino acid on the binding affinity of an antibody–antigen complex.

Computational affinity maturation usually focuses on residues in the CDRs. However, as we learned in previous sections of this chapter, some residues in the framework can also play a role in the binding affinity and maintain the canonical conformations of antibodies. Although some mutations in noninteracting regions resulted in improved binding affinity [113, 114], the strategies to modify the CDR to increase antibody affinity are highlighted. Some examples of *in silico* affinity maturation of antibodies performed comprehensive computational CDR mutagenesis targeting all residues in CDRs or CDR H3 [115–118]. There are also examples of monitoring all interactions between the 3D structure of an antibody and its cognate target to determine the most relevant CDR residues in the binding by considering their stabilizing energies, inter and intra molecules distance, bonds formation or breakage, and overall complex stability [119].

These techniques still present deviations from the experimental data; however, they demonstrate that in some scenarios, computational approaches alone can be used for affinity maturation, decreasing the timeframe and costs of antibody engineering.

4. Analyses of mAbs' properties (solubility, stability, aggregation, chemical degradation, glycosylation)

In vitro antibody affinity maturation frequently results in a destabilizing process, needing compensatory modifications for preserving the thermodynamic stability of mAbs [120]. Emerging *in silico* tools are significant resources to promote the balance

between affinity and stability during antibody engineering (**Table 1**). Before proceeding to available resources to deal with the destabilizing process, we should mention two types of stability in antibodies: physical and chemical. The physical stability of a protein is related to conformational changes and also to its colloidal stability. Concerning the conformational changes, we relate the free energy (ΔG) of the protein in its unfolded and folded-state, and the folded-state should present less energy than the unfolded state ($G_{folded-state} < G_{unfolded-state}$) [121]. One of the *in silico* methods used to investigate folded and unfolded-state energies was mentioned earlier, e.g., molecular dynamics.

Among the numerous *in silico* tools for predicting conformational stability, DeepDDG [45] proved to be quite efficient compared to eight other methods (**Table 1**). DeepDGG is a machine learning method trained from 5444 experimental data. This tool allows the calculation of the energy difference between the mutated protein and its native state. This calculation allows us to observe whether the proposed mutations, for example, for an improvement in affinity, cause structure destabilization. Experimentally, the conformational change of a protein it is accessed indirectly through its melting temperature (Tm), and it can be measured by different experimental techniques, such as scanning calorimetry (DSC), differential scanning fluorometry (DSF), and circular dichroism (CD). The changes between folded and unfolded-state can be reversible, unlike the process known as aggregation, related to colloidal stability.

Although aggregation is different from solubility, the solubility of a molecule is usually calculated for aggregation prediction. In computational chemistry, aggregation and solubility are commonly treated as the same parameter. The aggregation tendency of some mAbs that could impair their efficacy might be prevented through aggregation-prone regions (APRs) analyses. APR assays rely on the hydrophobicity scales and residues' charge annotations. Among several predictors of solubility and APRs for proteins, it is possible to highlight two endeavors successfully applied to antibodies: Wang et al. [122] combined tools to predict APRs in commercial mAbs. They found similar aggregation-prone motifs among commercial and non-commercial antibodies, without correlation with 3D structures.

In 2011, Agrawal et al. [123] compared several aggregation prediction tools demonstrating their usefulness in drug discovery and development, especially when screening a large number of molecules by fast and low cost *in silico* assays. Recently, Raybould et al. [53] launched Therapeutic Antibody Profiler (TAP), a web application that compares candidates' sequences with natural antibody sequences, as natural antibodies are assumed to display favorable biophysical properties. TAP, notably, depends on the previous data of clinical-stage antibody therapeutics (CST). So, the robustness of this method is directly affected by the input improvement of the CST database. One modern and elegant approach drove the development of AggreRATE-Disc [54], a machine learning-based tool that can predict, within the sequences, mutations that can promote or mitigate aggregation. Although *in silico* tools can highlight sequences with aggregation issues, they do not substitute experimental assays; however, they can be managed, reducing the totality of necessary tests. These tools and databases help the screening steps across the development/discovery of new therapeutic drugs.

