

# Molecular Modeling Methods in the Development of Affine and Specific Protein-Binding Agents

Shamsudin Sh. Nasaev<sup>1</sup>, Artem R. Mukanov<sup>1</sup>, Ivan V. Mishkorez<sup>1,2</sup>, Ivan I. Kuznetsov<sup>1</sup>, Iosif V. Leibin<sup>3</sup>, Vladislava A. Dolgusheva<sup>1</sup>, Gleb A. Pavlyuk<sup>1</sup>, Artem L. Manasyan<sup>1</sup>, and Alexander V. Veselovsky<sup>2,a\*</sup>

<sup>1</sup>Research & Development Department, Xelari Ltd., 121601 Moscow, Russia

<sup>2</sup>Institute of Biomedical Chemistry, 119121 Moscow, Russia

<sup>3</sup>Skolkovo Institute of Science and Technology, Skolkovo Innovation Center, 121205 Moscow, Russia

<sup>a</sup>e-mail: veselov@ibmh.msk.su

Received April 12, 2024

Revised June 12, 2024

Accepted July 11, 2024

**Abstract**—High-affinity and specific agents are widely applied in various areas, including diagnostics, scientific research, and disease therapy (as drugs and drug delivery systems). It takes significant time to develop them. For this reason, development of high-affinity agents extensively utilizes computer methods at various stages for the analysis and modeling of these molecules. The review describes the main affinity and specific agents, such as monoclonal antibodies and their fragments, antibody mimetics, aptamers, and molecularly imprinted polymers. The methods of their obtaining as well as their main advantages and disadvantages are briefly described, with special attention focused on the molecular modeling methods used for their analysis and development.

DOI: 10.1134/S0006297924080066

**Keywords:** antibody, antibody mimetics, aptamers, molecularly imprinted polymers, molecular modeling

## INTRODUCTION

Since the first mentioning of using antibodies for protein recognition in 1968, there has been an increasing interest in the development of high-affinity and specific agents for their application in diagnostics and, later, in therapy and other fields [1]. In the 1970s and 1980s, technologies for production of monoclonal antibodies have been developed, the structure of immunoglobulins have been determined, and the methods for designing chimeric, humanized, and then fully human monoclonal antibodies have been created. Antibodies have taken their place among the main specific protein-binding agents in modern biology and medicine [2]. Currently, antibodies are used as immobilizing and detecting agents in diagnostics, as well as drugs and drug delivery systems, labeling and staining reagents, research tools, etc. [3-5].

However, even modern monoclonal antibodies have their drawbacks, which limits their diversity and areas of application.

The necessity to optimize approaches for development of binding agents, in particular, the requirements for their targeted and rational design, have led to a widespread application of molecular modeling methods. These methods have demonstrated a high efficiency in studying high-affinity protein-binding agents, increasing the efficacy and reducing the cost of their development, improving the binding parameters, and enhancing the stability of such agents. Several studies have shown the possibility of rational design of affinity-binding agents based on the target structure. Molecular modeling methods are used at different stages in the development of protein-binding agents, from the creation of primary monomer libraries to system optimization and identification of molecular binding mechanisms. Both common (homology modeling, molecular docking, molecular dynamics) and specialized (optimized for a specific type of binding agent) methods can be used.

**Abbreviations:** CDRs, complementarity-determining regions; MIPs, molecularly imprinted polymers.

\* To whom correspondence should be addressed.

This review summarizes main advantages, disadvantages, and limitations of known protein-binding agents, as well as key experimental approaches to their development. Special attention is given to modern molecular modeling methods used for investigation, optimization, and rational design of high-affinity protein-binding agents in order to overcome the limitations of experimental approaches and to evaluate the modeling results.

### PROTEIN-BASED PROTEIN-BINDING AFFINE AGENTS

**Antibodies: structure, methods of production, and applications.** Antibodies are globular proteins composed of four polypeptide chains: two light chains (50-60 kDa) and two heavy chains (100-120 kDa) connected by disulfide bonds. The antigen-binding properties of an antibody are determined by the variable domain (Fv) formed by the heavy and light chains (Fig. 1a). The variable domain contains conserved regions responsible for maintaining the structure of the binding site and variable loops, or CDRs (complementarity-determining regions), that provide interaction with the antigen (Fig. 1b). Antibodies are produced by the immune cells; they are also components of B-lymphocyte receptors [6].

The first use of antibodies as high-affinity specific agents can be traced back to 1890 when E. von Behring and S. Kitasato used a serum containing polyclonal antibodies to treat diphtheria in animals. Antibodies have become the most popular antitoxins; polyclonal sera are still used nowadays. A significant break-

through in the research and application of antibodies occurred in 1975, when G. Köhler and C. Milstein proposed the hybridoma technology for producing monoclonal antibodies. Monoclonal antibodies have enabled to develop unique, reproducible methods for protein isolation and identification, protein concentration measurement, and labeling of cells based on their antigenic composition. Since 1985, monoclonal antibodies have been used in drug therapy. Today, they remain the most common protein-binding high-affinity specific agents [7].

The most frequently used method for developing monoclonal antibodies is the hybridoma technology, which involves the fusion of genetically modified myeloma cells with B lymphocytes, resulting in the formation of cells capable of continuously producing antibodies against a target protein. The hybridoma technology is costly, resource-intensive, and cannot be easily automated, partly due to the need for continuous maintenance of cell cultures in bioreactors and monitoring of genetic changes during cell proliferation [8].

A more versatile and cost-effective alternative method for obtaining antibodies is phage display. In this technique, a gene of the antibody variable domain is inserted into a bacteriophage (e.g., M13) genome. The recombinant bacteriophage is used for infecting bacteria, and the antibody fragment is expressed as a part of the virus envelope. Only bacteriophages containing the antibody fragment on the surface can bind to the corresponding antigen. The bacteriophages with a low specificity are removed using a mixture of non-target proteins. This allows to select a variable domain that can subsequently be

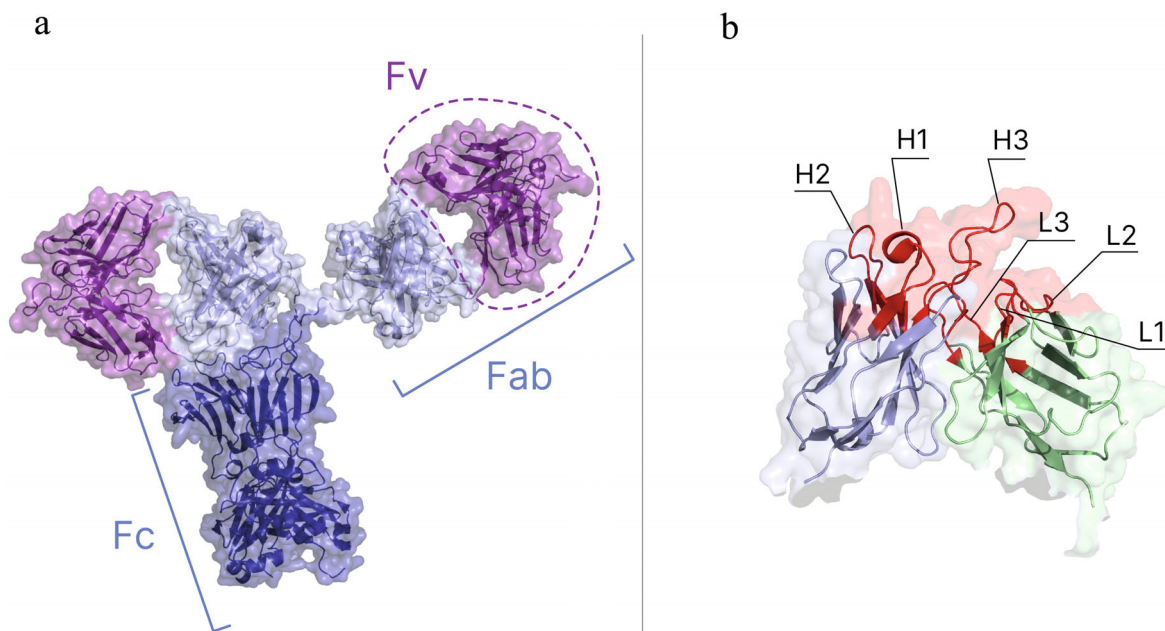


Fig. 1. The structure of immunoglobulin G (PDB: 1HZH) (a) and variable loops (CRDs) of the Fv domain (b).

modified to increase its affinity and specificity through directed mutagenesis and repeated selection [9].

Despite the existence of two methods for producing monoclonal antibodies, the costs of development and, more significantly, production are substantial, since both methods require the use of bioreactors. Consequently, the price of these antibodies can be high, limiting their application. Moreover, due to the difficulties in controlling the batch-to-batch reproducibility, the affinity, specificity, and stability of the produced antibodies often differ from those claimed by the manufacturer [10].

**Molecular modeling methods for the development and study of antibodies.** In recent years, molecular modeling methods have been actively used for optimization and development of the structure of antibodies in order to enhance their affinity, specificity, and other properties.

**Modeling of antibody structure.** Five out of six hypervariable regions (CDRs) in the variable domain exhibit a limited number of possible conformations, i.e., differ only slightly between antibodies with different specificity and affinity (RMSD,  $\sim 0.7$  Å) [11, 12]. Analysis of the antibody tertiary structures has shown that there are eight main templates for the variable region conformation – two for the structure of the light and heavy chains, five for the hypervariable loops within the CDRs (L1, L2, L3, H1, H2), and one used as a basis for modeling the H3 loop – which can vary significantly between different antibodies (Fig. 1b). The structures of most variable antibody fragments can be predicted by combining these loop conformations. Five of the six loops within the CDRs are typically modeled by the homology-based methods. The main challenge is modeling of the H3 loop and relative orientation of the H and L chains. Since H3 is a part of the HL interface, modeling of both regions is interdependent [13]. The H3 loop is built using *ab initio* methods that involve either searching for similar loops in known protein three-dimension structures or constructing the loop by sequential addition of amino acid residues, followed by optimization of the loop structure using molecular mechanics and molecular dynamics methods. Methods based on stochastic approaches, such as the Monte Carlo algorithm, have been proposed, in which a set of different geometries is generated and the most stable ones are selected by evaluating their energy. The structure is then optimized by minimizing the internal energy, and the most energetically favorable structure is chosen [14].

The conserved fragments of the antibody are designed by homology modeling based on the known antibody tertiary structure using the programs such as Modeller, I-TASSER, and Rosetta [15-18]. Based on these approaches, several software packages have been developed, both commercial (products from Schröding-

er Inc., Chemical Computer Group, and Accelrys Inc.) and freely available (PIGS, WAM, SAbPred) [17, 19-21]. A combination of homology and *ab initio* methods forms the basis of the RosettaAntibody software package, which provides a high correspondence between the modeled and experimentally determined structures [22, 23].

The most significant result in the application of methods developed for modeling the antigen-antibody interactions has become the ability to alter the affinity of such interactions by introducing point mutations in the antibody sequence and rapid estimation of the binding affinity *in silico* [24, 25]. Kiyoshi et al. [25] modeled 1178 point mutations including sequential substitutions of each of 62 amino acid residues in the CDRs with other 19 residues in the structure of the 11K2 antibody directed against the chemokine MCP-1. *In silico* selection demonstrated that twelve of these substitutions resulted in increased affinity. Subsequent verification by the surface plasmon resonance method confirmed that five of them indeed had an increased affinity. For example, one of the mutants demonstrated a 4.7-fold higher affinity compared to the wild-type antibody. The authors noted that the presence of charged residues in the variable domain light chain fragment played the most significant role in enhancing the antibody affinity [25]. Therefore, the used methods have shown the ability to identify amino acid residues most important for the antibody-antigen binding.

**Modeling of antigen-antibody complexes.** In the aforementioned study [25], the authors improved the affinity of the antibody with already known structure of the antibody-antigen complex [25], while the tertiary structures of complexes of new antibodies, as well as most commercially available antibodies, are unknown [26]. Therefore, the ability to correctly identify the antibody-binding site on the target protein has become of a particular importance. Various experimental and computational approaches have been developed to solve this problem [27]. *In silico* methods use both the tertiary structures of target proteins and their amino acid sequences as input data. Although prediction of linear epitopes based on their amino acid sequence yields fine results, they represent only  $\sim 10\%$  of all epitopes [27, 28]. The attempts to predict conformational epitopes based on amino acid sequences have been mostly unsuccessful due to the difficulties in accounting for the spatial interactions between different fragments of protein sequences [29-31]. Therefore, the most common method is macromolecular (protein-protein) docking, in which two proteins are docked as rigid structures based on the assessment of their electrostatic and steric complementarity. Formation of protein-protein complexes is accompanied by mutually induced adjustments of proteins surfaces, which makes rigid protein-protein docking low-informative.

For this reason, approaches for modeling protein complexes with rapid mutual geometry optimization have been developed [32, 33]. Thus, RosettaFold performs rapid geometry adjustments by energy minimization using the Monte Carlo algorithm. The SnugDock algorithm developed based on the RosettaFold and RosettaAntibody software, was specifically designed for modeling the antibody-antigen complexes and is capable of performing rapid Monte Carlo optimization to obtain higher-quality models [34]. It was found that an antibody-antigen complex close to the native one is usually more energetically favorable, resulting in the global energy minimum for the modeled system, which makes selection more efficient [33, 35].

However, the use of this approach is limited by its low accuracy. The complex optimization step is coarse and the energy evaluation function is simplified, leading to poor correlation between the experimental data and predictions [36]. Therefore, the results of docking are often further evaluated using molecular dynamics simulations. For instance, in modeling the antibody complexes with beta-amyloid, after docking of crenezumab (a humanized antibody against beta-amyloid peptides 1-40 and 1-42 developed as a treatment for Alzheimer's disease) to beta-amyloid, about 200 complex models were generated and divided into eight clusters. Subsequent 200-ns molecular dynamics simulations were performed on the representatives of these clusters in order to optimize the complex structure, analyze the structure of the obtained complexes, and evaluate their interaction energy. Two of these complexes were stable, and they explained well the experimental data [35].

The Monte Carlo method is stochastic, and its predictive power depends on the number of iterations, which increases the number of possible configurations of the complexes and leads to a challenging problem of selecting the most probable structures. Special algorithms are employed to increase the conformational sampling in order to reduce random energy fluctuations, re-evaluate complex energies by increasing the importance of electrostatic and desolvation energy, and cluster the complexes. The programs implementing these algorithms are FiberDock, ZRANK, and pyDock [37-40].

