



OPEN ACCESS

EDITED BY

Yong Huang,
Northwest A&F University, China

REVIEWED BY

Shuai Lu,
Zhengzhou University, China
Simon Evan Hufton,
National Institute for Biological
Standards and Control (NIBSC),
United Kingdom

*CORRESPONDENCE

Jianxin Sui
suijianxin@ouc.edu.cn

SPECIALTY SECTION

This article was submitted to
Microbial Immunology,
a section of the journal
Frontiers in Immunology

RECEIVED 02 October 2022

ACCEPTED 07 November 2022

PUBLISHED 21 November 2022

CITATION

Liu C, Lin H, Cao L, Wang K and Sui J
(2022) Research progress on unique
paratope structure, antigen binding
modes, and systematic mutagenesis
strategies of single-domain antibodies.
Front. Immunol. 13:1059771.
doi: 10.3389/fimmu.2022.1059771

COPYRIGHT

© 2022 Liu, Lin, Cao, Wang and Sui.
This is an open-access article
distributed under the terms of the
[Creative Commons Attribution License
\(CC BY\)](https://creativecommons.org/licenses/by/4.0/). The use, distribution or
reproduction in other forums is
permitted, provided the original
author(s) and the copyright owner(s)
are credited and that the original
publication in this journal is cited, in
accordance with accepted academic
practice. No use, distribution or
reproduction is permitted which does
not comply with these terms.

Research progress on unique paratope structure, antigen binding modes, and systematic mutagenesis strategies of single-domain antibodies

Chang Liu, Hong Lin, Limin Cao, Kaiqiang Wang
and Jianxin Sui*

College of Food Science and Engineering, Ocean University of China, Qingdao, Shandong, China

Single-domain antibodies (sdAbs) showed the incredible advantages of small molecular weight, excellent affinity, specificity, and stability compared with traditional IgG antibodies, so their potential in binding hidden antigen epitopes and hazard detection in food, agricultural and veterinary fields were gradually explored. Moreover, its low immunogenicity, easy-to-carry target drugs, and penetration of the blood-brain barrier have made sdAbs remarkable achievements in medical treatment, toxin neutralization, and medical imaging. With the continuous development and maturity of modern molecular biology, protein analysis software and database with different algorithms, and next-generation sequencing technology, the unique paratope structure and different antigen binding modes of sdAbs compared with traditional IgG antibodies have aroused the broad interests of researchers with the increased related studies. However, the corresponding related summaries are lacking and needed. Different antigens, especially hapten antigens, show distinct binding modes with sdAbs. So, in this paper, the unique paratope structure of sdAbs, different antigen binding cases, and the current maturation strategy of sdAbs were classified and summarized. We hope this review lays a theoretical foundation to elucidate the antigen-binding mechanism of sdAbs and broaden the further application of sdAbs.

KEYWORDS

single-domain antibody, structure, binding modes, mutagenesis, epitope, paratope

1 Introduction

The traditional heterotetrameric structures of conserved IgG antibodies have been challenged by the natural discovery of antibodies that are only retained heavy chains and devoid of light chains in the sera of *Camelidae* called heavy-chain antibodies (HCAbs) (1) and *Chondrichthyes* called Ig New Antigen Receptors (IgNARs) (2). With the continuous maturity of molecular biologies, such as hybridoma technology (3), DNA recombinant (4), phage display technology (5), and next-generation sequencing (6), genetic engineering antibodies, represented by sdAbs, could be in-depth investigated and developed.

SdAbs are the recombinant antibodies, which are screened from a library and ultimately heterologously expressed, only retaining the variable region of HCAbs or IgNARs. In nearly 20 years of research, the advantages like high affinity and specificity (7), excellent thermostability (8), and organic reagent tolerance have been gradually put on the map (9). More importantly, after the unique prolate “rugby ball shaped” paratope structures have been discovered (10), SdAbs have significantly developed in the detection of hazard substances in food, agricultural and veterinary fields (11–13). Although it is only composed of 110–130 amino acids, it has the equivalent or higher antigen affinity to traditional antibodies, attained with affinities as low as in nanomolar range against an antigen epitope (14). Maximum recorded associations were achieved in the picomolar array in the binding case of anti-albumin (15).

