Multidisciplinary Reviews

REVIEW ARTICLE
Published Online: December 17, 2022
https://doi.org/10.31893/multirev.2022012

Antibody engineering by computational approach

Mujahed I. Mustafa[a*]

[a]Department of Biotechnology, College of Applied and Industrial Sciences, University of Bahri, Khartoum, Sudan.

*Corresponding author: mujahedimustafa@gmail.com

Abstract In the pre-era of synthetic antibodies, pharmaceutical companies depended on finding novel drugs from medicinal plants and other traditional resources; while at present, technological advances in biology, computer, and robotics give researchers the ability to rewrite and edit DNA to synthesize extensive sets of drug candidates; these novel and improved candidates serve the basis for creating another library of drug candidates and so on until we find the right biomolecule for the disease of interest. All these technologies combined to synthesize therapeutic antibodies for many types of cancer, autoimmune diseases, and infectious diseases that can address diseases much more readily to very rapidly get therapeutics into patients so that we can potentially impact disease. The antibody mechanism is recognized, binds to disease cells, and pinpoints the immune system to attack those cells effectively. Now a day, they depend on a computational approach to guide and accelerate the process of antibodies engineering by a combination of a selection system and the use of high-throughput data acquisition and analysis to build and construct populations of next-generation antibodies that are thermo-stable, non-immunogenic as possible, and to be administered to many humans as possible. In this review, I will discuss the latest in silico methods for antibody engineering.

Keywords: next generation antibodies, novel drugs, synthetic immunology

1. Introduction

Antibody research has been one of the main focus areas in modern molecular biology (Si et al 2018; Geering and Fussenegger 2014; Singh 2020), such as next-generation antibodies with an estimated global pharmaceutical market of USD 140 billion by 2024 (Ecker, Jones, and Levine 2014). It's well-known that antibodies are "Y"-shaped proteins produced by the immune system to detect antigens, including proteins from viruses, bacteria, and abnormal cells (Stanfield and Wilson 2014). Each antibody selectively recognizes and binds to its target antigen, triggering an immune response that destroys the cells displaying the antigen (Stanfield and Wilson 2014).

Paul Ehrlich first described the antibody as "magic bullets" in 1908 (Bosch and Rosich 2008). In 1952, antibodies isolated from human blood were tested to see if they could treat immune system deficiencies. In 1975 Köhler and Milstein developed a technique allowing large numbers of identical antibodies to be produced using a single cell line or clone obtained from mice (Alkan 2019); antibodies produced in this way are known as monoclonal antibodies (mAbs) (Freysd'ottir 2000). Building on this understanding, scientists harness this method to develop therapeutic antibodies, which could be engineered to recognize and binds to specific antigens on cells involved in the development of various diseases; this was a breakthrough in the development of medicine. The first approval mAbs was in 1986 for use of preventing kidney transplant rejection (Webster et al. 2017). As early mAbs were developed in cells derived from mice, their potential was restricted due to the fact patient's immune system will recognize the antibodies themselves self-as foreign molecules and thus inactivate or destroy them. By the 1990s chimeric monoclonal antibodies were developed, meaning that around two-third of the molecule contains humans rather than mouse protein, making them less likely to be attacked by the immune system ('Monoclonal antibody, chimeric' 2004). The first chimeric monoclonal antibody was approved in 1994 to treat cardiovascular disease (Faulds and Sorkin 1994). Further advances led to humanized monoclonal antibodies with less than 10% mouse protein and eventually to fully human monoclonal antibodies. Adalimumab is the first fully human monoclonal antibody approved in 2002 for treating Rheumatoid arthritis (Szechiński et al 2008).

The antibody holds two chains (VL and VH), each consisting of several domains. The antigen-binding site is located in each chain's 'variable' domains (Stanfield and Wilson 2014). The remainder of the variable domains is structurally well conserved at the backbone level. Therefore, the main focus of antibody design is dedicated to predicting the conformations of the CDR loops from their sequences (Locker and Herr 2020; Stanfield and Wilson 2014).