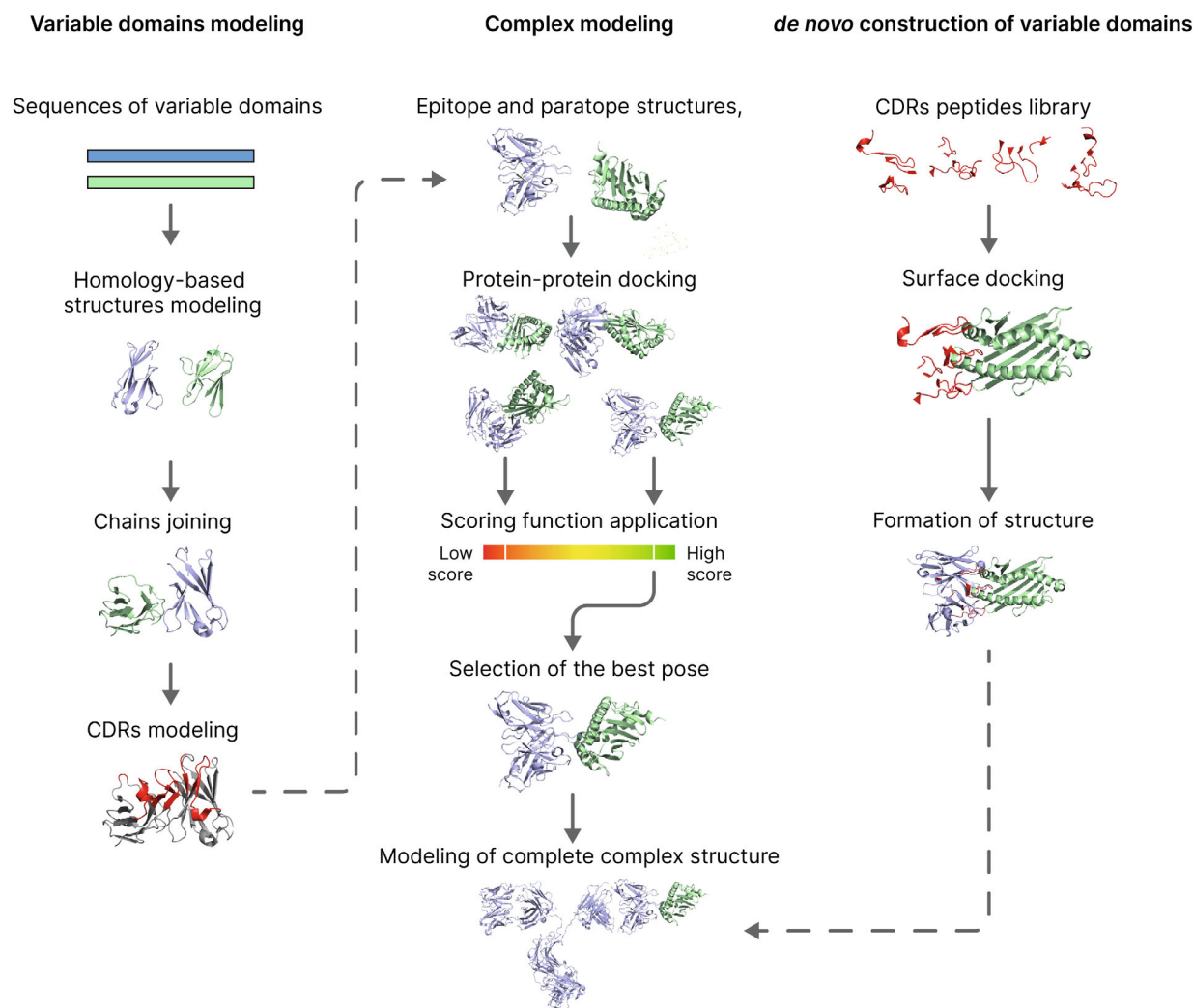
In addition to the Monte Carlo-based methods, direct approaches for searching the interacting surfaces have been developed using the fast Fourier transform (FFT). This algorithm can search for correspondences between the protein surfaces by transforming them into the frequency domain. Consequently, it identifies significantly fewer potential binding sites and conformations and in much lesser time, partly due to the ability to perform computations on graphics processors. However, the most significant factor in this algorithm is geometric matching (rather than

physicochemical compatibility), which substantially limits the applicability of the approach and requires additional modeling steps [41].

There are also methods based on predicting the interaction energy between two proteins, e.g., in antibody-antigen complexes, that are optimized to find correlations between the structure of interacting surfaces and experimentally determined affinity. Recently, the possibility of using the free energy perturbation method (FEP) to predict the influence of point mutations on the affinity and stability of antigen-antibody complexes has been demonstrated [42]. The FEP method is based on calculating the difference in changes in the free energy of a complex during conversion of a ligand (or amino acid residue) into another one via small changes in the ligand structure followed by optimization of each intermediate complex structure by molecular dynamics simulation. The DDMut-PPI program has been developed based on deep learning methods to predict with a higher efficiency the changes in the free binding energy for two proteins [43]. The increasing number of structures of protein complexes in the PDB database has made possible the development of frequency matrices. These matrices have shown the preference for polar and charged amino acid residues at the interaction interface, thus allowing rapid selection of potential interaction areas based on this criterion. This additional selection of complexes by their physicochemical properties significantly increases the probability of finding correct complex geometry [44, 45].

More universal combined approaches have also been developed. For instance, the modeling of antibody structure by homology, complex optimization, and application of sophisticated physicochemical evaluation functions have been used in the OptMAVEN-2.0 program for *de novo* modeling of variable antibody fragments for the interaction with specific epitopes on the protein antigen surface. This method implements protein-protein docking, pose optimization, side-chain positioning optimization, sequence selection, construction of statistical matrices, and more [46]. The method mimics the natural process of V(D)J gene recombination by designing antibodies from modular Ab parts (MAPs) with subsequent structure optimization to increase their affinity. When this approach was used for constructing antibodies against peptides, five designed antibodies that were tested experimentally demonstrated correct protein globule folding and stability in solution. Three of them exhibited a nanomolar-range affinity [47]. This approach was also used to design two antibodies against epitopes on the Zika virus envelope protein [46].

An intriguing prospect in the development of antibody modeling methods is the possibility of *de novo* design of antibodies against a specified epitope based



**Fig. 2.** Molecular modeling methods used for the development and optimization of monoclonal antibodies.

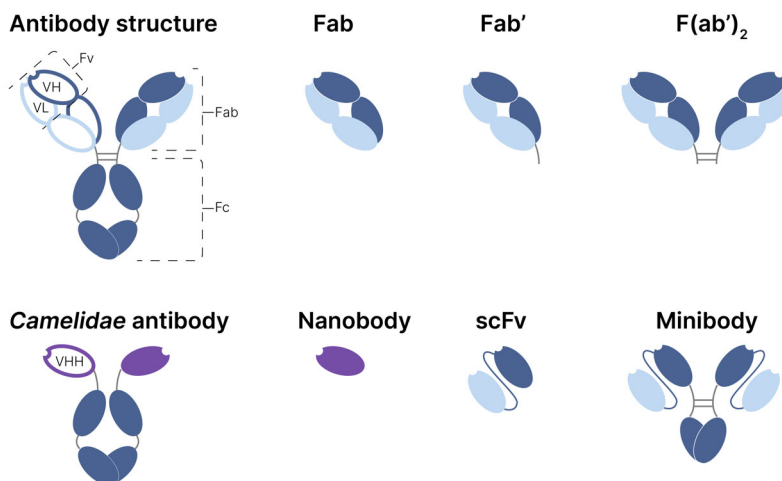
on computational models. A new approach has been developed that includes an assembly of a library of short seed sequences capable of binding to regions on the target protein surface. These sequences are docked onto the protein surface using the hot-spot method that involves the docking of a library of short fragments and selection of those showing the highest binding energy with the protein (the term “hot spots” often refers to amino acid residues that significantly contribute to the binding energy). Then, an antibody backbone is selected to link these short sequences. With a sufficiently diverse library, this approach could be used for the directed development of antibodies for specific epitopes; however, in the original work, its application was limited to the epitopes with the known antibody-epitope complex structures [48].

Combining statistical analysis and homology modeling is particularly applicable in the modeling methods based on deep learning. The AlphaFold2 neural network based on the transformer model and its

extension AlphaFold-Multimer have been utilized to model antibody-antigen complexes using primary sequences only. However, the modeling result still remained at ~30% success in reproducing the geometry of the experimentally obtained complex [49].

Various deep learning models based on graph neural networks, language models, etc., have been used for modeling antibodies for specific epitopes [50-52]. However, these methods are limited by the training datasets available from PDB, so that a very limited number of structures for both epitopes and antigen-binding domains have been used for training.

Other actively developed approaches are methods based on deep learning. They efficiently approximate hidden patterns and thus can provide a high accuracy in model construction, surpassing homology modeling methods, such as Modeller and I-TASSER [53]. However, high-quality training of such systems requires large training datasets, which are currently difficult to obtain. The small number of experimentally



**Fig. 3.** Comparison of antibody structures and examples of antibody fragments.

determined structures limits the efficiency of such approaches [49].

The methods used for the development and optimization of monoclonal antibodies are shown in Fig. 2.

In addition to the methods for designing the structure of antibodies and their complexes with antigens, various approaches have been developed to predict the antibody propensity for aggregation, immunogenicity, and pharmacokinetic clearance [54-56].

Despite a significant variety of methods for modeling interactions in the antibody-antigen complexes, most of them have a serious limitation, as they do not account for the role of solvent, which is crucial in the antibody-antigen interactions. The method most effectively addressing this issue is molecular dynamics simulations in explicit solvent that requires substantial computational resources. Potential solution to this problem might involve the use of accelerated dynamics methods, including those based on deep learning [57].

**Antibody fragments. Structure and methods of development.** As mentioned earlier, development and production of antibodies are costly processes that require continuous monitoring. Since only the antibody variable domain is involved in target binding, truncated antibody variants retaining the ability to bind their targets have been developed. Using genetic engineering methods, a variety of protein-binding antibody fragments have been created, such as single-chain variable fragments (scFVs), individual antigen-binding fragments (Fab's), minibodies, nanobodies, and others [58-60]. The comparison of antibody structures and their fragments is shown in Fig. 3.

The development of all above-mentioned monoclonal antibody derivatives has started with the application of hybridoma or phage display technologies. These technologies enable production of high-affinity and specific agents and allow further increase in their affinity during selection [61, 62]. Gene regions encod-

ing constant domains responsible for maintaining the antibody structure and binding to Fc receptors of the immune system are removed by genetic engineering methods, while the light and heavy chains of the immunoglobulin variable fragments are connected by the designed linker sequences.

Nanobodies are based on immunoglobulins with only heavy chain in the variable domain and, therefore, structurally simpler antigen-binding region. Such antibodies have been found in representatives of the Camelidae family. Reducing the size of a molecule decreases the cost and complexity of its production and increases the tissue permeability for this molecule, which is especially important in therapeutic applications. At the same time, this significantly reduces the stability of the antigen-binding region, thus worsening its binding parameters [58].

Methods for maturation on phage or yeast display are employed to solve the problem of reduced affinity and specificity of antibody fragments. In these methods, single substitutions are introduced into the sequences encoding antigen-binding regions, resulting in the creation of a diverse library from which the most affine and specific structures are then selected. Due to lower costs and effort associated with working with antibody fragments, this approach proves to be quite efficient [63].

**Molecular modeling methods in the development and research of antibody fragments.** Due a smaller size of antibody fragments, the use of molecular modeling methods in the development and optimization of antibodies has proven to be more efficient [64, 65].

Utilization of small antibody fragments makes it possible to design bispecific binding systems composed of fragments from two different antibodies connected by a linker. Molecular modeling methods have been successfully used to create bispecific scFVs that inhibit formation of the T-cell receptor (TCR) complex

with the major histocompatibility complex (MHC-II) in the presence of staphylococcal enterotoxin B (SEB) [66]. The most important modeling tool in this study was the microsecond molecular dynamics simulation which helped to decipher the mechanism of allosteric inhibition of the SEB-TCR complex formation.

In [66], a bispecific system was created from the variable domains of known antibodies. However, predicting the structure of new protein-binding antibody fragments requires specialized methods. Tools used for modeling complete antibodies are ineffective for this task, as the absence of stable domains increases the antibody fragment mobility [67]. For these tasks, special programs, like NanoNet, have been developed. This deep learning model can accurately and efficiently (about 1 million structures in four hours on a standard CPU) predict the structures of nanobodies using only their amino acid sequences as an input. These models, combined with flexible protein-protein docking, can serve as virtual screening tools to find new high-affinity and specific protein-binding agents [68].

Thus, despite a smaller size of antibody fragments compared to full monoclonal antibodies, their modeling poses similar challenges and requires significant computational resources.

**Antibody mimetics. Structure and methods of development.** Antibodies and their fragments are the most commonly used protein-based high-affinity and specific binding agents. However, they have several disadvantages that limit their application. High costs of production, low stability (which leads to problems with aggregation), immunogenicity, and high costs of storage and logistics are the main limiting factors that hinder wider use of antibodies. Most of these problems are related to the native structure of antibodies, which are not meant to exist outside their native environment. Despite the small size of antibody fragments, their structure still remains suboptimal and includes relatively large regions whose function is limited to maintaining the structure of the binding site [69].

Therefore, it was proposed to develop protein-binding agents non-homologous to antibodies; they are known as antibody mimetics (or non-immunoglobulin epitope binders). Designing simpler protein structures with an enhanced affinity and specificity makes it possible to overcome the limitations of antibodies while retaining their useful features.

Protein antibody mimetics have a simple rigid backbone composed of alpha-helices or beta-sheets, while their protein-binding sites are usually free loops [70]. Examples of protein scaffolds for antibody mimetics are shown in Fig. 4. Their development typically employs site-directed or random mutagenesis. Protein antibody mimetics are produced using simple bacterial reactors that provide a higher yield of these molecules compared to antibodies [63].

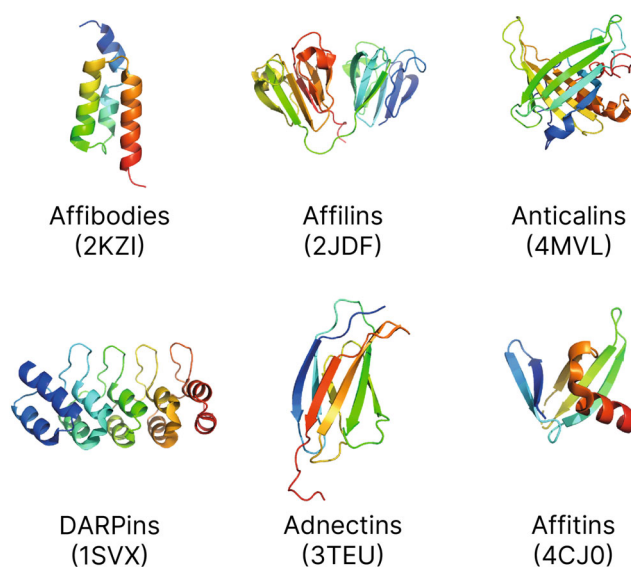


Fig. 4. Examples of backbones for proteinaceous antibody mimetics (PDB codes are given in parenthesis).