There are extensive related review articles about the structural characterization, physicochemical properties and different application fields of sdAbs. Muyldermans et al. (16) and Juma et al. (17) reviewed the typical structures of HCAbs-derived and IgNAR-derived sdAbs and their corresponding heavy chain antibodies detailly. Goldman et al. (18) reviewed the strategies to improve the stability of sdAbs, which showed that the excellent performance of sdAbs enables them to have development potential in many fields. Hoey et al. (19) and Khalid et al. (20) looked forward to the potential of sdAbs in the field of disease treatment, clinical diagnosis and immune detection, respectively. Meanwhile, Leow et al. (21) reviewed the potential of sdAbs in medical imaging. Although the potential of sdAbs in various application scenarios is vast, the actual binding situation between sdAbs and certain antigens with high affinity remained unclear. Studies have shown that the binding modes of sdAbs are different when they bind to certain antigens, especially haptens. The existing binding cases between specific antigens and sdAbs are needed to be summarized. Clarifying the binding modes between sdAbs and different antigens could achieve systematic maturation and widen their application of sdAbs.

Nowadays, with the continuous development and maturity of crystallography, protein analysis software with different algorithms, and next-generation sequencing technology, the research on the structure-activity relationship and systematic

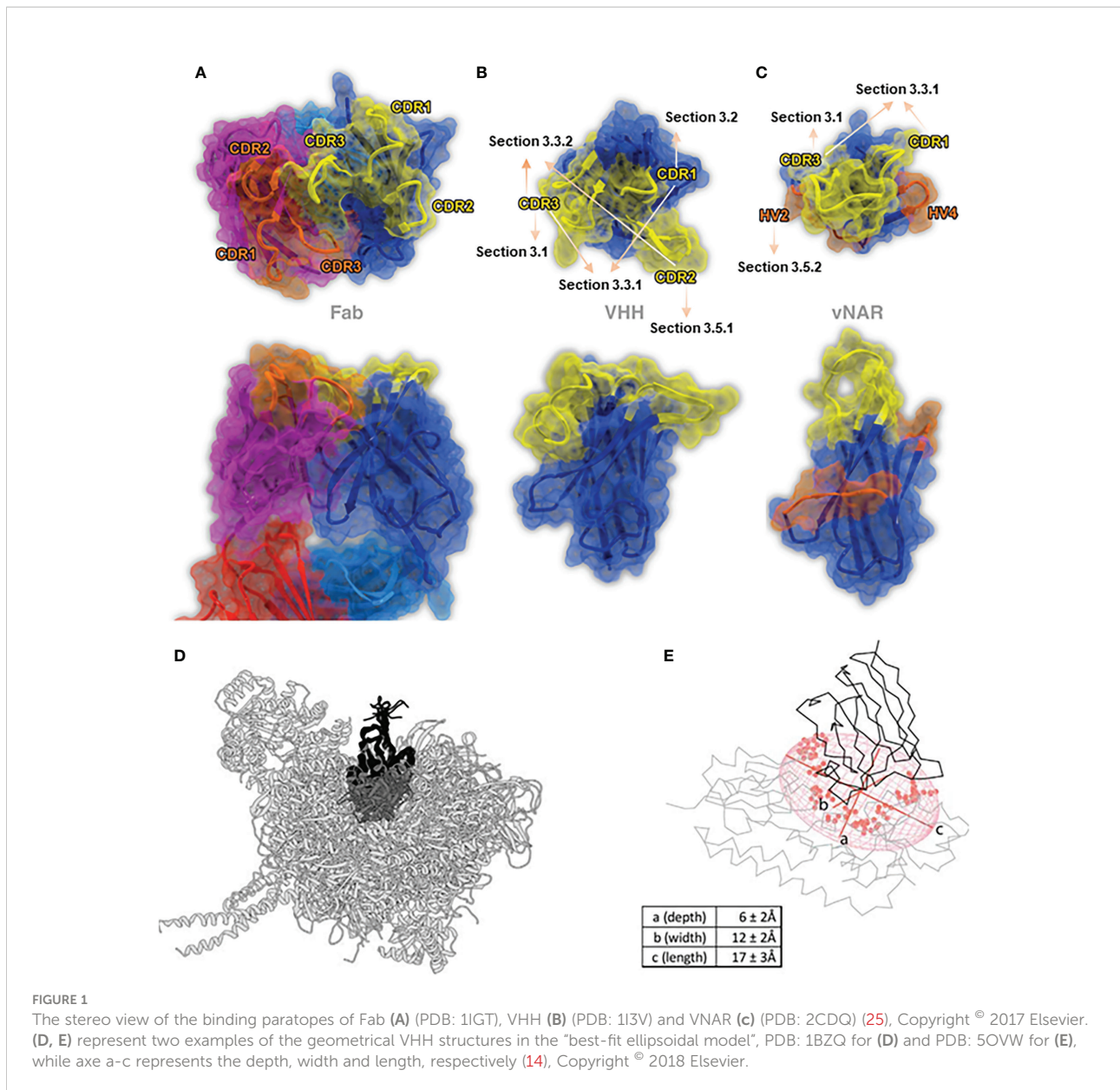
maturation of antibodies has become a focused area in sdAbs research (22). Therefore, we reviewed the existing binding modes between sdAbs to certain antigens, including macromolecule antigens and hapten antigens, and the strategies of systematic maturation of sdAbs, in hoping of providing a theoretical basis for further elucidating the antigen binding mechanism of sdAbs and broadening the application of sdAbs.