In recent years, antibody engineering by in silico method is nearly worldwide accepted as a critical approach to guide and accelerate the process of engineering antibodies by a combination of a selection system and the use of high-throughput data acquisition and analysis to build and construct populations of next-generation antibodies that are thermo-stable, non-

immunogenic as possible, and to be administered to many humans as possible (Almagro et al. 2014). Engineering antibodies by computational approach are mainly for combining high-throughput data accusation, high-through genomic sequencing, and robotized high-throughput selection along with synthetic natural DNA sources to carefully consider designs to accelerate the rate of antibodies discovery as drugs (Almagro et al. 2014) (Figure. 1a). The history of that has been initially a pioneering technique that uses high-throughput sequencing to investigate why are libraries don't produce more hits, and how to improve the quality of these hits, to make them better drugs, which are shelf-stable, thermo-stable, aggregation-resistant, and non-immunogenic (Kumar et al 2017; Datta 2020; Viola et al 2018).

The increasing knowledge of sequence-structure co-relation in antibodies and the advancement in the in silico approach particularly in protein modeling has facilitated the growth of in silico methods that can aid in engineering antibodies for desired alterations (Almagro et al 2014; Almagro et al 2011; Teplyakov et al 2014). The main target area of a biological engineer is complementarity-determining regions (CDRs) because antibodies function through them (Kuroda et al 2012). The recent advancement in synthetic biology has led to manipulating the amino acid sequences by enhancing the affinity of CDRs (Farhadi et al 2017). Nevertheless, antibody engineering concerns the trade-off between binding affinity and other properties. To conquer such an issue, a standard approach to enhance the properties of antibodies is random mutagenesis based on in vitro libraries (Kuroda and Tsumoto 2020).

Now a day, it's a relatively easy task to engineer binding affinities with other properties through an in vitro, library-based method by elevating temperature and controlling solution conditions during the selection process. However, due to the advances in computational power, deep sequencing, and artificial intelligence, in silico strategies are becoming an alternate approach in engineering antibodies (Zhao et al. 2018). Yet, such predictions are not completely pleasing for good two explanations, there is no straightforward protocol to design by, and the precision of in silico approach is not as good as that of in vitro libraries approach because the biophysical properties of biomolecules are not well understood (Böldicke 2018; Kuroda and Tsumoto 2020). In this review, I discuss the frontier in antibody engineering by computational approach. Among several properties, I focused on antibodies' stability, viscosity, and immunogenicity, all of which have gathered much devotion in antibody engineering by computational approach (Figure 1a).

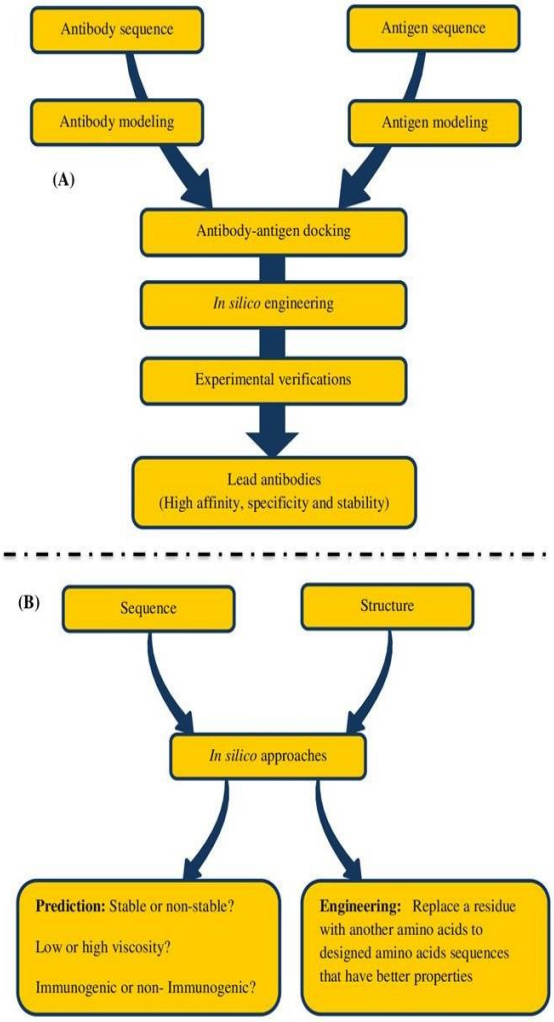

Figure 1 (A) Development procedure of antibodies engineering and the roles of the computational approach. (B) The main workflow of the computational approach.