Regarding the chemical stability of antibodies, it is possible to mention the degradation by chemical modification of amino acids, such as asparagine (Asn) deamidation, aspartate (Asp) isomerization, methionine (Met) oxidation, and lysine (Lys) glycation [124, 125]. The IgGs are commonly N-glycosylated at Asp297 residue in each Fc-CH$_2$ domain [126]. These Fc N-glycan are associated with correct

folding, stability, aggregation, immunogenicity, and serum half-life of the mAbs. The conformational changes at the CH_2 antibody portion by multiple hydrophobic and polar non-covalent interactions harnesses the Fc binding to preferences of binding to C1q and FcγRs [126]. There are no specific mAb glycosylation's webtools. Still, some web platforms (**Table 1**) designed to predict glycosylation sites on human protein sequences could also be useful for mAbs. The IgGs have a conservative N-glycan site; consequently, it needs attention in the engineering process that could accidentally create or remove a glycosylation site and interfering in the mAb chemical stability. In other instances, the glycosylation site is intentionally removed.

To evaluate any possible glycosylation spots, the NetNGlyc 1.0 [56] predicts N-glycosylation sites in human proteins using a trained neural network to distinguish between the acceptor and non-acceptor residue sequences. The N-GlyDE is a two-stage N-glycan prediction tool trained by the human proteome datasets. An algorithm generates a score between N-glycosylation proteins and non-N-linked glycoproteins in the first step. In the second stage, the prediction uses a support vector machine to evaluate if each asparagine-Xaa-serine/threonine (being Xaa different to proline) sequence can be glycosylated [57]. Further, the GlycoSiteAlign [59] is a tool that aligns amino acid sequences regarding its glycosylation site using the GlyConnect databank. This tool can be useful to compare a high number of mAbs sequences derived from different clones or expression conditions.

In a linear amino acid sequence of an antibody, it is possible to find numerous regions prone to modification. However, one must note that many of these regions may be buried due to the molecule conformation. Therefore, a conformational study is essential to highlight the residues liable to the chemical change. Chemical stability is generally based on statistical analysis derived from experiments or databases available in the literature, although some computational methods are being used [124, 127–133]. Statistics-based methods depend on data from previous experiments and provide valuable information about the behavior of proteins, being excellent guides during the development of new antibodies.

Currently, there are tools to predict the most varied protein characteristics. Many of them are free for academic purposes (**Table 1**). A difficulty still faced during the development of an antibody lies in the complexity of details and how one parameter influences another. For example, modifications to improve binding affinity may interfere with the stability of the molecule or even generate/remove a glycosylation spot. In the same way, a structural change for stability can impair binding affinity. There has been an immeasurable evolution of *in silico* methods, allowing analyzes to be carried out more quickly and at a lower cost than traditional experimental methods.

5. Highlights

Advances in bioinformatics allow us to outline different strategies in the discovery of new therapeutic antibodies. There has been significant progress in online tools in recent years, and probably the refinement of the techniques will be increased, bringing more accurate and reliable results.

Online platforms can present a long wait and execution times. The use of those platforms requires a good internet connection, and also a robust computer for analysis and treatment of the generated data.

Bioinformatics is a notably promising field, and indeed, has a prominent place on the innovation frontier.

Acknowledgments

Author details

Tania M. Manieri[1*], Carolina G. Magalhaes[1], Daniela Y. Takata[1], João V. Batalha-Carvalho[1] and Ana M. Moro[1,2]

1 Biopharmaceuticals Lab, Butantan Institute, Sao Paulo, Brazil

2 iii/INCT, São Paulo, Brazil

*Address all correspondence to: tania.manieri@butantan.gov.br

IntechOpen

References

[1] Rees AR. Understanding the human antibody repertoire. MAbs. 2020;12(1):1729683.

[2] Boyd SD, Joshi SA. High-Throughput DNA Sequencing Analysis of Antibody Repertoires. Microbiol Spectr. 2014;2(5).

[3] Schroeder HW, Jr. Similarity and divergence in the development and expression of the mouse and human antibody repertoires. Dev Comp Immunol. 2006;30(1-2):119-35.

[4] Berman HM, Westbrook J, Feng Z, Gilliland G, Bhat TN, Weissig H, et al. The Protein Data Bank. Nucleic Acids Res. 2000;28(1):235-42.

[5] Lefranc MP, Pommie C, Kaas Q, Duprat E, Bosc N, Guiraudou D, et al. IMGT unique numbering for immunoglobulin and T cell receptor constant domains and Ig superfamily C-like domains. Developmental and comparative immunology. 2005;29(3):185-203.

[6] Ye J, Ma N, Madden TL, Ostell JM. IgBLAST: an immunoglobulin variable domain sequence analysis tool. Nucleic Acids Res. 2013;41(Web Server issue):W34-40.