2 Structural features of sdAbs

2.1 VHH domains

At present, the basic structure of sdAbs and their corresponding encoding genes have been investigated comprehensively. SdAbs derived from camelid heavy-chain antibodies are called VHH domains. The gene encoding VHH domain length is about 360 bp, which allows expanded functionality through the creation of modularity *via* genetic fusions to a wide array of proteins, like the creation of multi-specific antibody fusions (23, 24). VHH domains comprise 9 β -strands, one 4-stranded β -sheet and another 5-stranded β -sheet, connected by a conserved disulfide bond between Cys residue at position 23 (23Cys) and Cys residue at position 94 (94Cys) to stabilize the structure, packed against a conserved Trp residue (16). An additional disulfide bond connects the CDR3 loop and CDR1 in camels or CDR2 in llamas, resulting in a more constrained conformation (25). Different from the interface region comprised of highly conserved hydrophobic residues, usually 47Val, 49Gly, 50Leu, 52Trp. In VHH domains, these residues are replaced by smaller or hydrophilic amino acids, primarily 47Phe, 49Glu, 50Arg, and 52Gly (26–29). As a result, the water solubility is improved while the tendency to form the aggregate is reduced compared to traditional IgG antibodies (25) (Figures 1A, B).

Under the existing research, VHH domains typically rely heavily on the elongated CDR3 regions in antigens bindings. There's a lot of evidence proving that the extended CDR3 region participates in intramolecular interactions with the VHH framework (19). Based on the length of CDR3, it can be generally divided into three types: concave type (about 6 aa), loop type (about 12 aa), and convex type (about 16 aa). According to the definition of IMGT, in VHH domains, the length of the CDR3 region is approximately twice than that of CDR1 and CDR2 regions, as CDR1 and CDR2 regions are usually 7 aa long, respectively (30). The convex form of CDR3 provides a sufficiently large antigen interacting surface for about 600–800 Å (31), implying a more extensive versatility and flexibility in binding target antigens (19). Thereby, it is often found to enhance the additional interaction strength and penetrate deeply into the cavity of the target antigen. Compared to monoclonal antibodies, VHH domains are also



exhibited to improve tissue penetration and significantly increase the stability (32).

2.2 VNAR domains

Compared to VHH domains, sdAbs derived from IgNARs are called VNAR domains. The features of smaller molecular weight and stronger ion tolerance of VNAR domains compared to VHH domains have attracted widespread of interests. The most distinctive feature of the VNAR domain is the deletion of C' and C'' strands that typically comprise the CDR2 region as the somatic mutations result (15), making it only consisted of 8 β -strands, becoming the most minor antigen binding domain

naturally (33, 34) (Figure 1C). A conserved disulfide bond connects two β -strands between FR1 (22Cys) and FR3 (83Cys) (35) [in some studies numbered as 21Cys & 82Cys (20)]. The absence of the CDR2 region is compensated by two loops, known as hypervariable region 2 region (HV2) and hypervariable region 4 region (HV4), with a high diversity of amino acids. Structurally speaking, the HV2 region forms a “belt-like” surrounding the VNAR domain, while the HV4 region lies at the top of the VNAR domain, opposite to the CDR1 region (33, 36).