The most common method for developing antibody mimetics involves a cycle of directed evolution followed by candidate selection. Initially, a DNA library encoding homologous protein structures is amplified by error-prone PCR (achieved by adding manganese ions to the reaction mixture) to significantly increase the diversity of amino acid sequences relative to the original library. The most promising sequences with an increased affinity and specificity are selected using the phage display method, and the procedure is repeated [71]. After a limited number of cycles, a highly affine and specific protein-binding agent with high stability can be obtained.

However, this method has limitations that hinder a large-scale entry of antibody mimetics to the market. One of the main challenges in the design of protein-based antibody mimetics is finding an appropriate protein scaffold. Despite a large variety of known scaffolds, such as affibodies (B-domain of staphylococcal protein A), adnectins (extracellular domain of human fibronectin III), and affitins (variants of DNA-binding protein Sac7d), each protein backbone has its drawbacks, and discovering new ones is constrained by the ability to solve protein tertiary structures. Since most of these proteins have structures foreign for an organism, they tend to be immunogenic, which limits their application *in vivo*. Also, due to the absence of Fc fragments, they lack effector functions [63]. Consequently, although protein-based antibody mimetics are free from many drawbacks typical for antibodies, their quantity and areas of application are currently limited.

**Molecular modeling methods in research and development of protein antibody mimetics.** Molecular modeling methods address several serious challenges

in the development of protein antibody mimetics. The antibody mimetic structure typically includes stable, water-soluble fragments of known proteins that can adopt a correct conformation in bacterial cells. This simplifies their identification, since the amino acid sequences of these fragments predominantly consist of hydrophilic residues, have a low cysteine content, etc. [72]. Analysis of statistic distribution of amino acid residues allows to select a limited set of residues that can be used to form structures with the desired properties, such as solubility and pH sensitivity, without compromising the binding parameters, resulting in the production of proteins retaining stable structures over a wide range of temperatures and pH values.

The tertiary structure of such proteins can be predicted with programs like AlphaFold2 and RosettaFold. However, the accuracy of these predictions remains uncertain as these programs are trained mostly on native structures [73]. A potential solution to this issue may be the application of diffusion-based deep learning models, such as AlphaFold3 and Chroma, which are capable of predicting protein structures based on the desired parameters [74, 75]. Due to the ability of diffusion models to make small iterative steps during generation, the imposed constraints can cumulatively direct the modeling towards the desired properties. Thus, the structures predicted by Chroma showed minimal differences with the experimentally determined structures (RMSD,  $\sim 1$  Å). Since Chroma has a subquadratic computational complexity from  $O(N)$  to  $O(N \log N)$ , this tool could be very useful for developing the backbones for protein-based antibody mimetics [75]. A significant limitation of diffusion models is generation of the so-called “hallucinations,” i.e., unrealistic structures and connections that contradict the physical properties [76]. Therefore, expert evaluation of modeling results is crucial.

However, the methods used for modeling protein structures do not address the issue of potential immunogenicity of designed proteins. Based on proteomics data, databases such as SEDB, Epitome, IEDB, AntiJen, and Bcipep have been created, that store information about epitopes recognized by T-cell and B-cell receptors and antibodies, as well as those incorporated into MHC molecules [77-81]. This information can help to identify potential epitopes and replace them when designing protein scaffolds, thereby reducing their potential immunogenicity.

Despite a smaller size and lower structural complexity of antibody mimetics compared to antibodies, there are currently few examples of application of molecular modeling methods for their development. The primary approach in the *de novo* design of protein antibody mimetics is the hot spot method, which combines the docking of potentially binding peptides and subsequent selection of the scaffolds [82]. Using this

method, several mimetics have been developed that bind with a high affinity to the conserved regions of the influenza virus hemagglutinin (HA). As a result of modeling, two high-affinity proteins were designed, and the structure of the HA complex with the antibody mimetic showing the highest affinity was analyzed by X-ray crystallography. The structure of the predicted complex closely matched the experimentally determined one [83].

A combination of several approaches, including high-molecular flexible docking, molecular dynamics simulation, deep learning methods for the tertiary structure prediction, and immunogenicity prediction algorithms, can significantly improve existing approaches and lead to the development of new protein antibody mimetics. These mimetics offer a number of advantages and lack the limitation of monoclonal antibodies and their derivatives.

## NON-PROTEINACEOUS AFFINE PROTEIN-BINDING AGENTS

**Aptamers.** *Structure and application.* The disadvantages of protein affinity agents are related to the cost of production, problems with storage and logistics, immunogenicity, etc., have stimulated the search for non-proteinaceous protein-binding agents. The most promising of them are aptamers.

Aptamers are short single-stranded molecules of DNA, RNA, or synthetic nucleic acids, usually 20 to 60 nucleotides long, with a high affinity and specificity for various high- and low-molecular-weight targets [84]. Aptamers were independently discovered by two groups in 1990 as a result of the SELEX (systematic evolution of ligand by exponential enrichment) process [84, 85]. Later, aptamers have been found in nature as components of bacterial riboswitches [86].

The idea of using nucleic acids for the recognition of protein targets has emerged in the studies of human immunodeficiency virus (HIV). It was shown that the transactivation regulatory element (TAR) containing RNA sequences can inhibit HIV replication by binding to the viral protein Tat with a high affinity and specificity [87]. Despite that aptamers are completely different from antibodies; it is believed that they can compete with antibodies in both diagnostic and therapeutic applications [88]. Although aptamers recognize and bind targets in the same way as antibodies, they have a number of advantages, such as shorter production time and lower cost, easier modifications, better thermal and chemical stability, smaller size, and low immunogenicity [89].

Due to their advantages over antibodies, aptamers have found applications in various fields of molecular biology, biotechnology, diagnostics, and therapy.



One of the promising areas of research using aptamers is development of optical sensors. Numerous studies have shown that fluorescent dyes attached to the conformationally flexible regions of the aptamers can produce an optical signal in response to ligand binding [90]. However, the number of aptamers on the market is limited compared to antibodies, primarily due to the lack of an urgent need for infrastructure restructuring, which is economically costly and currently impractical for most applications [91].

Many aptamers designed for therapeutic purposes are currently tested in clinical trials or have already been released to the market, including drugs for the treatment of age-related macular degeneration, blood clotting disorders, cancer, inflammatory processes, etc. [89, 92].

Besides being used as therapeutic agents, aptamers can be employed for targeted drug delivery into human cells to enhance drug efficacy and reduce side effects [93]. Human prostate-specific membrane antigen (PSMA, a transmembrane protein associated with prostate cancer) that is overexpressed on the surface of tumor cells, has become the first model system for the aptamer-based drug delivery [94]. Bispecific aptamers against gp120 and CD4 receptor deliver a small interfering RNA-based active agent that inhibits HIV activity *in vitro* [95].

*Methods for designing and producing aptamers.* A standard method of aptamer selection is SELEX, which is an analogue of directed evolution followed by screening of protein-binding candidates, but instead of phage display and expression in bacterial cells, it involves *in vitro* incubation and PCR. During SELEX, the target is first incubated with a set ( $\sim 10^9$ - $10^{11}$ ) of single-stranded random oligonucleotides known as the primary library. The oligonucleotides in the SELEX library typically consist of 40-100 nucleotides due to limitations of chemical synthesis, with a random region in the middle and common sequences at both ends for primer annealing. After incubation of the primary library with the target, unbound oligonucleotides are removed, and the aptamer-protein complexes are separated. The released DNA sequences are amplified by the error-prone PCR. Standard SELEX usually includes several rounds of this procedure [85]; the selected aptamers are sequenced and parameters of their binding to the target are assessed by various methods, such as surface plasmon resonance or isothermal calorimetry. Up to twenty selection rounds are typically carried out to enrich aptamers with a high affinity to the target [96].

Although traditional SELEX is the main approach for generating aptamers, it is not exempt from limitations. The primary library should include as many diverse structures as possible, while at the same time, it should not contain double-stranded nucleic acids, as

well as single-stranded linear structures incapable of providing reliable folding. The problem of finding the optimal primary library remains unsolved [97]. Typically, to increase the library diversity, a large number of structures with a high proportion of GC pairs are used. It has been shown that increasing the GC content of the primary library leads to a greater complexity of tertiary structures, resulting in a greater diversity of potential aptamers and higher average affinity [98].

Another approach to increasing the diversity of the primary library is an approach developed by SomaLogic (a subsidiary of Standard BioTools). The company's primary library for the production of special modified aptamers, the so-called SOMAmers, contains a large number of structures in which thymidines are replaced by C5'-ethynyl-2'-deoxyuridines. During SELEX, these residues are modified by side chains of amino acids or other groups to expand the properties of the obtained aptamers [99]. However, the evident limitations of this method are its high cost and possibility to insert only a single chemical modification into a SOMAmer structure.

The PCR process used for enrichment in SELEX is not ideal. Due to the presence of a large variety of secondary structures in the primary library and a higher rate of synthesis of short sequences, there is a bias that can lead to excessive enrichment with the structures possessing the structure most advantageous for polymerization reaction, rather than with structures with the optimal required parameters. To solve this problem, emulsion PCR is used, in which each individual structure is amplified in a microsome, thus not competing for the polymerase active site [100].

Despite numerous approaches to the SELEX optimization, its efficiency currently remains at a 30% probability of detecting the desired aptamer [101].

One of the serious limitations of aptamers compared to antibodies is the related to the absence of hydrophobic properties in natural nucleic acids. Among the solutions are the above-mentioned SOMAmers that can be modified, in particular, by side radicals of hydrophobic amino acids (e.g., benzyl radical) (Fig. 5a). However, this modification is labor-intensive and not universal, since the obtained aptamers cannot be not amplified directly by PCR, and additional steps are required. More promising is addition of a synthetic base, such as dNaM-d5SICS (Fig. 5a). Due to the presence of polyaromatic hydrocarbon group, this base is hydrophobic, but still maintains its ability for complementary replication in PCR using modified DNA polymerases [102]. However, adding a new base pair to the primary library for SELEX requires an increased diversity and significantly raises the costs of production.

Another limitation of aptamers as therapeutic agents is the lack of effector functions. Sequentially cross-linking of free ends of two aptamers, one of

which bound to the target protein and the second interacted to FcRyIII, resulted in the effector aptamer capable of inducing immune response via binding to the target protein [103].

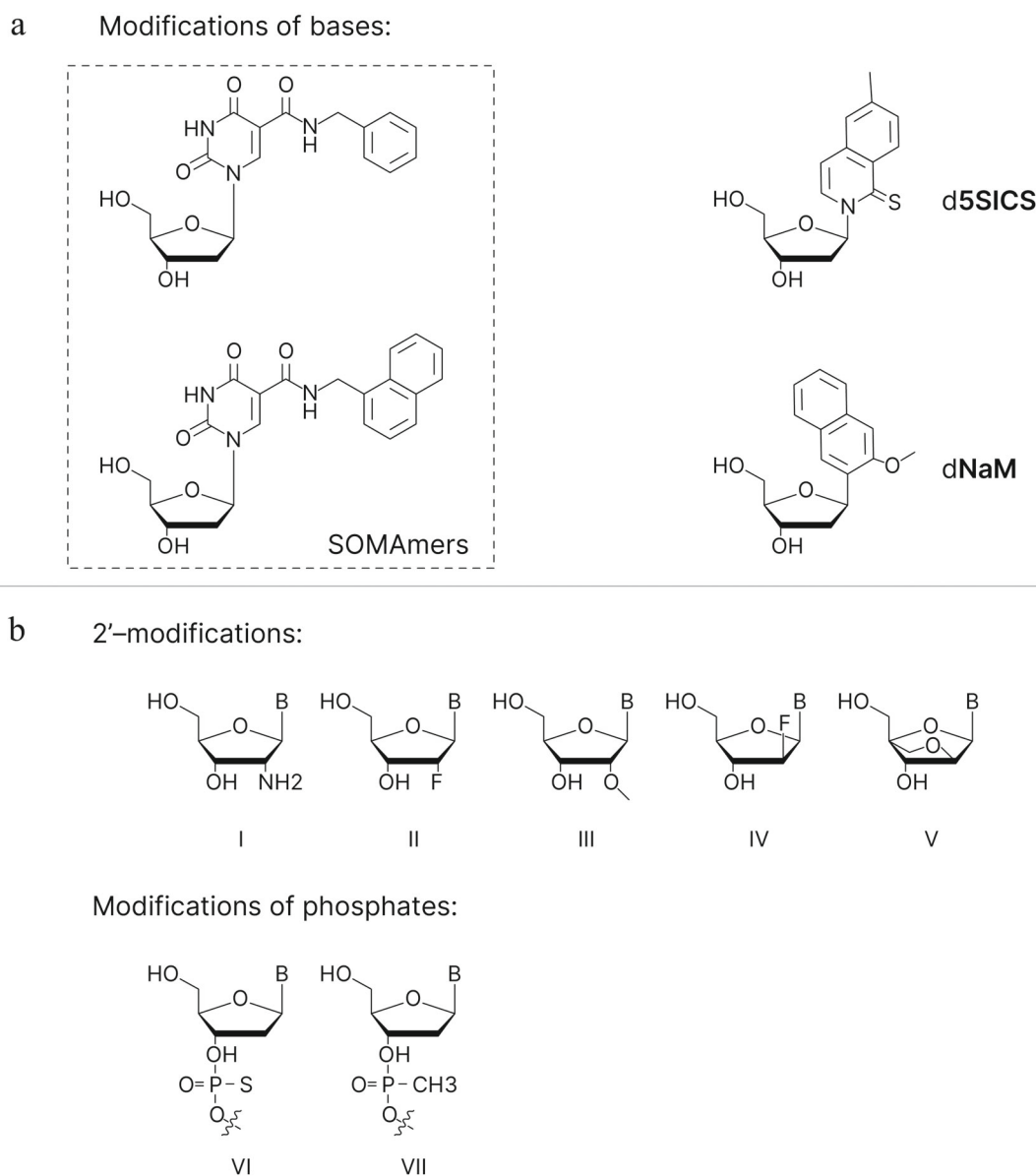
Since aptamers are short oligonucleotides, they are susceptible to degradation. RNA-based aptamers are highly unstable due to self-degradation caused by formation of transient bonds between the 2'-hydroxyl group and phosphate, as well as due to the high level of environment contamination with ribonucleases that catalyze RNA cleavage [104, 105]. DNA aptamers prove to be more practical in terms of stability; however, they can also be degraded by deoxyribonucleases *in vivo*. Numerous modifications of aptamers have been developed to prevent both autodegradation and the action of various nucleases (Fig. 5b). Substitution of the 2'-hydroxyl group with amino group (I), fluorine (II), or O-methyl group (III) prevented both autodegradation and cleavage by nucleases. However, this modification significantly enhanced the flexibility of nucleic acid, making it similar to the flexibility of RNA, which may affect the structure of the resulting molecules. 2'-Deoxy-2'-fluoro-D-arabinonucleic acid (IV) and locked nucleic acids (LNAs) (V) are used to significantly increase the resistance of aptamers to nucleases, while maintaining the flexibility of DNA molecule [106]. More effective in this case is replacement of the phosphate group oxygen with sulfur or methyl group, which does not affect the structure of deoxyribose. In this case, the aptamer acquires resistance to nucleases, but alteration of the molecule charge increases its ability for non-specific binding to the target [106, 107]. Also, modifications increasing the resistance to nucleases or increasing their affinity can prevent polymerization in PCR, which complicates the SELEX process [108].