[7] Brochet X, Lefranc MP, Giudicelli V. 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. 2008;36(Web Server issue):W503-8.

[8] Norman RA, Ambrosetti F, Bonvin A, Colwell LJ, Kelm S, Kumar S, et al. Computational approaches to therapeutic antibody design: established methods and emerging trends. Brief Bioinform. 2019.

[9] Marks C, Nowak J, Klostermann S, Georges G, Dunbar J, Shi J, et al. Sphinx: merging knowledge-based and ab initio approaches to improve protein loop prediction. Bioinformatics. 2017;33(9):1346-53.

[10] Leem J, Dunbar J, Georges G, Shi J, Deane CM. ABodyBuilder: Automated antibody structure prediction with data–driven accuracy estimation. mAbs. 2016;8:1259-68.

[11] Klausen MS, Anderson MV, Jespersen MC, Nielsen M, Marcatili P. LYRA, a webserver for lymphocyte receptor structural modeling. Nucleic Acids Res. 2015;43(W1):W349-55.

[12] Schritt D, Li SL, Rozewicki J, Katoh K, Yamashita K, Volkmuth W, et al. Repertoire Builder: high-throughput structural modeling of B and T cell receptors. Mol Syst Des Eng. 2019;4(4):761-8.

[13] Lyskov S, Chou FC, Conchuir SO, Der BS, Drew K, Kuroda D, et al. Serverification of molecular modeling applications: the Rosetta Online Server that Includes Everyone (ROSIE). PLoS One. 2013;8(5):e63906.

[14] Leem J, Georges G, Shi J, Deane CM. Antibody side chain conformations are position-dependent. Proteins. 2018;86(4):383-92.

[15] Krivov GG, Shapovalov MV, Dunbrack RL, Jr. Improved prediction of protein side-chain conformations with SCWRL4. Proteins. 2009;77(4):778-95.

[16] Ryu J, Lee M, Cha J, Laskowski RA, Ryu SE, Kim DS. BetaSCPWeb: side-chain prediction for protein structures using Voronoi diagrams and geometry prioritization. Nucleic Acids Research. 2016;44(W1):W416-W23.

[17] Choi Y, Deane CM. FREAD revisited: Accurate loop structure

prediction using a database search algorithm. Proteins-Structure Function and Bioinformatics. 2010;78(6):1431-40.

[18] Jacobson MP, Pincus DL, Rapp CS, Day TJF, Honig B, Shaw DE, et al. A hierarchical approach to all-atom protein loop prediction. Proteins-Structure Function and Bioinformatics. 2004;55(2):351-67.

[19] Wong WK, Georges G, Ros F, Kelm S, Lewis AP, Taddese B, et al. SCALOP: sequence-based antibody canonical loop structure annotation. Bioinformatics. 2019;35(10):1774-6.

[20] Kunik V, Ashkenazi S, Ofran Y. Paratome: An online tool for systematic identification of antigen-binding regions in antibodies based on sequence or structure. Nucleic Acids Research. 2012;40:521-4.

[21] Krawczyk K, Baker T, Shi J, Deane CM. Antibody i-Patch prediction of the antibody binding site improves rigid local antibody-antigen docking. Protein Engineering, Design and Selection. 2013;26:621-9.

[22] Ambrosetti F, Olsen TH, Olimpieri PP, Jiménez-García B, Milanetti E, Marcatilli P, et al. proABC-2: PRediction Of AntiBody Contacts v2 and its application to information-driven docking. Bioinformatics. 2020.

[23] Liberis E, Velickovic P, Sormanni P, Vendruscolo M, Lio P. Parapred: Antibody paratope prediction using convolutional and recurrent neural networks. Bioinformatics. 2018;34:2944-50.

[24] Krawczyk K, Liu X, Baker T, Shi J, Deane CM. Improving B-cell epitope prediction and its application to global antibody-antigen docking. Bioinformatics. 2014;30:2288-94.

[25] Sela-Culang I, Ashkenazi S, Peters B, Ofran Y. PEASE: Predicting B-cell epitopes utilizing antibody sequence. Bioinformatics. 2015;31:1313-5.

[26] Vita R, Mahajan S, Overton JA, Dhanda SK, Martini S, Cantrell JR, et al. The Immune Epitope Database (IEDB): 2018 update. Nucleic Acids Res. 2019;47(D1):D339-D43.

[27] Chou PY, Fasman GD. Prediction of the secondary structure of proteins from their amino acid sequence. Adv Enzymol Relat Areas Mol Biol. 1978;47:45-148.

