Available online at www.sciencedirect.com

ScienceDirect

Current Opinion in Structural Biology

Computational design of antibodies

Sharon Fischman[1] and Yanay Ofran[1,2]

Antibody design aims to create new antibodies with biological activity that can be used in therapy and research. Traditional methods for antibody discovery, such as animal immunization and large-scale library screening, generate antibodies that bind to the target of interest, but do not necessarily have the desired functional effect. Computational methods can be utilized as a means to guide the search for biologically relevant antibodies, focusing on specificity and affinity determinants to target a particular region of the antigen. Such an approach would allow for the design of epitope-specific antibodies that will have the desired effect on the function of the targeted protein.

Addresses
[1] Biolojic Design LTD, 9 Bialik Street, Givat Shmuel, Israel
[2] The Goodman Faculty of Life Sciences, Nanotechnology Building, Bar Ilan University, Ramat Gan 52900, Israel

Corresponding author: Ofran, Yanay (yanay@ofranlab.org)

Current Opinion in Structural Biology 2018, **51**:156–162

This review comes from a themed issue on **Engineering & design**

Edited by **Giovanna Ghirlanda** and **Ivan Korendovych**

https://doi.org/10.1016/j.sbi.2018.04.007

Introduction

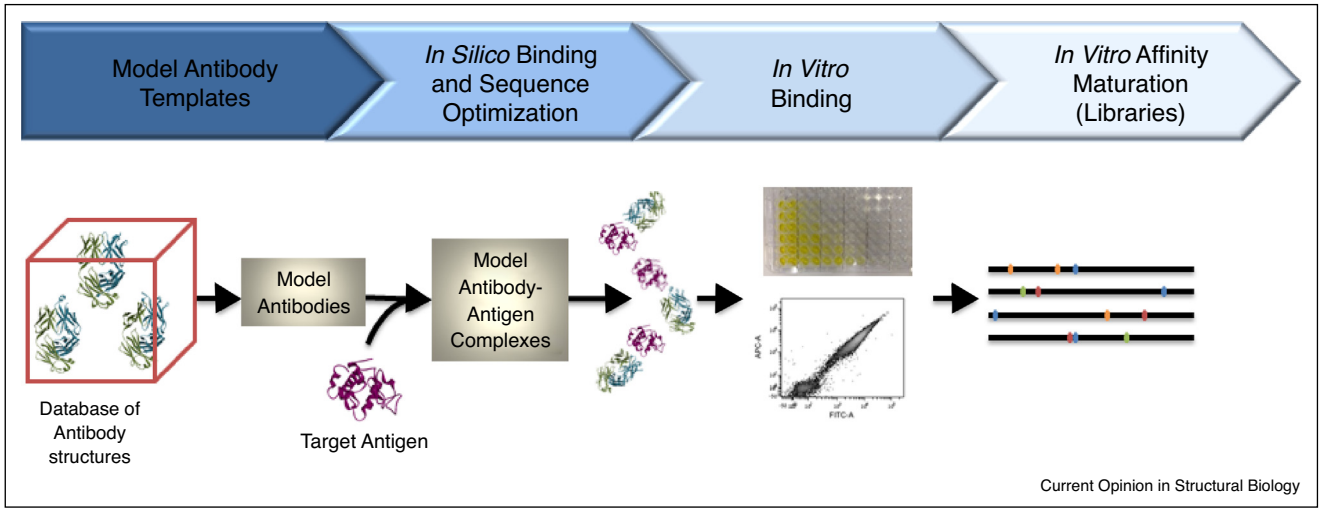
Antibodies are the fastest growing class of therapeutics [1]. However, despite tremendous discovery efforts, existing technologies fail to generate biologically active antibodies against many of the most promising targets. The essential goal of Ab design, particularly in the context of drug design, is to design a novel antibody that has a biological effect. However, most approaches focus getting a specific binder to the target, not on eliciting a desired biological activity. Immunization and screening of large libraries can be employed to obtain binders to a target of interest. Different approaches to the design of such libraries, including restricted codons [2] or combinations of germline H and L chain genes [3,4], have succeeded in producing antibodies with novel binding specificities and in some cases, biological activity [5]. These methods, however, select for the tightest binders, typically to immunodominant epitopes, precluding the discovery of antibodies with lower affinities that may bind to other, functionally relevant sites. Targeting specific sites within a target antigen, for example, those known to agonize or antagonize a biological pathway, remains a challenge in antibody design.