**Molecular modeling methods in design and studies of aptamers.** *Method for aptamer selection from a primary sequence library.* To optimize the development of aptamers, numerous methods based on molecular modeling have been proposed. Commonly used computational approaches include docking and molecular dynamics, which allow to simulate interactions between the aptamer and its target. These methods are sometimes supplemented by quantum mechanics or hybrid energy calculation methods to provide a more precise estimation of the binding energy [109]. Methods for predicting the secondary and tertiary structure of RNA and DNA play an important role in the design of aptamers, since the tertiary structure is required to model the protein-aptamer complexes [110]. Modern *in silico* tools make it possible to model aptamers for both small molecules and complex biopolymers. The only limitation of these methods is that they cannot use cells as targets (unlike SELEX). A typical cycle of aptamer modeling *in silico* starts with the

secondary structure prediction, followed by the tertiary structure prediction and its optimization. Next, rigid or flexible docking of the target and aptamer is performed, leading to the selection of complexes with the highest binding scores. The next important, but not mandatory step is molecular dynamics simulation to assess the complex stability and to determine the binding energy with a higher accuracy [111]. Analysis of the aptamer interaction with the target allows to perform nucleotide substitutions or chemical modifications, after which the previous steps can be repeated with new candidates [112]. In addition to the structure and interaction modeling methods, quantitative structure-activity relationship (QSAR) models have been utilized in aptamer design for rapid prediction of the desired feature and scanning of the sequence space [113-115].

All the above computational methods can be used in combination with the experimental ones, complementing each other and increasing the efficiency of aptamer design. For example, the diversity of the primary SELEX library can be increased by using the methods for prediction of secondary structures that allow estimating the structure complexity and selecting sequences with the greatest potential. Such methods can also be applied to select a correct ratio of structures in the primary library to prevent an uneven PCR amplification [116]. However, the methods for the secondary structure prediction are generally low-performance and unable to quickly predict a large number of different structures. Also, they are mostly applicable to RNA, while aptamers are predominantly based on DNA due to its greater structural stability and cost-effectiveness. Furthermore, most secondary structure prediction methods ignore the common pseudoknot structure due to the high computational cost of its prediction [116, 117]. Therefore, deep learning methods have been proposed to develop models capable of predicting nucleic acid structures with a high speed and accuracy [118]. However, currently, experimental data on single-stranded DNA structures are still insufficient for a high-quality model training.

A primary sequence library in SELEX contains a large number of diverse sequences; however, an aptamer selected for a specific target may still be suboptimal, i.e., have low tertiary structure stability or affinity. A multistep approach involving introduction of point substitutions into the aptamer followed by the molecular dynamics simulation was applied in the development of an aptamer for binding with the antibiotic sulfadimethoxine [119]. The authors used an aptamer generated through SELEX as a starting component. Its interaction with the target was modeled using molecular dynamics, and the affinity of interaction was estimated as a change in the Gibbs free energy during interaction between the aptamer and the target molecule.



**Fig. 5.** Modifications of nucleic acids to increase the affinity of aptamers and their resistance to nucleases: a) modification adding molecule hydrophobicity; b) modifications increasing resistance to autodegradation and hydrolysis by nucleases: 2'-amino- (I), 2'-fluoro- (II), 2'-O-methyl-deoxyribose (III), 2'-deoxy-2'-fluoro-D-arabinonucleic acid (IV), LNA (V), and phosphate modifications (VI, VII). B, any base.

Point substitutions were introduced into the aptamer, after which molecular dynamics simulation and assessment of binding energy were performed again. The best sequences were selected at each step, which allowed to increase the aptamer affinity without resorting to additional SELEX experiments.

Another example of this approach is development of an aptamer against prostate-specific antigen (PSA) [120]. The researchers utilized five sequences obtained by SELEX as a basis for modeling and adopted the functions of reproduction and crossover in the genetic algorithm to produce next-generation sequences, whose interaction with the target was assessed using

protein-protein docking. The best candidates were synthesized and their interaction with PSA was evaluated using a quartz crystal microbalance to select the most promising candidates. Hence, in this study, the sequences were iteratively optimized for the target by utilizing both real experiments, as well as simpler and more high-performance computer models. The resulting aptamer showed a three-fold higher affinity for the target than the original aptamers obtained by SELEX.

*Method for aptamer modeling based on the target structure.* In the studies described above, computer modeling has been used mainly to improve already existing experimentally obtained aptamers. However,

approaches to *de novo* modeling of aptamers based on the information about the target structure have been proposed. To design an aptamer for cytochrome P450, Shcherbinin et al. [121] used the hot spot method and defined two distinct regions in the nucleotide chain. The first one was involved in the formation of bonds with the protein through nucleotides, making the principal contribution to the molecule affinity and specificity. The second region was responsible for preserving the aptamer's tertiary structure. After determining the protein surface regions with positively charged amino acids, the structures of trinucleotides binding to the protein were obtained using small molecule docking and molecular dynamics simulation. The fragments exhibiting the highest affinity were fused with helical complementary regions to form a loop structure. The synthesized aptamers showed a high specificity by exclusively binding to proteins within the same family; the experimentally determined binding energies showed a strong correlation with the calculated ones [121].

The modeling of molecules usually starts with a creation of a set of candidate molecules and assessment of their efficiency for further selection. The most accurate, but at the same time, the most expensive approach for such assessment is experimental estimation of the binding energy for the candidate interaction with the target. Hence, molecular docking and dynamics methods are used instead, although, simulations for a large number of candidates can be time-consuming. An alternative approach involves utilizing machine learning methods to predict a required parameter. This approach was used to develop an aptamer against aminopeptidase CD13. The aptamer sequences were generated iteratively using a genetic algorithm. Next-generation sequences were selected by employing an evaluation function based on the machine learning model pretrained on the characteristics of the primary and secondary structures [114]. In another study, a neural network model trained on experimental data was used to explore a set of sequences to develop aptamers against neutrophil gelatinase-associated lipocalin. The authors developed and synthesized aptamers with a high affinity, but also found shorter nucleic acid sequences with a comparable affinity and greater stability [115]. However, such neural networks can identify dependencies and improve already known structures, but when searching for new compounds, they can exclude promising structures from consideration. For this reason, the results of these studies require thorough validation and good interpretation.

Molecular modeling is most promising in the development of modified aptamers more suitable for therapeutic applications. Thus, using *in silico* approaches to the aptamer development allows to select an optimal combination of modifications that can enhance the

binding affinity, resistance to nuclease degradation, while maintaining the binding specificity. In [122], sequential application of different computer programs allowed to model modified aptamers and their interactions with target proteins. The authors used the classical prediction of the aptamer tertiary structure, followed by the nucleotide modification, high-molecular docking, and analysis of interacting bases. In particular, Mfold and 3dRNA were used to predict the tertiary structures of nucleic acids. However, Mfold cannot predict pseudoknots and also demonstrates low performance [123]; 3dRNA also has low performance and shows a significant discrepancy between the experimental and predicted structures (RMSD,  $\sim 4$  Å) [124]. Thus, due to the low prediction accuracy and low performance, the suggested approach is limited to modeling new aptamers for a given target.

Another obvious limitation of this method is the absence of molecular dynamics stage for optimization of the complex structure. Even a single modification can profoundly affect the tertiary structure of nucleic acid [125]. For this reason, after modifications, the structure should be investigated by molecular dynamics methods to identify conformational changes. However, conventional force fields for modeling nucleic acids cannot be used to simulate modified nucleic acids. Galindo-Murillo et al. [126] conducted parameterization of modified nucleotides, enabling molecular dynamics simulation of atypical bases, which yielded a good similarity between the modeling and experimental results. However, the range of parameterized modifications was limited. Existing universal force fields are capable of simulating molecules with many different modifications, but for biopolymers, these force fields are in poor agreement with experiments [127].

Therefore, there are two promising directions in the development of methods for molecular modeling of aptamers based on the target structure. One group of methods imitates the SELEX selection process using the tertiary structure prediction, docking, and molecular dynamics, while the other group of methods provide a targeted development of the nucleic acid structure capable of binding to the epitope on the protein surface (Fig. 6). Both approaches require high-quality prediction of the secondary and tertiary structure, as well as the use of molecular dynamics simulation. The development of these methods is crucial for the successful prediction of new aptamers.

**Molecularly imprinted polymers. Structure and methods of production.** Molecularly imprinted polymers (MIPs) were proposed in 1973 to facilitate production of antibody analogs [128]. MIPs are porous materials capable of selective recognition of a template (molecules intended to be recognized by MIP). They are obtained by self-assembly of functional monomers

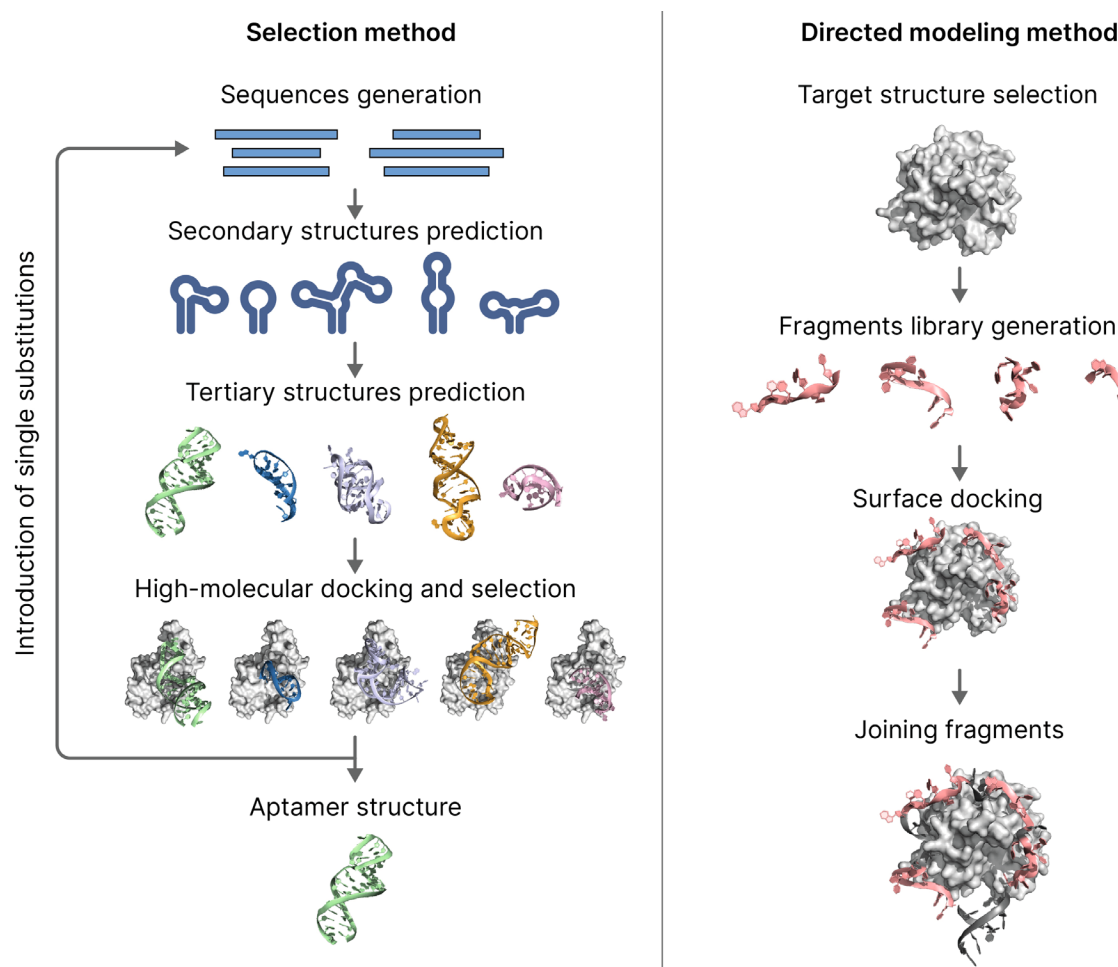


Fig. 6. Application of molecular modeling methods in the development and optimization of aptamers.

around the template molecule in a porogen (solvent for pore formation), after which polymerization is initiated in the presence of a cross-linking reagent [129]. MIPs have found many applications in diagnostics (e.g., in analogs of immunoassay), affinity separation, drug delivery, etc. [130-135]. MIPs were originally designed for low-molecular-weight compounds, but nowadays they start to find their application for interaction with protein molecules.

The basic approach to obtaining MIPs is a directed creation of a surface with the geometric and physico-chemical correspondence to a template. The first step is addition of functional monomers, which interact with the template and form an unstable pre-polymerization complex. At the second stage, polymerization of the resulting complex is carried out in the presence of an excess of cross-linking monomers and a porogen. At the last stage, the template is washed out from MIP, thus creating a cavity for specific binding of the target protein [136, 137].

There are two main approaches to obtaining MIPs. The first one includes the use of covalent interactions between the monomers and the template; after poly-

merization, the covalent bonds are cleaved. Due to the complexity of this procedure and the need for additional cleavage steps, this approach is rarely used for proteins.

The second approach is based on the formation of relatively weak noncovalent interactions between the template and functional monomers, such as hydrogen bonds, ionic interactions, van der Waals forces, dipole-dipole and hydrophobic interactions. Due to its simplicity, it is the most commonly used method. However, noncovalent binding requires a more careful selection of the composition and number of functional monomers [138].

Depending on the template structure and selected polymer, the synthesis of MIPs always needs with a careful selection of the reaction mixture composition and reaction conditions, which is one of the main disadvantages of this method. In addition to the unguided selection of possible components, functional and cross-linking monomers, porogens, and reaction conditions, this process is also poorly reproducible. Also, each new batch of MIPs requires the use of a template, which hinders the scalability of production [139].

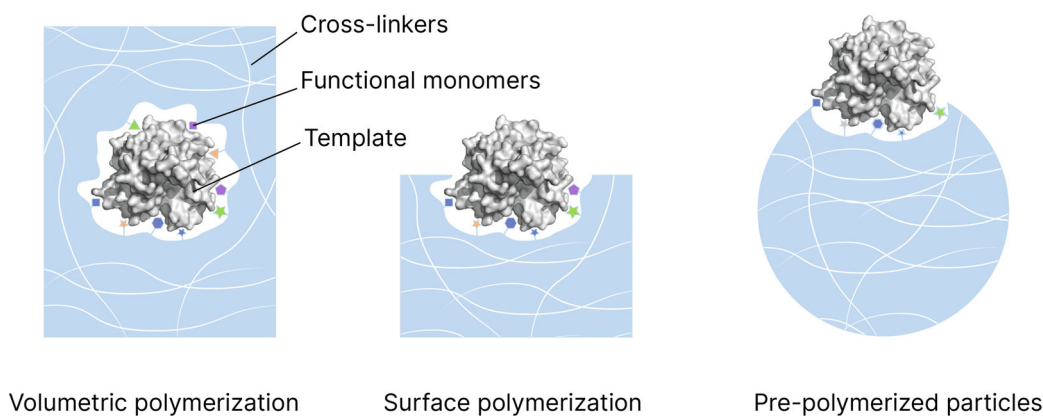


Fig. 7. Protein-binding MIPs.