[28] Jespersen MC, Peters B, Nielsen M, Marcatili P. BepiPred-2.0: improving sequence-based B-cell epitope prediction using conformational epitopes. Nucleic Acids Research. 2017;45(W1):W24-W9.

[29] Kringelum JV, Lundegaard C, Lund O, Nielsen M. Reliable B Cell Epitope Predictions: Impacts of Method Development and Improved Benchmarking. PLoS Computational Biology. 2012;8.

[30] Andersen PH, Nielsen M, Lund O. Prediction of residues in discontinuous B-cell epitopes using protein 3D structures. Protein Science. 2006;15:2558-67.

[31] Ponomarenko J, Bui HH, Li W, Fusseder N, Bourne PE, Sette A, et al. ElliPro: A new structure-based tool for the prediction of antibody epitopes. Bmc Bioinformatics. 2008;9:1-8.

[32] Schlessinger A, Ofran Y, Yachdav G, Rost B. Epitome: database of structure-inferred antigenic epitopes. Nucleic acids research. 2006;34.

[33] Bui HH, Sidney J, Li W, Fusseder N, Sette A. Development of an epitope conservancy analysis tool to facilitate the design of epitope-based diagnostics and vaccines. Bmc Bioinformatics. 2007;8.

[34] Dhanda SK, Vaughan K, Schulten V, Grifoni A, Weiskopf D, Sidney J, et al. Development of a novel clustering tool for linear peptide sequences. Immunology. 2018;155(3):331-45.

[35] Kozakov D, Hall DR, Xia B, Porter KA, Padhorny D, Yueh C, et al. The ClusPro web server for protein-protein docking. Physiology & behavior. 2017;12:255-78.

[36] Pierce BG, Wiehe K, Hwang H, Kim BH, Vreven T, Weng ZP. ZDOCK server: interactive docking prediction of protein-protein complexes and symmetric multimers. Bioinformatics. 2014;30(12):1771-3.

[37] Schneidman-Duhovny D, Inbar Y, Nussinov R, Wolfson HJ. PatchDock and SymmDock: Servers for rigid and symmetric docking. Nucleic Acids Research. 2005;33:363-7.

[38] Morris GM, Huey R, Lindstrom W, Sanner MF, Belew RK, Goodsell DS, et al. AutoDock4 and AutoDockTools4: Automated docking with selective receptor flexibility. J Comput Chem. 2009;30(16):2785-91.

[39] Ravindranath PA, Forli S, Goodsell DS, Olson AJ, Sanner MF. AutoDockFR: Advances in Protein-Ligand Docking with Explicitly Specified Binding Site Flexibility. PLoS Comput Biol. 2015;11(12):e1004586.

[40] Zhao Y, Sanner MF. FLIPDock: docking flexible ligands into flexible receptors. Proteins. 2007;68(3):726-37.

[41] Torchala M, Moal IH, Chaleil RAG, Fernandez-Recio J, Bates PA. SwarmDock: A server for flexible protein-protein docking. Bioinformatics. 2013;29:807-9.

[42] Sircar A, Gray JJ. SnugDock: Paratope structural optimization during antibody-antigen docking compensates for errors in antibody homology models. PLoS Computational Biology. 2010;6.

[43] van Zundert GCP, Rodrigues J, Trellet M, Schmitz C, Kastritis PL, Karaca E, et al. The HADDOCK2.2 Web Server: User-Friendly Integrative Modeling of Biomolecular Complexes. J Mol Biol. 2016;428(4):720-5.

[44] Ramirez-Aportela E, Lopez-Blanco JR, Chacon P. FRODOCK 2.0: fast protein-protein docking server. Bioinformatics. 2016;32(15):2386-8.

[45] Cao H, Wang J, He L, Qi Y, Zhang JZ. DeepDDG: Predicting the Stability Change of Protein Point Mutations Using Neural Networks. Journal of chemical information and modeling. 2019;59(4):1508-14.

[46] Pucci F, Kwasigroch JM, Rooman M. SCooP: an accurate and fast predictor of protein stability curves as a function of temperature. Bioinformatics. 2017;33(21):3415-22.

[47] Cheng J, Randall AZ, Sweredoski MJ, Baldi P. SCRATCH: a protein structure and structural feature prediction server. Nucleic Acids Res. 2005;33(Web Server issue):W72-6.