2. Methodology

The search strategies used to obtain information for the review were based on searching by certain keywords such as “Antibody Engineering”, “Physicochemical and Biological Properties of Antibodies”, “Prediction and Designing of Immunogenicity of Antibodies”.

3. Physicochemical and biological properties of antibodies

3.1. Stability

One of the most important properties in antibody drug discovery is protein stability, of which there are two types, physical and chemical stabilities. Commonly, protein stability can be categorized into conformational stability and colloidal stability. Proteins are only slightly stable, and proteins in solution are in equilibrium between folded and unfolded conformations (Figure 2). Many homology modeling (Nichols et al 2015; Perchiacca and Tessier 2012; Perchiacca, Bhattacharya, and Tessier 2012) have been sufficient for developing ability goals (excluding cases when antibodies present rare CDR in whichever heavy chain or light chain loops). Protein aggregate is a state occurring when a protein in folded-unfolded stability may accumulate into an oligomeric condition which is an irreversible process (Figure 2); on the protein surface, short hydrophobic fragments designated as aggregation-prone regions (APRs), researchers have considered dictates the aggregation tendencies of proteins, furthermore, any mutation in on this region can intensely affect the rate of aggregation.

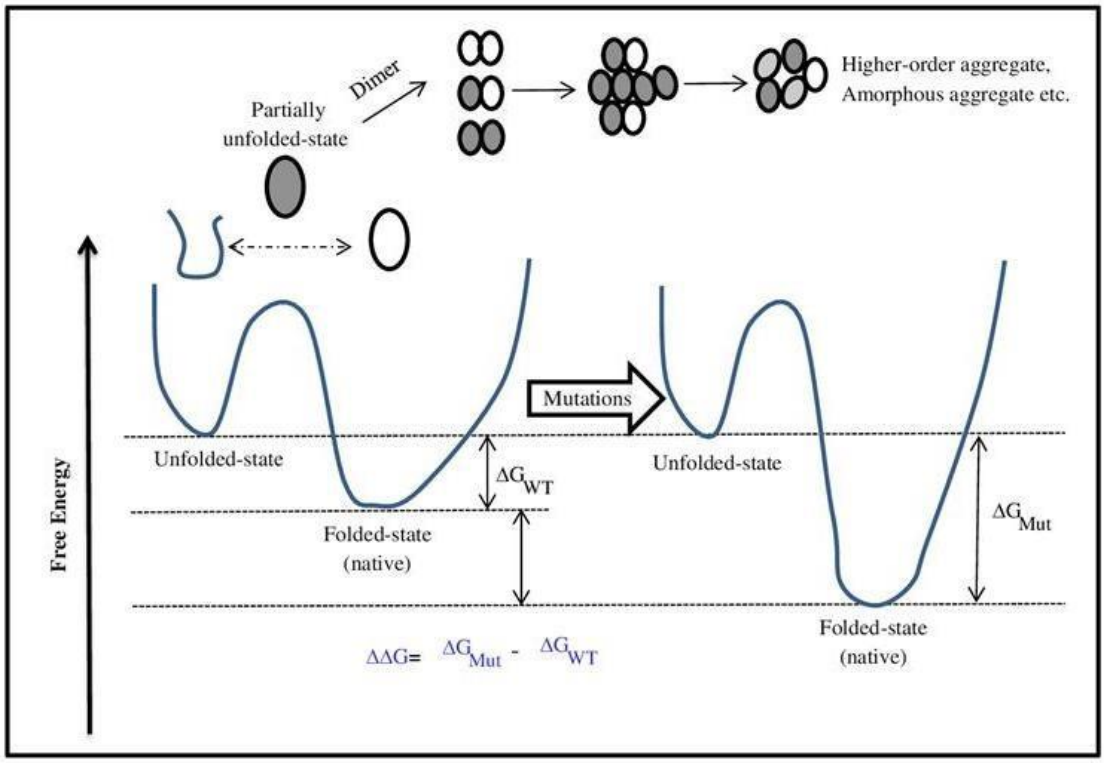

Figure 2 Equilibrium between folded and unfolded proteins and its relation to protein stability. *Source:* (Kuroda et al 2020).