A combinatorial chemistry method has been proposed to solve the problem of component selection, which involved creating a large variety of possible compositions and running experiments on production of MIPs in parallel [140, 141].

Other factors complicating MIP synthesis are polymerization conditions, including elevated temperatures, pH values unnatural for proteins, and organic solvents, which can lead to changes in the template structure [132]. The method of epitope imprinting was proposed as an alternative to using the whole protein. Most proteins have surface regions with a large number of polar and charged amino acids that are most commonly involved in binding. Instead of attempting polymerization around the whole protein, which requires selection of conditions ensuring retention of the protein structure, only fragments that include a given epitope are used. Due to a smaller size and often more stable structure of such epitopes, this approach increases the range of acceptable conditions for the polymerization reaction. However, in order to achieve a full correspondence between the given epitope and the native structure, it is necessary to carry out rather complicated design of smaller proteins that preserve the structure of the epitope [142].

Although it is possible to test a large number of components to obtain a required MIP structure, the strength of interactions between the functional monomers and the template is usually low, so the monomers are added in excess. This often results in a wide range of equilibrium states of the monomer–template complex, leading to the variety in the affinity and specificity of the resulting MIPs [143].

Polymerization around the template must allow subsequent unimpeded removal of the template protein from the binding site, leaving access for a new protein. Therefore, tightly polymerizing cross-linking monomers used to develop MIPs for small molecules are not suitable for proteins. For this reason, methods allowing to change reaction conditions in order

to regulate the pore size and, therefore, to control the binding and release of proteins have been proposed. In particular, it was suggested to carry out polymerization with the addition of pH-sensitive cross-linking elements of peptide nature, so that the increase in pH would lead to the increase in the pore size and binding site increase and provide the protein release [144].

A more popular approach is the surface imprint method, when the template is coated with a thin layer of polymer that does not completely cover the protein globule and does not interfere with the protein molecule release and binding. Further development of this method is nanoparticle pre-polymerization, when functional monomers are bound to the pre-prepared nanoparticles composed of a cross-linking polymer, and full polymerization occurs after interaction with the target protein. However, even with this modification, the surface imprint method significantly reduces the binding specificity compared to bulk polymerization [145]. The types of protein-binding MIPs are shown in Fig. 7.

*Molecular modeling in the development and research of MIPs.* MIP development is a long and costly process. It requires selection of optimal monomers, the ratio of components in the reaction mixture, and optimal reaction conditions. This complicates and lengthens process of develop MIP development, which limits their application. Computer modeling can significantly optimize this process.

To date, most studies on molecular modeling of MIPs have been performed for low-molecular-weight compounds [146-149]. The main efforts have been focused on the pre-polymerization stage. Both quantum mechanics and molecular dynamics methods have been used for this purpose in order to find the most suitable functional monomers, evaluate the binding process, and select the optimal ratio of the reaction mixture components [150-157].

Several computational approaches have been proposed to design MIPs that bind to proteins [158-160].

In this case, the most commonly used methods are molecular docking and molecular dynamics simulation.

One of the difficulties in creating MIPs is their possible influence on the structure of the template protein. Thus, molecular dynamics modeling of complex formation between ribonuclease A and functional monomers based on styrene and cross-linking polymer polyethylene glycol dimethylacrylate 400, showed possible conformational changes in the protein that led to its inactivation [161]. At the same time, no possible influence of functional monomers on the PSA structure has been revealed by molecular dynamics simulations in [162].

Kryscio et al. [158] studied the mechanisms underlying the effect of MIP synthesis conditions on the conformation of albumin by docking various functional and cross-linking monomers onto the protein surface. It was found that the monomers preferentially bound to a particular position at the protein surface, where they significantly affected the protein tertiary structure. Moreover, different functional monomers competed for the same amino acid residues.

Molecular modeling techniques are also used to select monomers and to find the optimal reagent ratios for the design of protein-specific MIPs [158, 163, 164]. Rajpal et al. [165] used docking to analyze the composition of a mixture of functional monomers in order to select the optimal reaction mixture for the synthesis of MIPs for several peptides prior to laboratory experiments. Docking followed by molecular dynamics simulation has been used to select the concentrations of functional and cross-linking monomers for the optimal binding to myoglobin [159].

Another important task is prediction of the MIP affinity to its target [166]. Lowdon et al. [166] evaluated several deep learning-based methods for the ability to estimate the affinity of MIPs to small molecules. Thus, the use of different deep learning models allowed to predict well the binding parameters of molecular imprints to 2-methoxyphenidine. A similar approach can be applied to the analysis of MIP binding to proteins, which can greatly simplify selection of components and polymerization conditions in the preparation of molecular imprints.

## CONCLUSION

The development of high-affinity and specific agents clearly evolves towards simpler, cheaper, and more technologically advanced solutions. Initially, these agents were monoclonal antibodies and their derivatives (chimeric, humanized, and human monoclonal antibodies). The need to simplify and cheapen the production and further development of antibody technologies have led to the use of antibody fragments,

such as scFvs, single antigen binding fragment (Fab's), minibodies, nanobodies, and others. Antibody mimetics became the next step in the development of proteinaceous high-affinity agents. Because of difficulties with the storage and application of protein systems, the next stage was the development of non-proteinaceous systems, the main ones being aptamers and MIPs. All developed high-affinity agents have their advantages and disadvantages. It can be expected that in the future, these technologies will be developed in parallel, providing researchers with a wide range of solutions to choose in each particular case depending on the tasks at hand.

In recent years, molecular modeling methods have been actively used in the development of high-affinity agents. Their application has allowed to abandon the traditional trial-and-error approach. Modeling the interaction of a high-affinity agent with its target provides a more detailed understanding of the molecular mechanisms underlying the binding and facilitates selection of the optimal structure of the system and conditions for its preparation. Molecular modeling has become an integral part of such developments.

The main methods are:

1. Homology modeling to construct the structures of antibodies, their fragments, and known antibody mimetics, as well as the structure of the target;
2. *Ab initio* methods for predicting variable binding regions of the antibodies, their fragments, and antibody mimetics;
3. Macromolecular docking for the pose selection for interacting bases in the design of protein-binding agents and aptamers (e.g., hot spot method) and in selection of components for MIP design;
4. Macromolecular docking for predicting the structures of protein–protein and protein–nucleic acid complexes;
5. Molecular mechanics, quantum chemistry, and statistical methods to estimate the interaction energies as a basis for selecting the binding agents with a higher affinity;
6. Molecular dynamics to simulate the *in vivo* behavior of molecular systems and to assess the stability of simulated structures and their complexes.

The classical modeling methods continue to improve; however, in recent years, there has been a dramatic increase in the interest in prediction systems based on deep learning. The ability of neural networks to recognize hidden patterns allows to create the programs with a high predictive power. Systems based on deep learning are used in modeling and *de novo* design of proteins, protein–protein complexes (including antibodies and antigen–antibody complexes), and nucleic acids, as they allow approximation of physicochemical characteristics, accelerate computationally expensive simulation processes, and more.

However, a serious limitation of deep learning-based methods is insufficient training datasets, often leading to a decreased accuracy or “hallucinations”, which prevents these methods from fully supplanting the classical approaches.

Modern software packages, including freeware and web-based ones, can perform the whole cycle of computations in the development of all types of high-affinity agents. Nevertheless, more data, algorithm improvement, and comparative studies of the applicability of different methods are required to further increase their accuracy, performance, and efficiency.

**Contributions.** Sh.Sh.N., A.R.M., I.V.M., and V.A.D. developed the study concept, wrote the text of the article, and prepared the figures; I.I.K., I.V.L., G.A.P., and A.L.M. wrote and edited the article; A.V.V. edited the article and supervised the study.

**Funding.** The work was performed within the framework of the Program of Fundamental Scientific Research in the Russian Federation for a Long term Period (2021-2030) (project no. 122030100170-5).

**Ethics declarations.** This article does not contain any research involving humans or using animals as subjects. The authors declare that they have no conflicts of interest.

## REFERENCES

- Berson, S. A., and Yalow, R. S. (1968) General principles of radioimmunoassay, *Clin. Chim. Acta*, **22**, 51-69, [https://doi.org/10.1016/0009-8981\(68\)90247-7](https://doi.org/10.1016/0009-8981(68)90247-7).
- Rhiel, L., and Becker, S. (2021) Applications of Antibodies in Therapy, Diagnosis, and Science, in *Introduction to Antibody Engineering*, Springer International Publishing, Cham, [https://doi.org/10.1007/978-3-030-54630-4\\_6](https://doi.org/10.1007/978-3-030-54630-4_6).
- Vengesai, A., Midzi, H., Kasambala, M., Mutandadzi, H., Mduluza-Jokonya, T. L., Rusakaniko, S., Mutapi, F., Naicker, T., and Mduluza, T. (2021) A systematic and meta-analysis review on the diagnostic accuracy of antibodies in the serological diagnosis of COVID-19, *Syst. Rev.*, **10**, 155, <https://doi.org/10.1186/s13643-021-01689-3>.
- Stern, M., and Herrmann, R. (2005) Overview of monoclonal antibodies in cancer therapy: present and promise, *Crit. Rev. Oncol. Hematol.*, **54**, 11-29, <https://doi.org/10.1016/j.critrevonc.2004.10.011>.
- Bekes, M., Langlely, D. R., and Crews, C. M. (2022) PROT-AC targeted protein degraders: the past is prologue, *Nat. Rev. Drug Discov.*, **21**, 181-200, <https://doi.org/10.1038/s41573-021-00371-6>.
- Rees, A. R. (2020) Understanding the human antibody repertoire, *MAbs*, **12**, 1729683, <https://doi.org/10.1080/19420862.2020.1729683>.
- Liu, J. K. (2014) The history of monoclonal antibody development – progress, remaining challenges and future innovations, *Ann. Med. Surg. (Lond)*, **3**, 113-116, <https://doi.org/10.1016/j.amsu.2014.09.001>.
- Yang, O., Prabhu, S., and Ierapetritou, M. (2019) Comparison between batch and continuous monoclonal antibody production and economic analysis, *Ind. Eng. Chem. Res.*, **58**, 5851-5863, <https://doi.org/10.1021/acs.iecr.8b04717>.
- Alfaleh, M. A., Alsaab, H. O., Mahmoud, A. B., Alkayyal, A. A., Jones, M. L., Mahler, S. M., and Hashem, A. M. (2020) Phage display derived monoclonal antibodies: from bench to bedside, *Front. Immunol.*, **11**, 1986, <https://doi.org/10.3389/fimmu.2020.01986>.
- Bradbury, A., and Pluckthun, A. (2015) Reproducibility: standardize antibodies used in research, *Nature*, **518**, 27-29, <https://doi.org/10.1038/518027a>.
- Chothia, C., and Lesk, A. M. (1987) Canonical structures for the hypervariable regions of immunoglobulins, *J. Mol. Biol.*, **196**, 901-917, [https://doi.org/10.1016/0022-2836\(87\)90412-8](https://doi.org/10.1016/0022-2836(87)90412-8).
- Sircar, A., Sanni, K. A., Shi, J., and Gray, J. J. (2011) Analysis and modeling of the variable region of camelid single-domain antibodies, *J. Immunol.*, **186**, 6357-6367, <https://doi.org/10.4049/jimmunol.1100116>.
- Almagro, J. C., Teplyakov, A., Luo, J., Sweet, R. W., Kodangattil, S., Hernandez-Guzman, F., and Gilliland, G. L. (2014) Second antibody modeling assessment (AMA-II), *Proteins*, **82**, 1553-1562, <https://doi.org/10.1002/prot.24567>.
- Zhu, K., and Day, T. (2013) Ab initio structure prediction of the antibody hypervariable H3 loop, *Proteins*, **81**, 1081-1089, <https://doi.org/10.1002/prot.24240>.
- Leem, J., Dunbar, J., Georges, G., Shi, J., and Deane, C. M. (2016) ABodyBuilder: Automated antibody structure prediction with data-driven accuracy estimation, *MAbs*, **8**, 1259-1268, <https://doi.org/10.1080/19420862.2016.1205773>.
- Schoeder, C. T., Schmitz, S., Adolf-Bryfogle, J., Sevy, A. M., Finn, J. A., Sauer, M. F., Bozhanova, N. G., Mueller, B. K., Sangha, A. K., Bonet, J., Sheehan, J. H., Kuenze, G., Marlow, B., Smith, S. T., Woods, H., Bender, B. J., Martina, C. E., del Alamo, D., Kodali, P., Gulsevin, A., et al. (2021) Modeling immunity with Rosetta: methods for antibody and antigen design, *Biochemistry*, **60**, 825-846, <https://doi.org/10.1021/acs.biochem.0c00912>.
- Marcatili, P., Rosi, A., and Tramontano, A. (2008) PIGS: automatic prediction of antibody structures, *Bioinformatics*, **24**, 1953-1954, <https://doi.org/10.1093/bioinformatics/btn341>.
- Roy, A., Kucukural, A., and Zhang, Y. (2010) I-TASSER: a unified platform for automated protein structure and function prediction, *Nat. Protoc.*, **5**, 725-738, <https://doi.org/10.1038/nprot.2010.5>.
- Zhu, K., Day, T., Warshaviak, D., Murrett, C., Friesner, R., and Pearlman, D. (2014) Antibody structure determina-