[48] Cheng JL, Randall A, Baldi P. Prediction of protein stability changes for single-site mutations using support vector machines. Proteins-Structure Function and Bioinformatics. 2006;62(4):1125-32.

[49] Folkman L, Stantic B, Sattar A, Zhou Y. EASE-MM: Sequence-Based Prediction of Mutation-Induced Stability Changes with Feature-Based Multiple Models. J Mol Biol. 2016;428(6):1394-405.

[50] Quan L, Lv Q, Zhang Y. STRUM: structure-based prediction of protein stability changes upon single-point mutation. Bioinformatics. 2016;32(19):2936-46.

[51] Pires DE, Ascher DB, Blundell TL. mCSM: predicting the effects of mutations in proteins using

graph-based signatures. Bioinformatics. 2014;30(3):335-42.

[52] Pandurangan AP, Ochoa-Montano B, Ascher DB, Blundell TL. SDM: a server for predicting effects of mutations on protein stability. Nucleic Acids Res. 2017;45(W1):W229-W35.

[53] Raybould MIJ, Marks C, Krawczyk K, Taddese B, Nowak J, Lewis AP, et al. Five computational developability guidelines for therapeutic antibody profiling. Proc Natl Acad Sci U S A. 2019;116(10):4025-30.

[54] Rawat P, Kumar S, Michael Gromiha M. An in-silico method for identifying aggregation rate enhancer and mitigator mutations in proteins. Int J Biol Macromol. 2018;118 (Pt A):1157-67.

[55] Peng HP, Lee KH, Jian JW, Yang AS. Origins of specificity and affinity in antibody-protein interactions. Proc Natl Acad Sci U S A. 2014;111(26):E2656-65.

[56] Blom N, Sicheritz-Ponten T, Gupta R, Gammeltoft S, Brunak S. Prediction of post-translational glycosylation and phosphorylation of proteins from the amino acid sequence. Proteomics. 2004;4(6):1633-49.

[57] Pitti T, Chen CT, Lin HN, Choong WK, Hsu WL, Sung TY. N-GlyDE: a two-stage N-linked glycosylation site prediction incorporating gapped dipeptides and pattern-based encoding. Sci Rep. 2019;9(1):15975.

[58] Cooper CA, Gasteiger E, Packer NH. GlycoMod--a software tool for determining glycosylation compositions from mass spectrometric data. Proteomics. 2001;1(2):340-9.

[59] Gastaldello A, Alocci D, Baeriswyl JL, Mariethoz J, Lisacek F. GlycoSiteAlign: Glycosite Alignment Based on Glycan Structure. J Proteome Res. 2016;15(10):3916-28.

[60] Julenius K. NetCGlyc 1.0: prediction of mammalian C-mannosylation sites. Glycobiology. 2007;17(8):868-76.

[61] Steentoft C, Vakhrushev SY, Joshi HJ, Kong Y, Vester-Christensen MB, Schjoldager KT, et al. Precision mapping of the human O-GalNAc glycoproteome through SimpleCell technology. EMBO J. 2013;32(10):1478-88.

[62] North B, Lehmann A, Dunbrack RL, Jr. A new clustering of antibody CDR loop conformations. J Mol Biol. 2011;406(2):228-56.

[63] Noel D, Bernardi T, Navarro-Teulon I, Marin M, Martinetto JP, Ducancel F, et al. Analysis of the individual contributions of immunoglobulin heavy and light chains to the binding of antigen using cell transfection and plasmon resonance analysis. Journal of immunological methods. 1996;193(2):177-87.

[64] Xu JL, Davis MM. Diversity in the CDR3 region of V(H) is sufficient for most antibody specificities. Immunity. 2000;13(1):37-45.

[65] Morea V, Tramontano A, Rustici M, Chothia C, Lesk AM. Conformations of the third hypervariable region in the VH domain of immunoglobulins. J Mol Biol. 1998;275(2):269-94.

[66] Weitzner BD, Jeliazkov JR, Lyskov S, Marze N, Kuroda D, Frick R, et al. Modeling and docking of antibody structures with Rosetta. Nat Protoc. 2017;12(2):401-16.

[67] Kunik V, Peters B, Ofran Y. Structural consensus among antibodies defines the antigen binding site. PLoS Computational Biology. 2012;8.

[68] Esmaielbeiki R, Krawczyk K, Knapp B, Nebel JC, Deane CM. Progress and challenges in predicting protein interfaces. Briefings in Bioinformatics. 2016;17:117-31.