In contrast to mammalian antibody genes that are typically organized in the translocon format, shark antibody genes are exclusively arranged in the cluster configuration (15). This cluster configuration and multiple re-arrangement events,

containing P-nucleotide addition, N-region addition, D-region, and exonuclease trimming provide diversification in sequence and length in VNAR domains, especially in the CDR3 region (15, 25, 37). The length of CDR3 could vary up to 34 amino acids, while it only generally comprises 8-12 amino acids in humans (38).

Unlike VHH domains, VNAR domains usually contain non-canonical cysteines, which could form additional disulfide bonds and dramatically alter the structure topology of variable loops of VNAR domains (39–42). Therefore, the divergence of additional disulfide bonds increased the VNAR domains' structural variability and the interaction of antigen epitopes (15). The inter-domain disulfide bonds formed by Cys residues determine the structures of the VNAR domains. To further distinguish different structures of VNAR domains due to the atypical disulfide bonds, VNAR domains are divided and classified into different subtypes (40). To date, VNAR domains are mainly classified into the following four subtypes, as shown in the following table (Table 1).

Recently, a sizeable next-generation sequencing combing Perl Script (a customized algorithm used to analyze and merge the sequences) was used together to analyze approximately 1.2 million full-length VNAR domain sequences gained from an unimmunized phage-display library constructed from six naïve nurse sharks (*Ginglymostoma cirratum*) (51). However, around 5% of VNAR domains cannot be classified in any subtypes mentioned above but also showed a remarkable binding affinity to specific antigen epitopes (41). Advancements and optimization are needed considering the limitation of this classification method and further understanding of the

biophysical properties and the binding modes of antigen epitopes.

2.3 The unique paratope of sdAbs

In conventional IgG antibodies, the variable fragment contains six hypervariable loops, also called complementary determining loops (CDRs), including three loops in light chains, called CDR-L1, L2, L3, and three loops in heavy chains, called CDR-H1, H2, H3, sustained by highly conserved β -sheet frameworks (52). In order to recognize the specific antigenic regions accurately and specifically, residues usually encoded by up to six different CDR regions, located at the interfaces of VH and VL interfaces, were defined as paratopes, and the corresponding complementary binding sites on the antigen surface were defined as epitopes (Figure 1A) (53).

In conventional IgG antibodies, the forms of paratopes are usually a cavity, groove, or flat surface, with an epitope of 600–900 Å (the size also depends on the amino acid composition, the loop size, and the difference of algorithm) (54). The entire structure of IgG antibodies is highly conserved, while the CDR regions, particularly H3 loops, differ extensively not only in terms of sequence but also in structures (52). Undoubtedly, the CDR-H3 loop of traditional IgG antibodies plays an essential role in binding and recognizing epitopes (55). However, as mentioned above, the forms of paratope in conventional antibodies are relatively flat binding surfaces. Thus, it is considered restricted and struggled in binding certain epitopes, like active sites of enzymes, parasite coat proteins, viral canyons,

TABLE 1 Subtypes of VNAR domains and their structural features.