- tion using a combination of homology modeling, energy-based refinement, and loop prediction, *Proteins*, **82**, 1646-1655, <https://doi.org/10.1002/prot.24551>.
20. Whitelegg, N. R., and Rees, A. R. (2000) WAM: an improved algorithm for modelling antibodies on the WEB, *Protein Eng.*, **13**, 819-824, <https://doi.org/10.1093/protein/13.12.819>.
  21. Dunbar, J., Krawczyk, K., Leem, J., Marks, C., Nowak, J., Regep, C., Georges, G., Kelm, S., Popovic, B., and Deane, C. M. (2016) SAbPred: a structure-based antibody prediction server, *Nucleic Acids Res.*, **44**, W474-W478, <https://doi.org/10.1093/nar/gkw361>.
  22. Sircar, A., Kim, E. T., and Gray, J. J. (2009) RosettaAntibody: antibody variable region homology modeling server, *Nucleic Acids Res.*, **37**, W474-W479, <https://doi.org/10.1093/nar/gkp387>.
  23. Weitzner, B. D., Kuroda, D., Marze, N., Xu, J., and Gray, J. J. (2014) Blind prediction performance of RosettaAntibody 3.0: grafting, relaxation, kinematic loop modeling, and full CDR optimization, *Proteins*, **82**, 1611-1623, <https://doi.org/10.1002/prot.24534>.
  24. Su, Z. Y. (2014) Ibalizumab-human CD4 receptor interaction: computational alanine scanning molecular dynamics studies, *Curr. Comput. Aided Drug Des.*, **10**, 217-225, <https://doi.org/10.2174/157340991003150302230124>.
  25. Kiyoshi, M., Caaveiro, J. M., Miura, E., Nagatoishi, S., Nakakido, M., Soga, S., Shirai, H., Kawabata, S., and Tsumoto, K. (2014) 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*, **9**, e87099, <https://doi.org/10.1371/journal.pone.0087099>.
  26. Ferdous, S., and Martin, A. C. R. (2018) AbDb: antibody structure database—a database of PDB-derived antibody structures, *Database (Oxford)*, **2018**, bay040, <https://doi.org/10.1093/database/bay040>.
  27. Kumar, N., Bajiya, N., Patiyal, S., and Raghava, G. P. S. (2023) Multi-perspectives and challenges in identifying B-cell epitopes, *Protein Sci.*, **32**, e4785, <https://doi.org/10.1002/pro.4785>.
  28. Zeng, X., Bai, G., Sun, C., and Ma, B. (2023) Recent Progress in Antibody Epitope Prediction, *Antibodies*, **12**, 52, <https://doi.org/10.3390/antib12030052>.
  29. Dai, B., and Bailey-Kellogg, C. (2021) Protein interaction interface region prediction by geometric deep learning, *Bioinformatics*, **37**, 2580-2588, <https://doi.org/10.1093/bioinformatics/btab154>.
  30. Da Silva, B. M., Myung, Y., Ascher, D. B., and Pires, D. E. V. (2022) epitope3D: a machine learning method for conformational B-cell epitope prediction, *Brief. Bioinform.*, **23**, bbab423, <https://doi.org/10.1093/bib/bbab423>.
  31. Cia, G., Pucci, F., and Rooman, M. (2023) Critical review of conformational B-cell epitope prediction methods, *Brief. Bioinform.*, **24**, bbac567, <https://doi.org/10.1093/bib/bbac567>.
  32. Chu, L. S., Ruffolo, J. A., Harmalkar, A., and Gray, J. J. (2024) Flexible protein-protein docking with a multitask iterative transformer, *Protein Sci.*, **33**, e4862, <https://doi.org/10.1002/pro.4862>.
  33. Gray, J. J., Moughon, S., Wang, C., Schueler-Furman, O., Kuhlman, B., Rohl, C. A., and Baker, D. (2003) Protein-protein docking with simultaneous optimization of rigid-body displacement and side-chain conformations, *J. Mol. Biol.*, **331**, 281-299, [https://doi.org/10.1016/s0022-2836\(03\)00670-3](https://doi.org/10.1016/s0022-2836(03)00670-3).
  34. Sircar, A., and Gray, J. J. (2010) SnugDock: paratope structural optimization during antibody-antigen docking compensates for errors in antibody homology models, *PLoS Comput. Biol.*, **6**, e1000644, <https://doi.org/10.1371/journal.pcbi.1000644>.
  35. Zhao, J., Nussinov, R., and Ma, B. (2017) Mechanisms of recognition of amyloid-beta (Abeta) monomer, oligomer, and fibril by homologous antibodies, *J. Biol. Chem.*, **292**, 18325-18343, <https://doi.org/10.1074/jbc.m117.801514>.
  36. Guest, J. D., Vreven, T., Zhou, J., Moal, I., Jeliazkov, J. R., Gray, J. J., Weng, Z., and Pierce, B. G. (2021) An expanded benchmark for antibody-antigen docking and affinity prediction reveals insights into antibody recognition determinants, *Structure*, **29**, 606-621.e605, <https://doi.org/10.1016/j.str.2021.01.005>.
  37. Mashiach, E., Nussinov, R., and Wolfson, H. J. (2010) FiberDock: a web server for flexible induced-fit backbone refinement in molecular docking, *Nucleic Acids Res.*, **38**, W457-W461, <https://doi.org/10.1093/nar/gkq373>.
  38. Pierce, B., and Weng, Z. (2007) ZRANK: reranking protein docking predictions with an optimized energy function, *Proteins*, **67**, 1078-1086, <https://doi.org/10.1002/prot.21373>.
  39. Cheng, T. M., Blundell, T. L., and Fernandez-Recio, J. (2007) pyDock: electrostatics and desolvation for effective scoring of rigid-body protein-protein docking, *Proteins*, **68**, 503-515, <https://doi.org/10.1002/prot.21419>.
  40. Lorenzen, S., and Zhang, Y. (2007) Identification of near-native structures by clustering protein docking conformations, *Proteins*, **68**, 187-194, <https://doi.org/10.1002/prot.21442>.
  41. Ritchie, D. W., and Venkatraman, V. (2010) Ultra-fast FFT protein docking on graphics processors, *Bioinformatics*, **26**, 2398-2405, <https://doi.org/10.1093/bioinformatics/btq444>.
  42. Zhu, F., Bourguet, F. A., Bennett, W. F. D., Lau, E. Y., Arrildt, K. T., Segelke, B. W., Zemla, A. T., Desautels, T. A., and Faissol, D. M. (2022) Large-scale application of free energy perturbation calculations for antibody design, *Sci. Rep.*, **12**, 12489, <https://doi.org/10.1038/s41598-022-14443-z>.

43. Zhou, Y., Myung, Y., Rodrigues, C. H. M., and Ascher, D. B. (2024) DDMut-PPI: predicting effects of mutations on protein–protein interactions using graph-based deep learning, *Nucleic Acids Res.*, **52**, W207–W214, <https://doi.org/10.1093/nar/gkae412>.
44. Wang, M., Zhu, D., Zhu, J., Nussinov, R., and Ma, B. (2018) Local and global anatomy of antibody-protein antigen recognition, *J. Mol. Recognit.*, **31**, e2693, <https://doi.org/10.1002/jmr.2693>.
45. Xue, L. C., Rodrigues, J. P., Kastritis, P. L., Bonvin, A. M., and Vangone, A. (2016) PRODIGY: a web server for predicting the binding affinity of protein-protein complexes, *Bioinformatics*, **32**, 3676–3678, <https://doi.org/10.1093/bioinformatics/btw514>.
46. Chowdhury, R., Allan, M. F., and Maranas, C. D. (2018) OptMAVEN-2.0: *de novo* design of variable antibody regions against targeted antigen epitopes, *Antibodies (Basel)*, **7**, 23, <https://doi.org/10.3390/antib7030023>.
47. Poosarla, V. G., Li, T., Goh, B. C., Schulten, K., Wood, T. K., and Maranas, C. D. (2017) Computational *de novo* design of antibodies binding to a peptide with high affinity, *Biotechnol. Bioeng.*, **114**, 1331–1342, <https://doi.org/10.1002/bit.26244>.
48. Hummer, A. M., Abanades, B., and Deane, C. M. (2022) Advances in computational structure-based antibody design, *Curr. Opin. Struct. Biol.*, **74**, 102379, <https://doi.org/10.1016/j.sbi.2022.102379>.
49. Yin, R., and Pierce, B. G. (2023) Evaluation of AlphaFold antibody-antigen modeling with implications for improving predictive accuracy, *bioRxiv*, <https://doi.org/10.1101/2023.07.05.547832>.
50. Olsen, T. H., Moal, I. H., and Deane, C. M. (2022) AbLang: an antibody language model for completing antibody sequences, *Bioinform. Adv.*, **2**, vbac046, <https://doi.org/10.1093/bioadv/vbac046>.
51. Shuai, R. W., Ruffolo, J. A., and Gray, J. J. (2023) IgLM: Infilling language modeling for antibody sequence design, *Cell Syst.*, **14**, 979–989, <https://doi.org/10.1016/j.cels.2023.10.001>.
52. Akbar, R., Robert, P. A., Pavlovic, M., Jeliakov, J. R., Snapkov, I., Slabodkin, A., Weber, C. R., Scheffer, L., Miho, E., Haff, I. H., Haug, D. T. T., Lund-Johansen, F., Safonova, Y., Sandve, G. K., and Greiff, V. (2021) A compact vocabulary of paratope-epitope interactions enables predictability of antibody-antigen binding, *Cell Rep.*, **34**, 108856, <https://doi.org/10.1016/j.celrep.2021.108856>.
53. Bertoline, L. M. F., Lima, A. N., Krieger, J. E., and Teixeira, S. K. (2023) Before and after AlphaFold2: an overview of protein structure prediction, *Front. Bioinform.*, **3**, 1120370, <https://doi.org/10.3389/fbinf.2023.1120370>.
54. Lai, P. K., Gallegos, A., Mody, N., Sathish, H. A., and Trout, B. L. (2022) Machine learning prediction of antibody aggregation and viscosity for high concentration formulation development of protein therapeutics, *MABs*, **14**, 2026208, <https://doi.org/10.1080/19420862.2022.2026208>.
55. Marks, C., Hummer, A. M., Chin, M., and Deane, C. M. (2021) Humanization of antibodies using a machine learning approach on large-scale repertoire data, *Bioinformatics*, **37**, 4041–4047, <https://doi.org/10.1093/bioinformatics/btab434>.
56. Grinshpun, B., Thorsteinson, N., Pereira, J. N., Ripmann, F., Nannemann, D., Sood, V. D., and Fomekong Nanfack, Y. (2021) Identifying biophysical assays and in silico properties that enrich for slow clearance in clinical-stage therapeutic antibodies, *MABs*, **13**, 1932230, <https://doi.org/10.1080/19420862.2021.1932230>.
57. Reis, P., Barletta, G. P., Gagliardi, L., Fortuna, S., Soler, M. A., and Rocchia, W. (2022) Antibody-antigen binding interface analysis in the big data era, *Front. Mol. Biosci.*, **9**, 945808, <https://doi.org/10.3389/fmolb.2022.945808>.
58. Crivianu-Gaita, V., and Thompson, M. (2016) Aptamers, antibody scFv, and antibody Fab' fragments: An overview and comparison of three of the most versatile biosensor biorecognition elements, *Biosens. Bioelectron.*, **85**, 32–45, <https://doi.org/10.1016/j.bios.2016.04.091>.
59. De Marco, A. (2020) Recombinant expression of nanobodies and nanobody-derived immunoreagents, *Protein Expr. Purif.*, **172**, 105645, <https://doi.org/10.1016/j.pep.2020.105645>.
60. Zelenovic, N., Filipovic, L., and Popovic, M. (2023) Recent developments in bioprocessing of recombinant antibody fragments, *Biochemistry (Moscow)*, **88**, 1191–1204, <https://doi.org/10.1134/s0006297923090018>.
61. Fink, M., Schimek, C., Cserjan-Puschmann, M., Reinisch, D., Brocard, C., Hahn, R., and Striedner, G. (2021) Integrated process development: the key to improve Fab production in *E. coli*, *Biotechnol. J.*, **16**, e2000562, <https://doi.org/10.1002/biot.202000562>.
62. Mei, Y., Chen, Y., Sivaccumar, J. P., An, Z., Xia, N., and Luo, W. (2022) Research progress and applications of nanobody in human infectious diseases, *Front. Pharmacol.*, **13**, 963978, <https://doi.org/10.3389/fphar.2022.963978>.
63. Yu, X., Yang, Y. P., Dikici, E., Deo, S. K., and Daunert, S. (2017) Beyond antibodies as binding partners: the role of antibody mimetics in bioanalysis, *Annu. Rev. Anal. Chem. (Palo Alto Calif.)*, **10**, 293–320, <https://doi.org/10.1146/annurev-anchem-061516-045205>.
64. Tang, J., Zhang, C., Castillo, N. C., Lalaurie, C. J., Gao, X., Dalby, P. A., and Kozielski, F. (2023) Crystal structures and molecular dynamics simulations of a humanised antibody fragment at acidic to basic pH, *Sci. Rep.*, **13**, 16281, <https://doi.org/10.1038/s41598-023-42698-7>.
65. Paloni, M., and Cavallotti, C. (2017) Molecular modeling of the interaction of protein L with antibodies, *ACS Omega*, **2**, 6464–6472, <https://doi.org/10.1021/acsomega.7b01123>.