3.2. Viscosity

Another important property in antibody discovery is antibody viscosity, due to its applied consequences concerning formulation and administration. The behavior of the concentration-dependent viscosity of antibodies depends on pairwise interactions or self-association, which further leads to higher-order intermolecular interactions (Kuroda and Tsumoto 2020) (Figure 2).

3.3. Immunogenicity

The term “immunogenicity” refers to the patient's immune response against the proteins. Immunogenicity evaluation usually carries out by animal testing, which costs a lot of time, cost, and effort. Therefore, using the *in silico* approach minimizes time and costs by facilitating sequence alignments to check the amino acid similarities of antibodies and targets of interest (Pham and Meng 2020).

4. Prediction and designing of physicochemical properties of antibodies

4.1. Overview of computational prediction and designing

Several *in silico* approaches for physicochemical properties prediction, such as viscosity, protein stability, and immunogenicity, have been established to enable predictive protein engineering. The input and output of these tools are summarized in (Figure 1b). Generally, these prediction models can be classified into two groups: statistical and physics-based. Physics-based approaches, unique, do not rely on any experimental data to achieve predictions because they are based on the laws of physics. While statistical methods depend on statistical data extracted from experimental information, the prediction accuracy is proportional to the quality of artificial intelligence used to train the prediction approaches (Dahiyat and Mayo 1997; Kuhlman and Baker 2000) (Table 1).

4.2. Prediction and designing of viscosity

Viscosity prediction has gathered a fair share of devotion as a target of interest in antibody design by in silico method. Some studies have revealed that high-antibody viscosities are better associated with negative than positive charges (Kramer et al 2012). In some cases, aggregation forms due to the unusual viscosity behavior of an antibody. This behavior led some scientists to suggest that the viscosity behaviors of antibodies are driven by their corresponding amino acid sequences. Keeping in mind that the constant domains of antibodies are highly conserved, the differences in behaviors probably result from differences in the variable regions (Li et al 2014).

Table 1 Features used in the machine learning model DeepDDG for predicting

Approaches	Categories	Features
Machine Learning Model DeepDDG for Predicting Thermo-stability	Sequence-based features	Amino acid compositions
		Protein length
		Secondary structures
	Structure-based features	Solubility-dependent statistical potentials
		Secondary structures
		Solvent-accessible surface area
Machine Learning Model SOLart for Predicting Solubility	Sequence-based features	Amino acid compositions
		Protein length
		Secondary structures
	Structure-based features	Solubility-dependent statistical potentials
		Secondary structures
		Solvent-accessible surface area

4.3. Prediction of colloidal stability and solubility

One of the hot zones in the research field is protein aggregation, specifically regards the capability to develop protein therapeutics. Theoretically, solubility and aggregation are different marvels, because they are reversible and irreversible, respectively (Agostini et al 2011; Sormanni et al 2014). Currently, many *in silico* tools are accessible to predict aggregation rates and aggregation-prone regions (Table 1).

4.3. Prediction of chemical stability

In therapeutic antibody prediction, many in silico approaches have been suggested to evaluate the chemical stability (Sydow et al 2014; Aledo et al 2017; Kumar et al 2017). One of the most frequent degradation procedures is the chemical modification of Asparagine and Aspartic acid residues, which have the same degradation pathway (Kumar et al 2017). There is

no protocol to predict such degradation. There are statistical and physics-based approaches; the first is experimental data-dependent, combined with homology modeling (Sydow et al 2014); the second is superior to the first by quantum-mechanical calculations dependent without the need for experimental data (Plotnikov et al 2017).

4.5. Prediction and designing of immunogenicity of antibodies

Predicting high physicochemical properties such as high stability that a chief complete unfolding is a critical factor in therapeutic antibodies for good immune responses in patients. Due to the recent advances in computational immunology, various in silico tools facilitated prediction and reduced protein immunogenicity as much as possible (Abhinandan and Martin 2007; Gao et al 2013; Olimpieri et al 2014; Choi et al 2017) (Table 2).