Subtypes	Non-canonical Cys Location	Featural Structural Description	Name (Species)	Refs
Type I	located on FR2 & FR4	1. relatively rigid antigen-binding surface; 2. Cysteines in CDR3 loops form intraloop disulfide bonds; 3. CDR3 pinned tightly against the side of the molecule	Nurse shark (<i>Ginglymostoma cirratum</i>); wobbegong shark (<i>Orectolobus ornatus</i>);	(10, 15), (33, 41), (43, 44)
Type II	located on CDR1 & CDR3	1. stable interloop disulfide bonds; 2. a protrusive “finger-like” CDR3 formation with an average of 15 and 21 residues 3. CDR1 played an apparent but minor contribution compared to CDR3 loop in antigen-binding;	Nurse Shark (<i>Ginglymostoma cirratum</i>); wobbegong shark (<i>Orectolobus ornatus</i>); Spiny dogfish (<i>Squalus acanthias</i>); Smooth dogfish (<i>Mustelus Canis</i>); Horn shark (<i>Heterodontus francisci</i>); Bamboo shark (<i>Chiloscyllium plagiosum</i>)	(36, 42), (43, 44), (45, 46)
Type IIB (Type IV)	lacking non-canonical cysteine residues and disulfide bonds	only two cysteine residues hold VNAR together (21Cys and 82Cys)	wobbegong shark (<i>Orectolobus ornatus</i>); Spiny dogfish (<i>Squalus acanthias</i>); Small-spotted catshark (<i>Scyliorhinus canicula</i>)	(40, 46), (47, 48)
Type III	located on CDR1 and CDR3 with a highly conserved Trp 31 residue on CDR1	two of the three diverse regions of type III VNARs are germline joined and showed less diversity in CDR3	only found in nurse sharks during neonatal (<i>Ginglymostoma cirratum</i>) and Spiny dogfish (<i>Squalus acanthias</i>); development before maturation of the antigen of antigen-driven response	(47, 49, 50)

and recessed cryptic epitopes (56–58). On the other hand, given the large size of traditional antibodies, it is almost impossible to achieve tissue penetration, like the blood-brain barrier, or combine with the target epitope using the conventional IgG antibodies (59) (Figures 1B, C).

Markedly different from the paratopes of traditional antibodies, the prolate “rugby ball shaped” paratope structure of sdAbs forms a distinctly convex surface, increasing the contact frequency of epitopes, making it highly suitable to bind the rigid, concave, clefts, cavities, and restricted epitopes and can access the hydrophobic core of epitope enriched with aromatic residues (60). This may be why sdAbs showed equivalent or higher binding affinities as conventional Abs with other excellent antibody properties (16). Recently, a study reported a “best-fit ellipsoidal model” to geometrically simulate and quantify the spatial situation of the sdAbs paratope (Figures 1D, E). In this best-fit ellipsoidal model, the average depth of the epitope is approximately 6 Å, while width and length are measured as 12 Å and 17 Å, respectively. In this study, a total of 28 residues in paratopes formed an approximately 1500 Å² antigen contacting surface area on average (14).

The CDR3 region usually plays an essential role in antibody epitopes. In VHH domains, the CDR3 region could form more convex and unique paratopes with CDR1, CDR1&2, and CDR2&FR, respectively (61–64). The paratopes in VHH domains are enriched in aromatic residues like conventional antibodies but bear a more hydrophobic character (14). In VNAR domains, despite restrictions on the formation of disulfide bonds by non-canonical Cys residues, a wide variety of VNAR formations can still be adopted, attributed to the enormous topological latitude inherent in the CDR3 region (40). The other interesting thing is that there is a significantly increased frequency for polar and charged amino acids on the paratope of sdAbs, which is consistent with the fact that sdAbs are more water-soluble than traditional antibodies (65). Overall, the diverse and complex paratope architecture of sdAbs provides more possibilities for binding to antigen epitopes (66, 67). The latest X-ray structure study showed that sdAbs and traditional antibodies, targeting the same homologous antigens, covered similar surface areas and formed similar non-covalent interactions with the antigens. Compared with traditional antibodies, sdAbs would preferentially enter different antigen areas on proteins (68).

There have been researches on analyzing large samples of camel-derived sdAb-antigen complex structures to look for trends in camel-derived sdAb-antigen binding and the paratope of sdAbs. In Mitchell’s research (69), 156 individual sdAb-antigen complex structures were compared with corresponding traditional antibody-antigen complex structures. The study showed that the paratope of sdAbs showed more substantial diversity in amino acid residues and binding forces with antigens than traditional antibodies. In sdAb-antigen binding cases, CDR3 regions are more

advantageous than other regions in mediating antigen interaction, in nearly 1/3 binding cases, sdAbs do not contact antigen through CDR1 and CDR2 regions. In Mitchell’s another research (70), 90 individual sdAb-antigen protein crystal structures were analyzed. Results showed that although sdAbs have only three variable loops, sdAbs could compensate by increasing the length of three CDR regions, the variation level of sequences and the diversity of amino acids. Compared with traditional antibodies, sdAbs have 7% more amino acid residues in CDR3 regions, so sdAbs showed the equivalent or higher antigen affinity to traditional antibodies.