66. Bai, G., Ge, Y., Su, Y., Chen, S., Zeng, X., Lu, H., and Ma, B. (2021) Computational construction of a single-chain bi-paratopic antibody allosterically inhibiting TCR-staphylococcal enterotoxin b binding, *Front. Immunol.*, **12**, 732938, <https://doi.org/10.3389/fimmu.2021.732938>.
67. Bird, R. E., Hardman, K. D., Jacobson, J. W., Johnson, S., Kaufman, B. M., Lee, S. M., Lee, T., Pope, S. H., Riordan, G. S., and Whitlow, M. (1988) Single-chain antigen-binding proteins, *Science*, **242**, 423-426, <https://doi.org/10.1126/science.3140379>.
68. Cohen, T., Halfon, M., and Schneidman-Duhovny, D. (2022) NanoNet: rapid and accurate end-to-end nanobody modeling by deep learning, *Front. Immunol.*, **13**, 958584, <https://doi.org/10.3389/fimmu.2022.958584>.
69. Harmsen, M. M., and De Haard, H. J. (2007) Properties, production, and applications of camelid single-domain antibody fragments, *Appl. Microbiol. Biotechnol.*, **77**, 13-22, <https://doi.org/10.1007/s00253-007-1142-2>.
70. David, T. I., Pestov, N. B., Korneenko, T. V., and Barlev, N. A. (2023) Non-immunoglobulin synthetic binding proteins for oncology, *Biochemistry (Moscow)*, **88**, 1232-1247, <https://doi.org/10.1134/s0006297923090043>.
71. Skerra, A. (2000) Engineered protein scaffolds for molecular recognition, *J. Mol. Recognit.*, **13**, 167-187, [https://doi.org/10.1002/1099-1352\(200007/08\)13:4<167::aid-jmr502>3.0.co;2-9](https://doi.org/10.1002/1099-1352(200007/08)13:4<167::aid-jmr502>3.0.co;2-9).
72. Moody, P., Chudasama, V., Nathani, R. I., Maruani, A., Martin, S., Smith, M. E., and Caddick, S. (2014) A rapid, site-selective and efficient route to the dual modification of DARPin, *Chem. Commun. (Camb.)*, **50**, 4898-4900, <https://doi.org/10.1039/c4cc00053f>.
73. Liu, S., Wu, K., and Chen, C. (2022) Obtaining protein foldability information from computational models of AlphaFold2 and RoseTTAFold, *Comput. Struct. Biotechnol. J.*, **20**, 4481-4489, <https://doi.org/10.1016/j.csbj.2022.08.034>.
74. Abramson, J., Adler, J., Dunger, J., Evans, R., Green, T., Pritzel, A., Ronneberger, O., Willmore, L., Ballard, A. J., Bambrick, J., Bodenstein, S. W., Evans, D. A., Hung, C.-C., O'Neill, M., Reiman, D., Tunyasuvunakool, K., Wu, Z., Žemgulytė, A., Arvaniti, E., Beattie, C., et al. (2024) Accurate structure prediction of biomolecular interactions with AlphaFold 3, *Nature*, **630**, 493-500, <https://doi.org/10.1038/s41586-024-07487-w>.
75. Ingraham, J. B., Baranov, M., Costello, Z., Barber, K. W., Wang, W., Ismail, A., Frappier, V., Lord, D. M., Ng-Thow-Hing, C., Van Vlack, E. R., Tie, S., Xue, V., Cowles, S. C., Leung, A., Rodrigues, J. V., Morales-Perez, C. L., Ayoub, A. M., Green, R., Puentes, K., Oplinger, F., et al. (2023) Illuminating protein space with a programmable generative model, *Nature*, **623**, 1070-1078, <https://doi.org/10.1038/s41586-023-06728-8>.
76. Anishchenko, I., Pellock, S. J., Chidyausiku, T. M., Ramelot, T. A., Ovchinnikov, S., Hao, J., Bafna, K., Norn, C., Kang, A., Bera, A. K., DiMaio, F., Carter, L., Chow, C. M., Montelione, G. T., Baker, D., Anishchenko, I., Pellock, S. J., Chidyausiku, T. M., Ramelot, T. A., Ovchinnikov, S., et al. (2021) De novo protein design by deep network hallucination, *Nature*, **600**, 7889, <https://doi.org/10.1038/s41586-021-04184-w>.
77. Ponomarenko, J., Papangelopoulos, N., Zajonc, D. M., Peters, B., Sette, A., and Bourne, P. E. (2011) IEDB-3D: structural data within the immune epitope database, *Nucleic Acids Res.*, **39**, D1164-D1170, <https://doi.org/10.1093/nar/gkq888>.
78. Schlessinger, A., Ofran, Y., Yachdav, G., and Rost, B. (2006) Epitome: database of structure-inferred antigenic epitopes, *Nucleic Acids Res.*, **34**, D777-D780, <https://doi.org/10.1093/nar/gkj053>.
79. Saha, S., Bhasin, M., and Raghava, G. P. (2005) Bcipep: a database of B-cell epitopes, *BMC Genomics*, **6**, 79, <https://doi.org/10.1186/1471-2164-6-79>.
80. Toseland, C. P., Clayton, D. J., McSparron, H., Hemsley, S. L., Blythe, M. J., Paine, K., Doytchinova, I. A., Guan, P., Hattotuwigama, C. K., and Flower, D. R. (2005) AntiJen: a quantitative immunology database integrating functional, thermodynamic, kinetic, biophysical, and cellular data, *Immunome Res.*, **1**, 4, <https://doi.org/10.1186/1745-7580-1-4>.
81. Peters, B., Sidney, J., Bourne, P., Bui, H. H., Buus, S., Doh, G., Fleri, W., Kronenberg, M., Kubo, R., Lund, O., Nemazee, D., Ponomarenko, J. V., Sathiamurthy, M., Schoenberger, S., Stewart, S., Surko, P., Way, S., Wilson, S., and Sette, A. (2005) The immune epitope database and analysis resource: from vision to blueprint, *PLoS Biol.*, **3**, e91, <https://doi.org/10.1371/journal.pbio.0030091>.
82. Fleishman, S. J., Corn, J. E., Strauch, E. M., Whitehead, T. A., Karanicolas, J., and Baker, D. (2011) Hotspot-centric de novo design of protein binders, *J. Mol. Biol.*, **413**, 1047-1062, <https://doi.org/10.1016/j.jmb.2011.09.001>.
83. Fleishman, S. J., Whitehead, T. A., Ekiert, D. C., Dreyfus, C., Corn, J. E., Strauch, E. M., Wilson, I. A., and Baker, D. (2011) Computational design of proteins targeting the conserved stem region of influenza hemagglutinin, *Science*, **332**, 816-821, <https://doi.org/10.1126/science.1202617>.
84. Ellington, A. D., and Szostak, J. W. (1990) In vitro selection of RNA molecules that bind specific ligands, *Nature*, **346**, 818-822, <https://doi.org/10.1038/346818a0>.
85. Tuerk, C., and Gold, L. (1990) Systematic evolution of ligands by exponential enrichment: RNA ligands to bacteriophage T4 DNA polymerase, *Science*, **249**, 505-510, <https://doi.org/10.1126/science.2200121>.
86. Mironov, A. S., Gusarov, I., Rafikov, R., Lopez, L. E., Shatalin, K., Kreneva, R. A., Perumov, D. A., and Nudler, E. (2002) Sensing small molecules by nascent RNA: a mechanism to control transcription in bacteria, *Cell*, **111**, 747-756, [https://doi.org/10.1016/s0092-8674\(02\)01134-0](https://doi.org/10.1016/s0092-8674(02)01134-0).

87. Sullenger, B. A., Gallardo, H. F., Ungers, G. E., and Gilboa, E. (1991) Analysis of trans-acting response decoy RNA-mediated inhibition of human immunodeficiency virus type 1 transactivation, *J. Virol.*, **65**, 6811-6816, <https://doi.org/10.1128/jvi.65.12.6811-6816.1991>.
88. Chen, A., and Yang, S. (2015) Replacing antibodies with aptamers in lateral flow immunoassay, *Biosens. Bioelectron.*, **71**, 230-242, <https://doi.org/10.1016/j.bios.2015.04.041>.
89. Zhou, J., and Rossi, J. (2017) Aptamers as targeted therapeutics: current potential and challenges, *Nat. Rev. Drug Discov.*, **16**, 181-202, <https://doi.org/10.1038/nrd.2016.199>.
90. Cho, E. J., Lee, J. W., and Ellington, A. D. (2009) Applications of aptamers as sensors, *Annu. Rev. Anal. Chem. (Palo Alto Calif.)*, **2**, 241-264, <https://doi.org/10.1146/annurev.anchem.1.031207.112851>.
91. Mahmoud, M., Deigner, H.-P., Mahmoud, M., and Deigner, H.-P. (2015) Aptamers in diagnostics: replacing or complementing antibodies? *J. Bioeng. Bioelectron.*, **5**, e129, <https://doi.org/10.4172/2153-0777.1000e129>.
92. Ng, E. W., Shima, D. T., Calias, P., Cunningham, E. T., Jr., Guyer, D. R., and Adamis, A. P. (2006) Pegaptanib, a targeted anti-VEGF aptamer for ocular vascular disease, *Nat. Rev. Drug Discov.*, **5**, 123-132, <https://doi.org/10.1038/nrd1955>.
93. Kruspe, S., Mittelberger, F., Szameit, K., and Hahn, U. (2014) Aptamers as drug delivery vehicles, *ChemMedChem*, **9**, 1998-2011, <https://doi.org/10.1002/cmdc.201402163>.
94. Olson, W. C., Heston, W. D., and Rajasekaran, A. K. (2007) Clinical trials of cancer therapies targeting prostate-specific membrane antigen, *Rev. Recent Clin. Trials*, **2**, 182-190, <https://doi.org/10.2174/157488707781662724>.
95. Zhou, J., Swiderski, P., Li, H., Zhang, J., Neff, C. P., Akkina, R., and Rossi, J. J. (2009) Selection, characterization and application of new RNA HIV gp 120 aptamers for facile delivery of Dicer substrate siRNAs into HIV infected cells, *Nucleic Acids Res.*, **37**, 3094-3109, <https://doi.org/10.1093/nar/gkp185>.
96. Darmostuk, M., Rimpelova, S., Gbelcova, H., and Ruml, T. (2015) Current approaches in SELEX: An update to aptamer selection technology, *Biotechnol. Adv.*, **33**, 1141-1161, <https://doi.org/10.1016/j.biotechadv.2015.02.008>.
97. Luo, X., McKeague, M., Pitre, S., Dumontier, M., Green, J., Golshani, A., Derosa, M. C., and Dehne, F. (2010) Computational approaches toward the design of pools for the *in vitro* selection of complex aptamers, *RNA*, **16**, 2252-2262, <https://doi.org/10.1261/rna.2102210>.
98. Gevertz, J., Gan, H. H., and Schlick, T. (2005) *In vitro* RNA random pools are not structurally diverse: a computational analysis, *RNA*, **11**, 853-863, <https://doi.org/10.1261/rna.7271405>.
99. Lipi, F., Chen, S., Chakravarthy, M., Rakesh, S., and Veedu, R. N. (2016) *In vitro* evolution of chemically-modified nucleic acid aptamers: *Pros* and *cons*, and comprehensive selection strategies, *RNA Biol.*, **13**, 1232-1245, <https://doi.org/10.1080/15476286.2016.1236173>.
100. Verma, V., Gupta, A., and Chaudhary, V. K. (2020) Emulsion PCR made easy, *BioTechniques*, **69**, 421-426, <https://doi.org/10.2144/btn-2019-0161>.
101. Kohlberger, M., and Gadermaier, G. (2022) SELEX: critical factors and optimization strategies for successful aptamer selection, *Biotechnol. Appl. Biochem.*, **69**, 1771-1792, <https://doi.org/10.1002/bab.2244>.
102. Betz, K., Malyshev, D. A., Lavergne, T., Welte, W., Diederichs, K., Dwyer, T. J., Ordoukhanian, P., Romesberg, F. E., and Marx, A. (2012) KlenTaq polymerase replicates unnatural base pairs by inducing a Watson-Crick geometry, *Nat. Chem. Biol.*, **8**, 612-614, <https://doi.org/10.1038/nchembio.966>.
103. Boltz, A., Piater, B., Toleikis, L., Guenther, R., Kolmar, H., and Hock, B. (2011) Bi-specific aptamers mediating tumor cell lysis, *J. Biol. Chem.*, **286**, 21896-21905, <https://doi.org/10.1074/jbc.m111.238261>.
104. Li, Y., and Breaker, R. R. (1999) Kinetics of RNA degradation by specific base catalysis of transesterification involving the 2'-hydroxyl group, *J. Am. Chem. Soc.*, **121**, 5364-5372, <https://doi.org/10.1021/ja990592p>.
105. Nagy, A., Gertsenstein, M., Vintersten, K., and Behringer, R. (2007) General procedures for avoiding contamination with RNase, *CSH Protoc.*, **2007**, pdb.ip43, <https://doi.org/10.1101/pdb.ip43>.
106. Darfeuille, F. (2004) LNA/DNA chimeric oligomers mimic RNA aptamers targeted to the TAR RNA element of HIV-1, *Nucleic Acids Res.*, **32**, 3101-3107, <https://doi.org/10.1093/nar/gkh636>.
107. Odeh, F., Nsairat, H., Alshaer, W., Ismail, M. A., Esawi, E., Qaqish, B., Bawab, A. A., and Ismail, S. I. (2019) Aptamers chemistry: chemical modifications and conjugation strategies, *Molecules*, **25**, 3, <https://doi.org/10.3390/molecules25010003>.
108. Chen, Z., Luo, H., Gubu, A., Yu, S., Zhang, H., Dai, H., Zhang, Y., Zhang, B., Ma, Y., Lu, A., and Zhang, G. (2023) Chemically modified aptamers for improving binding affinity to the target proteins via enhanced non-covalent bonding, *Front. Cell Dev. Biol.*, **11**, 1091809, <https://doi.org/10.3389/fcell.2023.1091809>.
109. Li, X., Chung, L. W., and Li, G. (2016) Multiscale simulations on spectral tuning and the photoisomerization mechanism in fluorescent RNA spinach, *J. Chem. Theory Comput.*, **12**, 5453-5464, <https://doi.org/10.1021/acs.jctc.6b00578>.
110. Sato, R., Suzuki, K., Yasuda, Y., Suenaga, A., and Fukui, K. (2022) RNAapt3D: RNA aptamer 3D-structural modeling database, *Biophys. J.*, **121**, 4770-4776, <https://doi.org/10.1016/j.bpj.2022.09.023>.
111. Ahmad, N. A., Mohamed Zulkifli, R., Hussin, H., and Nadri, M. H. (2021) *In silico* approach for