Table 2 *In silico* approaches to evaluate, predict, and reduce the immunogenicity of antibodies.

Assessment of Immunogenicity	Description	URL
T20 score analyzer	Humanness score based on sequence identity to the top 20 matched human antibody sequences	https://dm.lakepharma.com/bioinformatics/
SHAB	Humanness score based on sequence identity to human antibody sequences	http://www.bioinf.org.uk/abs/shab/
Humanization of antibodies	Framework template search followed by CDR grafting with back mutations	http://www.biocomputing.it/tabhu

5. Discussion

In the last decade, antibodies engineering by the computational approach has drawn so much attention by guiding and accelerating the process of antibodies engineering by a combination of a selection system and the use of high-throughput data acquisition and analysis to build and construct populations of next-generation antibodies that are thermo-stable, non-immunogenic as possible, and to be administered to many humans as possible (Kumar et al 2017; Datta 2020; Viola et al 2018). It's strongly suggested that homology modeling-based antibody techniques are dependable enough to be used in high-throughput, sequence-based computational analysis (Kuroda et al 2008). Yet, CDR-H3 structure prediction is still challenging. Since antibodies function has been highlighted on CDR-H3diversity, to engineer a better functional antibody. CDR-H3 structure prediction approaches and antibody-antigen complexes must be developed. Another critical area that needs to be improved in the line of antibody engineering by computational approach is protein flexibility.

6. Final considerations

In conclusion, despite no *in silico* unified antibody engineering protocol, the computational approach remains an indispensable method for antibody designing. There is no doubt that the combination of in silico and library-based approaches will assist in the production of therapeutics antibodies to assist the immune system in the battle against all types of life-threatening and debilitating disorders.

Conflict of Interest

I declare that there are no conflicts of interest.

Funding

The current review did not receive any financial support.

References

Abhinandan KR, Martin ACR (2007) Analyzing the "Degree of Humanness" of Antibody Sequences. J Mol Biol 369:852-62.

Agostini F, Vendruscolo F, Tartaglia G (2011) Sequence-Based Prediction of Protein Solubility. J Mol Biol 421:237-41.

Aledo J, Francisco R, Veredas F (2017). A machine learning approach for predicting methionine oxidation sites. BMC Bioinformatics 18.

Sefik A (2019) The Discovery of Monoclonal Antibodies (On Georges Köhler). Allergy 74.

Almagro J, Beavers M, Hernandez-Guzman F, Maier J, Shaulsky J, Butenhof K (2011) Antibody Modeling Assessment. Proteins 79:3050-66.

Almagro, Juan, Alexey Teplyakov, Jinquan Luo, Raymond Sweet (2014) Second antibody modeling assessment (AMA-II). Proteins 82.

Böldicke T (2018) Antibody Engineering.

Bosch F, Laia R (2008) The contributions of Paul Ehrlich to pharmacology: a tribute on the occasion of the centenary of his Nobel Prize. Pharmacology 82:171-79.

Choi Y, Verma D, Griswold K, Bailey-Kellogg C (2017) EpiSweep: Computationally Driven Reengineering of Therapeutic Proteins to Reduce Immunogenicity While Maintaining Function.

Dahiyat B, Mayo S (1997) De Novo Protein Design: Fully Automated Sequence Selection. Science 278:82-87.

Datta J (2020) Molecular modelling & simulation: a review from molecular informatics. doi: 10.6084/m9.figshare.12287606.v4

Ecker D, Jones S, Levine H (2014) The Therapeutic Monoclonal Antibody Market. mAbs 7.

Farhadi T, Fakharian A, Hashemian S (2017) Affinity Improvement of a Humanized Antiviral Antibody by Structure-Based Computational Design. International Journal of Peptide Research and Therapeutics, 25.

Faulds D, Sorkin EM (1994) Abciximab (c7E3 Fab). A review of its pharmacology and therapeutic potential in ischaemic heart disease. Drugs 48:583-98.