[69] Olimpieri PP, Chailyan A, Tramontano A, Marcatili P. Prediction of site-specific interactions in antibody-antigen complexes: The proABC method and server. Bioinformatics. 2013;29:2285-91.

[70] Chothia C, Lesk AM. Canonical structures for the hypervariable regions of T cell αβ receptors. J Mol Biol. 1987;196:901-17.

[71] Schatz DG, Swanson PC. V(D) J recombination: Mechanisms of initiation. Annual Review of Genetics. 2011;45:167-202.

[72] Deac A, Veličković P, Sormanni P. Attentive Cross-Modal Paratope Prediction. Journal of Computational Biology. 2019;26:536-45.

[73] Xu XL, Sun J, Liu Q, Wang XJ, Xu TL, Zhu RX, et al. Evaluation of spatial epitope computational tools based on experimentally-confirmed dataset for protein antigens. Chinese Science Bulletin. 2010;55:2169-74.

[74] El-Manzalawy Y, Honavar V. Recent advances in B-cell epitope prediction methods. Immunome Research. 2010;6:S2.

[75] Hopp TP, Woods KR. Prediction of protein antigenic determinants from amino acid sequences. Proceedings of the National Academy of Sciences of the United States of America. 1981;78:3824-8.

[76] Kolaskar AS, Tongaonkar PC. A semi-empirical method for prediction of antigenic determinants on protein antigens. FEBS Letters. 1990;276:172-4.

[77] Pellequer JL, Westhof E. PREDITOP: A program for antigenicity prediction. Journal of Molecular Graphics. 1993;11:204-10.

[78] Pellequer JL, Westhof E, Van Regenmortel MHV. Predicting location of continuous epitopes in proteins from their primary structures. Methods in Enzymology. 1991;203:176-201.

[79] Pellequer JL, Westhof E, Van Regenmortel MHV. Correlation between the location of antigenic sites and the prediction of turns in proteins. Immunology Letters. 1993;36:83-99.

[80] Rapberger R, Lukas A, Mayer B. Identification of discontinuous antigenic determinants on proteins based on shape complementarities. Journal of Molecular Recognition. 2007;20:113-21.

[81] Soga S, Kuroda D, Shirai H, Kobori M, Hirayama N. Use of amino acid composition to predict epitope residues of individual antibodies. Protein Engineering, Design and Selection. 2010;23:441-8.

[82] Zhao L, Li J. Mining for the antibody-antigen interacting associations that predict the B cell epitopes. BMC Structural Biology. 2010;10:1-13.

[83] Zhao L, Wong L, Li J. Antibody-specified B-cell epitope prediction in line with the principle of context-awareness. IEEE/ACM Transactions on Computational Biology and Bioinformatics. 2011;8:1483-94.

[84] Bourquard T, Musnier A, Puard V, Tahir S, Ayoub MA, Jullian Y, et al. MAbTope: A Method for Improved Epitope Mapping. The Journal of Immunology. 2018;201:3096-105.

[85] Jespersen MC, Mahajan S, Peters B, Nielsen M, Marcatili P.

Antibody specific B-cell epitope predictions: Leveraging information from antibody-antigen protein complexes. Frontiers in Immunology. 2019;10:1-10.

[86] Pagadala NS, Syed K, Tuszynski J. Software for molecular docking: a review. Biophysical Reviews. 2017;9:91-102.

[87] Ciemny M, Kurcinski M, Kamel K, Kolinski A, Alam N, Schueler-Furman O, et al. Protein–peptide docking: opportunities and challenges. Drug Discovery Today. 2018;23:1530-7.

[88] Guedes IA, de Magalhaes CS, Dardenne LE. Receptor-ligand molecular docking. Biophys Rev. 2014;6(1):75-87.

[89] Salmaso V, Moro S. Bridging molecular docking to molecular dynamics in exploring ligand-protein recognition process: An overview. Frontiers in Pharmacology. 2018;9:1-16.

[90] Katchalski-Katzir E, Shariv I, Eisenstein M, Friesem AA, Aflalo C, Vakser IA. Molecular surface recognition: Determination of geometric fit between proteins and their ligands by correlation techniques. Proceedings of the National Academy of Sciences of the United States of America. 1992;89:2195-9.

[91] Kotev M, Soliva R, Orozco M. Challenges of docking in large, flexible and promiscuous binding sites. Bioorg Med Chem. 2016;24(20):4961-9.