While large-scale, general purpose libraries may yield some functional antibodies, the size of the haystack in which these needles hide makes it difficult to identify them by meticulous functional screening of thousands of binders. Computational approaches offer another route to antibody design. The general scheme of current methods for the computational design of antibodies is presented in Figure 1. A first step toward identifying an antibody that binds the antigen is to model the 3-D structure of candidate antibodies, as well as the structure of the antibody–antigen complex. These antibodies are then tested experimentally for binding, and if necessary, are improved via *in vitro* affinity maturation. Better understanding of the structural basis of antigen binding by antibodies is a key to the success of this approach [6]. Here, we review the current state of computational technologies for antibody design, and suggest how new computational approaches can be applied to design libraries that are more likely to yield biologically active antibodies.

Modeling antibodies and antibody–antigen complexes

Structure-based computational protein design in general, and antibody design in particular, relies heavily on quality three-dimensional structural data for both the template for design (in this case, the antibody), the desired target (in this case, the antigen), and their complex. Antibody modeling has advanced to the state where the majority of the antibody variable domain can be modeled reliably. The success in modeling is in part due to structurally canonical conformations of most CDRs [7]. However obtaining accurate models of the variable CDR H3 and the relative orientation of the H and L chains, arguably the most important elements in determining binding, remains a challenge [8•] (for a review of antibody modeling and challenges see [9]). Among other reasons, this is due to the unique conformation of H3 in different Abs [10]. As H3 comprises part of the H-L interface, modeling both of these regions is interdependent. H3 modeling can be improved by implementing geometric constraints that describe a conserved structural kink [11] (For a review of H3 modeling see [12]). Addressing both H3 modeling and VH-VL orientation, Marze *et al.* [13] demonstrate improvements to antibody modeling accuracy by utilizing multiple templates of VH-VL orientation in addition to CDR grafting with RosettaAntibody [14]. Deane and colleagues implement a Random Forest classifier to

Figure 1



Computational design of antibodies — general scheme. Current methods for computational antibody design begin with modeling an antibody and an antibody–antigen complex. Selected antibody sequences are tested experimentally for antigen binding, for example, either with a soluble antigen or a cell-expressed antigen, and binders are further optimized by affinity maturation methods.

identify specific sequence positions that characterize the VH-VL orientation as a series of torsion and bend angles affecting the possible degrees of freedom, to improve orientation prediction [15,16].

However, even when a reliable model for the antibody is obtained, modeling the Ab-Ag complex is a difficult task. The community wide critical assessment of protein interactions (CAPRI), which assesses the performance of computational tools for modeling complexes, demonstrates this difficulty. In its recent experiment [17], 67 research teams using state-of-the-art methods attempted to model 20 complexes. The teams submitted >20 000 models (i.e. an average of 1000 models per complex), and yet for six out of the 20 complexes there was not a single model that was deemed 'acceptable' in its quality (e.g. identifying correctly 50% or more of the interface contacts). The success of docking that is based on models of the subunits, is even poorer [18]. These difficulties are encountered in antibody–antigen docking as well [19].

Importantly, even when complex modeling successfully generates a correct model among its best models, there is no straightforward way of telling which one it is. Consequently, attempts to design an antibody that are based on modeling the 3-D structure of the complex, cannot rely on a single model, and hence require the synthesis of dozens, sometimes even hundreds, of different sequences, hoping that one of them binds.

Selecting models as a basis for computational design, however, is only the first step. Methods to predict changes in the free energy of mutants are then used to improve antibodies or to introduce cross-reactivity. These methods use either crystal structures or models of the antibody–antigen complexes [20–22] as their starting point. A study by Sirin *et al.* [23] highlights the limited performance of these methods. This study used a large dataset of mutants to compare the experimentally determined and the computationally predicted effects of mutations on binding free energies of antibody–antigen complexes. The computational methods tested included those based on statistical potentials as well as all-atom force-fields. They conclude that some of the computational methods perform reasonably well in identifying mutations with a large effect on binding, but the problem of identifying mutations with moderate or small effects is still unresolved. Another study [24] found that using consensus scoring of some of these programs can improve the identification of mutations that weaken binding. However, the study did not distinguish between mutations that improve affinity and mutations that were neutral. A recent study by Clark *et al.* [25] on a small number of antibodies shows that predictions of binding energy changes correlate with experimental alanine scanning data. However, the authors conclude that their tool is not yet a "robust, automated protocol . . . suitable for application to an arbitrary protein–protein interaction." Taken together, the results of these studies demonstrate challenges that still exist for computational design of antibodies: predicting whether, and how, the designed proteins are going to interact is a major challenge and predicting which mutations can improve affinity is not easier. This is why existing approaches require many experimental attempts and large libraries for improving preliminary binders.