Freysd'ottir J (2000) Production of Monoclonal Antibodies. Methods in molecular medicine 40:267-79.

Gao S, Huang K, Tu H, Adler A (2013) Monoclonal antibody humanness score and its applications. BMC Biotechnol 13: 55.

Geering, B, Fussenegger M (2014) Synthetic immunology: Modulating the human immune system. Trends Biotechnol 33.

Kramer R (2012) Toward a Molecular Understanding of Protein Solubility: Increased Negative Surface Charge Correlates with Increased Solubility. Biophysical journal 102:1907-15.

Kuhlman B, Baker D (2000) Native protein sequences are close to optimal for their structures. Proc Natl Acad Sci USA 97:10383-8.

Kumar et al (2017) Biopharmaceutical Informatics: Supporting biologic drug development via molecular modelling and informatics. Journal of Pharmacy and Pharmacology 70.

Kuroda D et al (2012) Computer-aided antibody design. Protein engineering, design & selection: PEDS 25:507-22.

Kuroda D et al (2008) Structural classification of CDR-H3 revisited: A lesson in antibody modeling. Proteins 73:608-20.

Kuroda D, Tsumoto K (2020) Engineering Stability, Viscosity, and Immunogenicity of Antibodies by Computational Design. Journal of Pharmaceutical Sciences 109.

Li L et al (2014) Concentration Dependent Viscosity of Monoclonal Antibody Solutions: Explaining Experimental Behavior in Terms of Molecular Properties. Pharmaceutical research 31.

Locker K, HerrA (2020) Antibodies: Structure and Immune Effector Functions.

'Monoclonal antibody, chimeric.' in. (2004) Larry W. Moreland (ed.), Rheumatology and Immunology Therapy (Springer Berlin Heidelberg: Berlin, Heidelberg).

Nichols P et al (2015) Rational design of viscosity reducing mutants of a monoclonal antibody: Hydrophobic versus electrostatic inter-molecular interactions. mAbs 7: 212-30.

Olimpieri P et al (2014) Tabhu: Tools for antibody humanization. Bioinformatics 31.

Perchiacca J et al (2012) Aggregation-resistant domain antibodies engineered with charged mutations near the edges of the complementarity-determining regions. Protein engineering, design & selection: PEDS 25:591-602.

Perchiacca J et al (2012) Engineering Aggregation-Resistant Antibodies. Annual Review of Chemical and Biomolecular Engineeringn 3:263-86.

Pham N, Meng W (2020) Protein Aggregation and Immunogenicity of Biotherapeutics. International Journal of Pharmaceutics 585:119523.

Plotnikov N et al (2017) Quantifying Risks of Asparagine Deamidation and Aspartate Isomerization in Biopharmaceuticals by Computing Reaction Free Energy Surfaces. The journal of physical chemistry. B 121.

Si W et al (2018) Synthetic immunology: T-cell engineering and adoptive immunotherapy. Synthetic and Systems Biotechnology 3.

Singh S (2020) Systems and Synthetic Immunology. doi: 10.1007/978-981-15-3350-1

Sormanni P et al (2014) The CamSol Method of Rational Design of Protein Mutants with Enhanced Solubility. J Mol Biol 427.

Stanfield R, Wilson I (2014) Antibody Structure. Microbiology Spectrum 2.

Sydow J et al (2014) Structure-Based Prediction of Asparagine and Aspartate Degradation Sites in Antibody Variable Regions. PLoS One 9:e100736.

Szechiński J et al (2008) Adalimumab - The first fully human monoclonal antibody used in the treatment of rheumatoid arthritis. Reumatologia 46:151-58.

Teplyakov A et al (2014) Antibody modeling assessment II: Structures and Models. Proteins: Structure, Function, and Bioinformatics 82.

Viola M et al (2018) Subcutaneous delivery of monoclonal antibodies: How do we get there? Journal of Controlled Release 286.

Webster AC et al (2017) Polyclonal and monoclonal antibodies for treating acute rejection episodes in kidney transplant recipients. The Cochrane database of systematic reviews 7:CD004756-CD56.

Zhao J et al (2018) In Silico Methods in Antibody Design. Antibodies 7:22.