3 Antigen binding modes of sdAbs

To date, several methods have been used to investigate the binding modes of sdAbs targeting certain antigens. One way is through X-ray crystallography. After analyzing the crystal structure of the sdAbs-antigen compound, the actual binding mode and key amino acids were determined. In the early years, alanine scanning or other experimental methods were used to predict important residues or paratopes for antigen binding (61, 71, 72). While another approach is to analyze the possible binding mode and key amino acids through molecular simulation and docking analysis based on the sequence of antigen and antibody by protein analysis software with different algorithms (73–75). These methods have become a relatively mainstream and reliable way to analyze the modes of sdAbs-antigen binding. This section summarized the contributions of key amino acids in different regions and forces of sdAbs paratopes in binding to different antigens in the existing research, hoping to discover the combination rules of sdAbs with certain different antigens.

3.1 The domination of the CDR3 region in antigen binding

As clarified above, the most distinct characteristics of sdAbs are the prolate and ellipsoid paratope structures, while the most distinguishing feature of paratope is the massive topology of the CDR3 region. Undoubtedly, the CDR3 sequence is essential for binding most antigens. It has been proved that most binding to antigen epitopes is attributed to the sdAbs paratopes formed by its long and flexible CDR3 region. This section summarizes the current studies on the dominant role of the CDR3 region in antigen binding, and the relevant key amino acids are summarized in Table 2.

Lysozyme is a bacteriostatic protein that inhibits bacterial growth by hydrolyzing peptidoglycan on cell walls with a distinct concave shape epitope, which has been gradually a widely studied target protein in antibody binding research (80). In Desmyter’s research (76), the compound crystal structure of the

TABLE 2 The cases of the CDR3 region dominate in antigen binding.

Target Antigen	sdAbs type	Main Key amino acid in paratope	Main Binding Forces	Refs
Hen egg white lysozyme	VHH	98Ile, 100Ala, 100aSer, 100bTyr and 100nTyr,	–	(76)
Hen egg white lysozyme	VHH	100aSer, 100eArg	salt bridges, hydrogen bonds	(60)
Hen egg white lysozyme	VNAR	100nArg	salt bridges	(33)
Cry1 toxins	VHH	105Asp, 106Arg, 107Val, 114Arg	–	(77)
Deoxynivalenol	VHH	102Thr, 103Val, 104Pro, 105Tyr, 106Ser	–	(78)
Tumor necrosis factor TNF- α	VHH	114Trp	salt bridge and hydrogen bonds	(79)
Gingipain K protease	VNAR	86Tyr, 88Tyr, 96Phe, 97Asp, 98Glu, 99Tyr.	hydrogen bonds; polar and non-polar interactions	(45)

VHH domain and the antigen lysozyme was first discovered. In this study, the binding mode of sdAbs to antigen was firstly described at a molecular level. VHH domain enters the active site of lysozyme cryptic concave epitope through a prominent CDR3 region (~70%). The prolate VHH domain tends to bind the epitopes of antigens that are more conserved, concave, and rigid with richer aromatic residues (Figure 2A). Later in the year 2006, Genst discussed the structure of six VHH-lysozyme compounds. Six VHH domains all tend to insert into the critical active sites of lysozyme 35Glu and 52Asp via their long CDR3 loops. Residue Arg100e of the CDR3 region made a contribution in the binding that forms a salt bridge with lysozyme to stabilize the compounds, which was consistent with Desmyter's research (60) (Figure 2B). In 2004, Stanfield firstly discovered the compound crystal structure of the VNAR domain and the antigen lysozyme (33). Similar to the way the VHH domain binds to lysozyme, the VNAR domain binds through a prominent CDR3 region, especially 100Arg and 101Tyr, which are deeply embedded in the active center of lysozyme. The residue 100Arg in the CDR3 region forms a salt bridge with 52Asp of lysozyme and partially inhibits lysozyme activity. In addition, it is also found that the conformation of the CDR1 region of the VNAR domain is not similar to that of human or murine but instead converges found in VHH domains, and part of the CDR1 region is also involved in binding with the antigen (Figure 2C).