[92] de Vries SJ, van Dijk ADJ, Krzeminski M, van Dijk M, Thureau A, Hsu V, et al. HADDOCK versus HADDOCK: New features and performance of HADDOCK2.0 on the CAPRI targets. Proteins. 2007;69:726-33.

[93] Dominguez C, Boelens R, Bonvin AMJJ. HADDOCK: A protein-protein docking approach based on biochemical or biophysical information. Journal of the American Chemical Society. 2003;125:1731-7.

[94] van Zundert GCP, Melquiond ASJ, Bonvin AMJJ. Integrative Modeling of Biomolecular Complexes: HADDOCKing with Cryo-Electron Microscopy Data. Structure. 2015;23(5):949-60.

[95] Weitzner BD, Jeliazkov JR, Lyskov S, Marze N, Kuroda D, Frick R, et al. Modeling and docking of antibody structures with Rosetta. Nature Protocols. 2017;12:401-16.

[96] Yamashita K, Ikeda K, Amada K, Liang S, Tsuchiya Y, Nakamura H, et al. Kotai Antibody Builder: Automated high-resolution structural modeling of antibodies. Bioinformatics. 2014;30:3279-80.

[97] Janin J, Henrick K, Moult J, Eyck LT, Sternberg MJE, Vajda S, et al. CAPRI: A critical assessment of PRedicted interactions. Proteins: Structure, Function and Genetics. 2003;52:2-9.

[98] Méndez R, Leplae R, De Maria L, Wodak SJ. Assessment of blind predictions of protein-protein interactions: Current status of docking methods. Proteins: Structure, Function and Genetics. 2003;52:51-67.

[99] Pedotti M, Simonelli L, Livoti E, Varani L. Computational docking of antibody-antigen complexes, opportunities and pitfalls illustrated by influenza hemagglutinin. International Journal of Molecular Sciences. 2011;12:226-51.

[100] Ponomarenko JV, Bourne PE. Antibody-protein interactions: Benchmark datasets and prediction tools evaluation. BMC Structural Biology. 2007;7.

[101] Vajda S. Classification of protein complexes based on docking difficulty.

Proteins: Structure, Function and Genetics. 2005;60:176-80.

[102] Lo Conte L, Chothia C, Janin J. The atomic structure of protein-protein recognition sites. J Mol Biol. 1999;285:2177-98.

[103] Sela-Culang I, Kunik V, Ofran Y. The structural basis of antibody-antigen recognition. Frontiers in Immunology. 2013;4:1-13.

[104] Kuroda D, Shirai H, Jacobson MP, Nakamura H. Computer-aided antibody design. Protein Engineering, Design and Selection. 2012;25:507-21.

[105] Barderas R, Desmet J, Timmerman P, Meloen R, Casal JI. Affinity maturation of antibodies assisted by in silico modeling. Proceedings of the National Academy of Sciences of the United States of America. 2008;105:9029-34.

[106] Moal IH, Moretti R, Baker D, Fernández-Recio J. Scoring functions for protein-protein interactions. Current Opinion in Structural Biology. 2013;23:862-7.

[107] Desmet J, De Maeyer M, Hazes B, Lasters I. The dead-end elimination theorem and its use in protein side-chain positioning. Nature. 1992;356:539-42.

[108] Metropolis N, Rosenbluth AW, Rosenbluth MN, Teller AH, Teller E. Equation of state calculations by fast computing machines. The Journal of Chemical Physics. 1953;21:1087-92.

[109] González MA. Force fields and molecular dynamics simulations. Collection SFN. 2011;12:169-200.

[110] Schymkowitz J, Borg J, Stricher F, Nys R, Rousseau F, Serrano L. The FoldX web server: An online force field. Nucleic Acids Research. 2005;33:382-8.

[111] Tuncbag N, Keskin O, Gursoy A. HotPoint: Hot spot prediction server for protein interfaces. Nucleic Acids Research. 2010;38:402-6.

[112] Zhu X, Mitchell JC. KFC2: A knowledge-based hot spot prediction method based on interface solvation, atomic density, and plasticity features. Proteins: Structure, Function and Bioinformatics. 2011;79:2671-83.

[113] Fukunaga A, Tsumoto K. Improving the affinity of an antibody for its antigen via long-range electrostatic interactions. Protein Engineering, Design and Selection. 2013;26:773-80.

[114] Yanaka S, Moriwaki Y, Tsumoto K, Sugase K. Elucidation of potential sites for antibody engineering by fluctuation editing. Scientific Reports. 2017;7:1-10.