- Post-SELEX DNA aptamers: a mini-review, *J. Mol. Graph. Model.*, **105**, 107872, <https://doi.org/10.1016/j.jmgm.2021.107872>.
112. Buglak, A. A., Samokhvalov, A. V., Zherdev, A. V., and Dzantiev, B. B. (2020) Methods and applications of in silico aptamer design and modeling, *Int. J. Mol. Sci.*, **21**, 8420, <https://doi.org/10.3390/ijms21228420>.
113. Kumar, P., and Kumar, A. (2020) Nucleobase sequence based building up of reliable QSAR models with the index of ideality correlation using Monte Carlo method, *J. Biomol. Struct. Dyn.*, **38**, 3296-3306, <https://doi.org/10.1080/07391102.2019.1656109>.
114. Torkamanian-Afshar, M., Nematzadeh, S., Tabar-zad, M., Najafi, A., Lanjanian, H., and Masoudi-Nejad, A. (2021) *In silico* design of novel aptamers utilizing a hybrid method of machine learning and genetic algorithm, *Mol. Divers.*, **25**, 1395-1407, <https://doi.org/10.1007/s11030-021-10192-9>.
115. Bashir, A., Yang, Q., Wang, J., Hoyer, S., Chou, W., McLean, C., Davis, G., Gong, Q., Armstrong, Z., Jang, J., Kang, H., Pawlosky, A., Scott, A., Dahl, G. E., Berndt, M., Dimon, M., and Ferguson, B. S. (2021) Machine learning guided aptamer refinement and discovery, *Nat. Commun.*, **12**, 2366, <https://doi.org/10.1038/s41467-021-22555-9>.
116. Matarrese, M. A. G., Loppini, A., Nicoletti, M., Filippi, S., and Chiodo, L. (2023) Assessment of tools for RNA secondary structure prediction and extraction: a final-user perspective, *J. Biomol. Struct. Dyn.*, **41**, 6917-6936, <https://doi.org/10.1080/07391102.2022.2116110>.
117. Sato, K., Akiyama, M., and Sakakibara, Y. (2021) RNA secondary structure prediction using deep learning with thermodynamic integration, *Nat. Commun.*, **12**, 941, <https://doi.org/10.1038/s41467-021-21194-4>.
118. Nasaev, S. S., Mukanov, A. R., Kuznetsov, I. I., and Veselovsky, A. V. (2023) AliNA – a deep learning program for RNA secondary structure prediction, *Mol. Inform.*, **42**, e202300113, <https://doi.org/10.1002/minf.202300113>.
119. Khoshbin, Z., and Housaindokht, M. R. (2021) Computer-aided aptamer design for sulfadimethoxine antibiotic: step by step mutation based on MD simulation approach, *J. Biomol. Struct. Dyn.*, **39**, 3071-3079, <https://doi.org/10.1080/07391102.2020.1760133>.
120. Hsieh, P. C., Lin, H. T., Chen, W. Y., Tsai, J. J. P., and Hu, W. P. (2017) The combination of computational and biosensing technologies for selecting aptamer against prostate specific antigen, *Biomed. Res. Int.*, **2017**, 5041683, <https://doi.org/10.1155/2017/5041683>.
121. Shcherbinin, D. S., Gnedenko, O. V., Khmeleva, S. A., Usanov, S. A., Gilep, A. A., Yantsevich, A. V., Shkel, T. V., Yushkevich, I. V., Radko, S. P., Ivanov, A. S., Veselovsky, A. V., and Archakov, A. I. (2015) Computer-aided design of aptamers for cytochrome p450, *J. Struct. Biol.*, **191**, 112-119, <https://doi.org/10.1016/j.jsb.2015.07.003>.
122. Oliveira, R., Pinho, E., Sousa, A. L., Dias, O., Azevedo, N. F., and Almeida, C. (2022) Modelling aptamers with nucleic acid mimics (NAM): from sequence to three-dimensional docking, *PLoS One*, **17**, e0264701, <https://doi.org/10.1371/journal.pone.0264701>.
123. Binet, T., Padiolleau-Lefevre, S., Octave, S., Avalle, B., and Maffucci, I. (2023) Comparative study of single-stranded oligonucleotides secondary structure prediction tools, *BMC Bioinformatics*, **24**, 422, <https://doi.org/10.1186/s12859-023-05532-5>.
124. Zhao, Y., Huang, Y., Gong, Z., Wang, Y., Man, J., and Xiao, Y. (2012) Automated and fast building of three-dimensional RNA structures, *Sci. Rep.*, **2**, 734, <https://doi.org/10.1038/srep00734>.
125. Adler, A., Forster, N., Homann, M., and Goringe, H. U. (2008) Post-SELEX chemical optimization of a trypanosome-specific RNA aptamer, *Comb. Chem. High Throughput Screen.*, **11**, 16-23, <https://doi.org/10.2174/138620708783398331>.
126. Galindo-Murillo, R., Cohen, J. S., and Akabayov, B. (2021) Molecular dynamics simulations of acyclic analogs of nucleic acids for antisense inhibition, *Mol. Ther. Nucleic Acids*, **23**, 527-535, <https://doi.org/10.1016/j.omtn.2020.11.023>.
127. Sponer, J., Banas, P., Jurecka, P., Zgarbova, M., Kuhrova, P., Havrila, M., Krepl, M., Stadlbauer, P., and Otyepka, M. (2014) Molecular dynamics simulations of nucleic acids. From tetranucleotides to the ribosome, *J. Phys. Chem. Lett.*, **5**, 1771-1782, <https://doi.org/10.1021/jz500557y>.
128. Wulff, G., Sarhan, A., and Zabrocki, K. (1973) Enzyme-analogue built polymers and their use for the resolution of racemates, *Tetrahedron Lett.*, **14**, 4329-4332, [https://doi.org/10.1016/s0040-4039\(01\)87213-0](https://doi.org/10.1016/s0040-4039(01)87213-0).
129. Haupt, K., Medina Rangel, P. X., and Bui, B. T. S. (2020) Molecularly imprinted polymers: antibody mimics for bioimaging and therapy, *Chem. Rev.*, **120**, 9554-9582, <https://doi.org/10.1021/acs.chemrev.0c00428>.
130. Shumyantseva, V. V., Bulko, T. V., Baychorov, I. Kh., and Archakov, A. I. (2015) Molecularly imprinted polymers in electro analysis of proteins, *Biomed. Khim.*, **61**, 325-331, <https://doi.org/10.18097/pbmc20156103325>.
131. Peeters, M., Eersels, K., Junkers, T., and Wagner, P. (2016) *Molecularly Imprinted Polymers: Synthetic Receptors for Diagnostic Medical Devices*, Elsevier, Amsterdam, <https://doi.org/10.1016/b978-0-12-801301-4.00013-x>.
132. Bossi, A., Bonini, F., Turner, A. P., and Piletsky, S. A. (2007) Molecularly imprinted polymers for the recognition of proteins: the state of the art, *Biosens. Bioelectron.*, **22**, 1131-1137, <https://doi.org/10.1016/j.bios.2006.06.023>.
133. Scorrano, S., Mergola, L., Del Sole, R., and Vasapollo, G. (2011) Synthesis of molecularly imprinted polymers for amino acid derivatives by using different functional monomers, *Int. J. Mol. Sci.*, **12**, 1735-1743, <https://doi.org/10.3390/ijms12031735>.

134. Piletsky, S. A., Turner, N. W., and Laitenberger, P. (2006) Molecularly imprinted polymers in clinical diagnostics – future potential and existing problems, *Med. Eng. Phys.*, **28**, 971-977, <https://doi.org/10.1016/j.medengphy.2006.05.004>.
135. Puoci, F., Iemma, F., and Picci, N. (2008) Stimuli-responsive molecularly imprinted polymers for drug delivery: a review, *Curr. Drug Deliv.*, **5**, 85-96, <https://doi.org/10.2174/156720108783954888>.
136. Emi Shaqiza, A., Mohd Noor, A., Islam, A. K. M. S., Dahyar, A., and Iqmal, T. (2011) Porogen effect towards the quality of curcumin imprinted polymer, *Indones. J. Chem.*, **11**, 207-211, <https://doi.org/10.22146/ijc.21381>.
137. Ramstrom, O., and Mosbach, K. (1999) Synthesis and catalysis by molecularly imprinted materials, *Curr. Opin. Chem. Biol.*, **3**, 759-764, [https://doi.org/10.1016/s1367-5931\(99\)00037-x](https://doi.org/10.1016/s1367-5931(99)00037-x).
138. Parisi, O. I., Francomano, F., Dattilo, M., Patitucci, F., Prete, S., Amone, F., and Puoci, F. (2022) The evolution of molecular recognition: from antibodies to molecularly imprinted polymers (MIPs) as artificial counterpart, *J. Funct. Biomater.*, **13**, 12, <https://doi.org/10.3390/jfb13010012>.
139. Vasapollo, G., Sole, R. D., Mergola, L., Lazzoi, M. R., Scardino, A., Scorrano, S., and Mele, G. (2011) Molecularly imprinted polymers: present and future prospective, *Int. J. Mol. Sci.*, **12**, 5908-5945, <https://doi.org/10.3390/ijms12095908>.
140. Dirion, B., Cobb, Z., Schillinger, E., Andersson, L. I., and Sellergren, B. (2003) Water-compatible molecularly imprinted polymers obtained via high-throughput synthesis and experimental design, *J. Am. Chem. Soc.*, **125**, 15101-15109, <https://doi.org/10.1021/ja0355473>.
141. Hoshino, Y., Kodama, T., Okahata, Y., and Shea, K. J. (2008) Peptide imprinted polymer nanoparticles: a plastic antibody, *J. Am. Chem. Soc.*, **130**, 15242-15243, <https://doi.org/10.1021/ja08062875>.
142. Lee, M.-H., Thomas, J. L., Liao, C.-L., Jurcevic, S., Crnogorac-Jurcevic, T., and Lin, H.-Y. (2017) Polymers imprinted with three REG1B peptides for electrochemical determination of Regenerating Protein 1B, a urinary biomarker for pancreatic ductal adenocarcinoma, *Microchim. Acta*, **184**, 1773-1780, <https://doi.org/10.1007/s00604-017-2169-4>.
143. Urraca, J. L., Hall, A. J., Moreno-Bondi, M. C., and Sellergren, B. (2006) A stoichiometric molecularly imprinted polymer for the class-selective recognition of antibiotics in aqueous media, *Angew. Chem. Int. Ed. Engl.*, **45**, 5158-5161, <https://doi.org/10.1002/anie.200601636>.
144. Xu, R., Tian, J., Guan, Y., and Zhang, Y. (2019) Peptide-cross-linked protein-imprinted polymers: easy template removal and excellent imprinting effect, *CCS Chem.*, **1**, 544-552, <https://doi.org/10.31635/ccschem.019.20190030>.
145. Ansari, S., and Masoum, S. (2019) Molecularly imprinted polymers for capturing and sensing proteins: current progress and future implications, *Trends Anal. Chem.*, **114**, 29-47, <https://doi.org/10.1016/j.trac.2019.02.008>.
146. Rajpal, S., Mishra, P., and Mizaikoff, B. (2023) Rational *in silico* design of molecularly imprinted polymers: current challenges and future potential, *Int. J. Mol. Sci.*, **24**, 6785, <https://doi.org/10.3390/ijms24076785>.
147. Nicholls, I. A., Golker, K., Olsson, G. D., Suriyanarayanan, S., and Wiklander, J. G. (2021) The use of computational methods for the development of molecularly imprinted polymers, *Polymers (Basel)*, **13**, 2841, <https://doi.org/10.3390/polym13172841>.
148. Liu, Z., Xu, Z., Wang, D., Yang, Y., Duan, Y., Ma, L., Lin, T., and Liu, H. (2021) A review on molecularly imprinted polymers preparation by computational simulation-aided methods, *Polymers (Basel)*, **13**, 2657, <https://doi.org/10.3390/polym13162657>.
149. Olsson, G. D., Wiklander, J. G., and Nicholls, I. A. (2021) Using molecular dynamics in the study of molecularly imprinted polymers, *Methods Mol. Biol.*, **2359**, 241-268, [https://doi.org/10.1007/978-1-0716-1629-1\\_21](https://doi.org/10.1007/978-1-0716-1629-1_21).
150. Douykhunklaw, C., and Sutthibutpong, T. (2022) A molecular dynamics study on the diffusion and imprint ability of spectinomycin under different sizes of aniline oligomers, *J. Mol. Model.*, **28**, 387, <https://doi.org/10.1007/s00894-022-05371-w>.
151. Suryana, S., Mutakin, M., Rosandi, Y., and Hasanah, A. N. (2022) Molecular dynamic study of mechanism underlying nature of molecular recognition and the role of crosslinker in the synthesis of salmeterol-targeting molecularly imprinted polymer for analysis of salmeterol xinafoate in biological fluid, *Molecules*, **27**, 3619, <https://doi.org/10.3390/molecules27113619>.
152. Rebelo, P., Pacheco, J. G., Voroshylova, I. V., Seguro, I., Cordeiro, M., and Delerue-Matos, C. (2022) Computational modelling and sustainable synthesis of a highly selective electrochemical MIP-based sensor for citalopram detection, *Molecules*, **27**, 3315, <https://doi.org/10.3390/molecules27103315>.
153. Daniels, E., Mustafa, Y. L., Herdes, C., and Leese, H. S. (2021) Optimization of cortisol-selective molecularly imprinted polymers enabled by molecular dynamics simulations, *ACS Appl. Bio Mater.*, **4**, 7243-7253, <https://doi.org/10.1021/acsbm.1c00774>.
154. Garcia, Y., Usuga, B. A., Campos, C. H., Alderete, J. B., and Jimenez, V. A. (2021) NanoMIPs design for fucose and mannose recognition: a molecular dynamics approach, *J. Chem. Inf. Model.*, **61**, 2048-2061, <https://doi.org/10.1021/acs.jcim.0c01446>.
155. Chianella, I., Lotierzo, M., Piletsky, S. A., Tothill, I. E., Chen, B., Karim, K., and Turner, A. P. (2002) Rational design of a polymer specific for microcystin-LR using a computational approach, *Anal. Chem.*, **74**, 1288-1293, <https://doi.org/10.1021/ac010840b>.

156. Karlsson, B. C., O'Mahony, J., Karlsson, J. G., Bengtsson, H., Eriksson, L. A., and Nicholls, I. A. (2009) Structure and dynamics of monomer-template complexation: an explanation for molecularly imprinted polymer recognition site heterogeneity, *J. Am. Chem. Soc.*, **131**, 13297-13304, <https://doi.org/10.1021/ja902087t>.
157. Nicholls, I. A., Andersson, H. S., Charlton, C., Henschel, H., Karlsson, B. C., Karlsson, J. G., O'Mahony, J., Rosengren, A. M., Rosengren, K. J., and Wikman, S. (2009) Theoretical and computational strategies for rational molecularly imprinted polymer design, *Biosens. Bioelectron.*, **25**, 543-552, <https://doi.org/10.1016/j.bios.2009.03.038>.
158. Kryscio, D. R., Shi, Y., Ren, P., and Peppas, N. A. (2011) Molecular docking simulations for macromolecularly imprinted polymers, *Ind. Eng. Chem. Res.*, **50**, 13877-13884, <https://doi.org/10.1021/ie201858n>.
159. Sullivan, M. V., Dennison, S. R., Archontis, G., Reddy, S. M., and Hayes, J. M. (2019) Toward rational design of selective molecularly imprinted polymers (MIPs) for proteins: computational and experimental studies of acrylamide based polymers for myoglobin, *J. Phys. Chem. B*, **123**, 5432-5443, <https://doi.org/10.1021/acs.jpch.9b03091>.
160. Boroznjak, R., Reut, J., Tretjakov, A., Lomaka, A., Opik, A., and Syrinski, V. (2017) A computational approach to study functional monomer-protein molecular interactions to optimize protein molecular imprinting, *J. Mol. Recognit.*, **30**, e2635, <https://doi.org/10.1002/jmr.2635>.
161. Hsu, C. Y., Lin, H. Y., Thomas, J. L., Wu, B. T., and Chou, T. C. (2006) Incorporation of styrene enhances recognition of ribonuclease A by molecularly imprinted polymers, *Biosens. Bioelectron.*, **22**, 355-363, <https://doi.org/10.1016/j.bios.2006.05.008>.
162. Mazouz, Z., Mokni, M., Fourati, N., Zerrouki, C., Barbault, F., Seydou, M., Kalfat, R., Yaakoubi, N., Omezine, A., Bouslema, A., and Othmane, A. (2020) Computational approach and electrochemical measurements for protein detection with MIP-based sensor, *Biosens. Bioelectron.*, **151**, 111978, <https://doi.org/10.1016/j.bios.2019.111978>.
163. Dhinesh Kumar, M., Karthikeyan, M., Kaniraja, G., Muthukumar, K., Muneeswaran, G., and Karunakaran, C. (2024) Computational modelling and optimization studies of electropentamer for molecular imprinting of DJ-1, *J. Mol. Graph. Model.*, **128**, 108715, <https://doi.org/10.1016/j.jmgm.2024.108715>.
164. Cubuk, H., Ozbil, M., and Cakir Hatir, P. (2021) Computational analysis of functional monomers used in molecular imprinting for promising COVID-19 detection, *Comput. Theor. Chem.*, **1199**, 113215, <https://doi.org/10.1016/j.comptc.2021.113215>.
165. Rajpal, S., and Mizaikoff, B. (2022) An *in silico* predictive method to select multi-monomer combinations for peptide imprinting, *J. Mater. Chem. B*, **10**, 6618-6626, <https://doi.org/10.1039/d2tb00418f>.
166. Lowdon, J. W., Ishikura, H., Kvernenes, M. K., Caldara, M., Cleij, T. J., van Grinsven, B., Eersels, K., and Diliën, H. (2021) Identifying potential machine learning algorithms for the simulation of binding affinities to molecularly imprinted polymers, *Computation*, **9**, 103, <https://doi.org/10.3390/computation9100103>.

**Publisher's Note.** Pleiades Publishing remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.