Besides, Jiao screened the VHH clone A8 specifically binding various Cry1 toxins of *Bacillus thuringiensis* (Bt) from the library and confirmed that the key amino acids in binding Cry1 toxins were 105Asp, 106Arg, 107Val, and 114Arg in the CDR3 region by homology modeling and molecular docking techniques (77) (Figure 2D). Qiu confirmed that the amino acid residues 104Pro, 105Tyr, and 106Ser in the CDR3 region of VHH clone N-28 are essential for binding the target antigen deoxynivalenol (78) (Figure 2E). Nie screened a sdAbs called NT-3 that can effectively inhibit the tumor necrosis factor TNF- α , which could induce autoimmune diseases and inflammation, with an IC₅₀ of 0.804 μ M (79). Molecular docking showed that the CDR3 region in this clone dominated by 114Trp played an essential role in affinity function, while other amino acids 37Arg, 39Phe,

40Ser, 66Gly, 67Ser, and 75Lys located on CDR1 and CDR2 also participate in the TNF- α binding reaction through salt bridge and hydrogen bonds (81) (Figure 2F). It is found that although the CDR3 region plays a leading role in antigen binding, but it is unlikely to bind the antigen alone.

The VNAR domain 12A-9 was initially screened from a combinatorial library derived from wobbegong sharks (*Orectolobus maculatus*), with the specific binding ability for the Gingipain K protease from *Porphyromaonas gingivalis* (45). In the modeling investigation of type III 12A-9, a large number of aromatic amino acids were discovered in the CDR3 region, including 86Tyr, 88Tyr, 96Phe, and 99Tyr. The amino acid 96Phe can form multiple antigen-binding conformations with different orientations with residues with larger side chains or charged residues, forming complex and diverse paratopes. When the epitopes approached the conserved acidic amino acids 97Asp and 98Glu, this region formed a large acid area around 96Phe. Residues 86Tyr, 91Ala, 92Glu, 93Leu, 94 Asp, and 95Ser were also involved in antigen binding.

3.2 The contribution of CDR1 region in antigen binding

To date, the contribution of CDR1 region in antigen binding is quite different for macromolecular antigens and small molecular antigens. For some macromolecular antigens, directed mutations and variations on CDR1 region usually seems to have less effect on increasing the affinity in antigen binding, indicating CDR1 region may be less involved in these macromolecular antigen binding cases. In Dooley's research, the KD values of the mutant (Ala30Val) changed from 9.5 nM to 10 nM, which did not improve significantly (82). Similarly, there was no significant increasing affinity of VNAR mutant Ala27Thr on the binding of *Plasmodium falciparum* AMA1, the Kd value did not change significantly (from 10⁻⁸ M to 1.47 \times 10⁻⁷ M) (83). Also, the mutation of Phe29Leu did lead to an increase in binding affinity to *Plasmodium falciparum* AMA1 (approximately 7-fold enhanced compared to the wild type), but a large proportion of incorrectly folded proteins were produced, resulting in a large loss of protein stability (84).