[115] Clark LA, Boriack-Sjodin PA, Eldredge J, Fitch C, Friedman B, Hanf KJM, et al. Affinity enhancement of an in vivo matured therapeutic antibody using structure-based computational design. Protein Science. 2006;15:949-60.

[116] Kiyoshi M, Caaveiro JMM, Miura E, Nagatoishi S, Nakakido M, Soga S, et al. Affinity improvement of a therapeutic antibody by structure-based computational design: Generation of electrostatic interactions in the transition state stabilizes the antibody-antigen complex. PLoS ONE. 2014;9:1-9.

[117] Lippow SM, Wittrup KD, Tidor B. Computational design of antibody-affinity improvement beyond in vivo maturation. Nat Biotechnol. 2007;25(10):1171-6.

[118] Willis JR, Sapparapu G, Murrell S, Julien JP, Singh V, King HG, et al. Redesigned HIV antibodies exhibit enhanced neutralizing potency and breadth. Journal of Clinical Investigation. 2015;125:2523-31.

[119] Ebrahimi Z, Asgari S, Ahangari Cohan R, Hosseinzadeh R,

Hosseinzadeh G, Arezumand R. Rational affinity enhancement of fragmented antibody by ligand-based affinity improvement approach. Biochemical and Biophysical Research Communications. 2018;506:653-9.

[120] Julian MC, Li L, Garde S, Wilen R, Tessier PM. Efficient affinity maturation of antibody variable domains requires co-selection of compensatory mutations to maintain thermodynamic stability. Sci Rep. 2017;7:45259.

[121] Kazlauskas R. Engineering more stable proteins. Chemical Society reviews. 2018;47(24):9026-45.

[122] Wang X, Das TK, Singh SK, Kumar S. Potential aggregation prone regions in biotherapeutics: A survey of commercial monoclonal antibodies. MAbs. 2009;1(3):254-67.

[123] Agrawal NJ, Kumar S, Wang X, Helk B, Singh SK, Trout BL. Aggregation in protein-based biotherapeutics: computational studies and tools to identify aggregation-prone regions. J Pharm Sci. 2011;100(12):5081-95.

[124] Kumar S, Plotnikov NV, Rouse JC, Singh SK. Biopharmaceutical Informatics: supporting biologic drug development via molecular modelling and informatics. The Journal of pharmacy and pharmacology. 2018;70(5):595-608.

[125] Lu X, Nobrega RP, Lynaugh H, Jain T, Barlow K, Boland T, et al. Deamidation and isomerization liability analysis of 131 clinical-stage antibodies. MAbs. 2019;11(1):45-57.

[126] Li W, Zhu ZY, Chen WZ, Feng Y, Dimitrov DS. Crystallizable Fragment Glycoengineering for Therapeutic Antibodies Development. Frontiers in Immunology. 2017;8.

[127] Yang R, Jain T, Lynaugh H, Nobrega RP, Lu X, Boland T, et al. Rapid assessment of oxidation via middle-down LCMS correlates with methionine side-chain solvent-accessible surface area for 121 clinical stage monoclonal antibodies. MAbs. 2017;9(4):646-53.

[128] Agrawal NJ, Dykstra A, Yang J, Yue H, Nguyen X, Kolvenbach C, et al. Prediction of the Hydrogen Peroxide-Induced Methionine Oxidation Propensity in Monoclonal Antibodies. J Pharm Sci. 2018;107(5):1282-9.

[129] Plotnikov NV, Singh SK, Rouse JC, Kumar S. Quantifying the Risks of Asparagine Deamidation and Aspartate Isomerization in Biopharmaceuticals by Computing Reaction Free-Energy Surfaces. J Phys Chem B. 2017;121(4):719-30.

[130] Jia L, Sun Y. Protein asparagine deamidation prediction based on structures with machine learning methods. PLoS One. 2017;12(7):e0181347.

[131] Delmar JA, Wang J, Choi SW, Martins JA, Mikhail JP. Machine Learning Enables Accurate Prediction of Asparagine Deamidation Probability and Rate. Mol Ther Methods Clin Dev. 2019;15:264-74.

[132] Sankar K, Hoi KH, Yin Y, Ramachandran P, Andersen N, Hilderbrand A, et al. Prediction of methionine oxidation risk in monoclonal antibodies using a machine learning method. MAbs. 2018;10(8):1281-90.

[133] Aledo JC, Canton FR, Veredas FJ. A machine learning approach for predicting methionine oxidation sites. Bmc Bioinformatics. 2017;18(1):430.