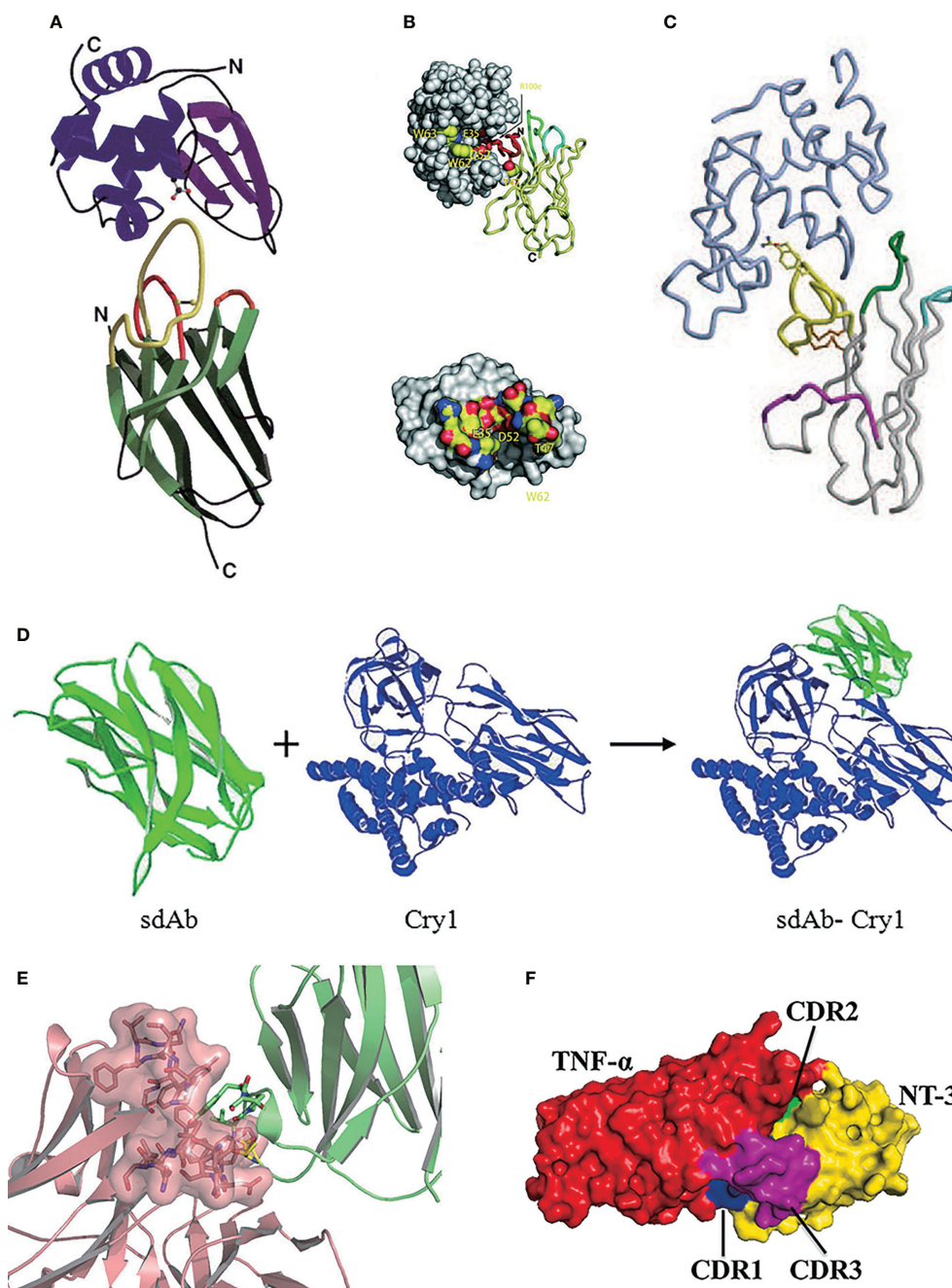


FIGURE 2

(A) Stereo view of the ribbon representation in the X-ray structure of VHH in the complex with lysozyme (76), Copyright © 1996 Nature. (B) One of six VHH-HEWL complex structures, where the active site residues are labeled and side-chain atoms color-coded (60), Copyright © 2006 The National Academy of Science of the USA. (C) The stereo view of the crystal structure of VNAR in the complex with HEL, where HEL is shown in light blue (33), Copyright © 2004 Science. (D) The stereo view of the structure modeling and molecular docking of sdAb-Cry1 toxin complex (77), Copyright © 2017 Springer. (E) The stereo view of molecular docking of sdAb-anti-DON scFv complex (78), Copyright © 2015 Springer. (F) The stereo view of molecular docking of sdAb-TNF- α complex (79), Copyright © 2021 Elsevier.

However, in some hapten antigen binding cases, sdAbs showed a special binding mode: the CDR3 region is rarely involved in the interactions with haptens, and the binding mainly contributes to the CDR1 region. In this binding

situation, the hapten antigen tends to insert its hydrophobic core into the “tunnel” structure formed by the CDR1 region (85) (Figure 3A). In Rosa’s research, three VHH domains, called T4, T9 and T10, which binds to triclocarban (TCC) specifically were

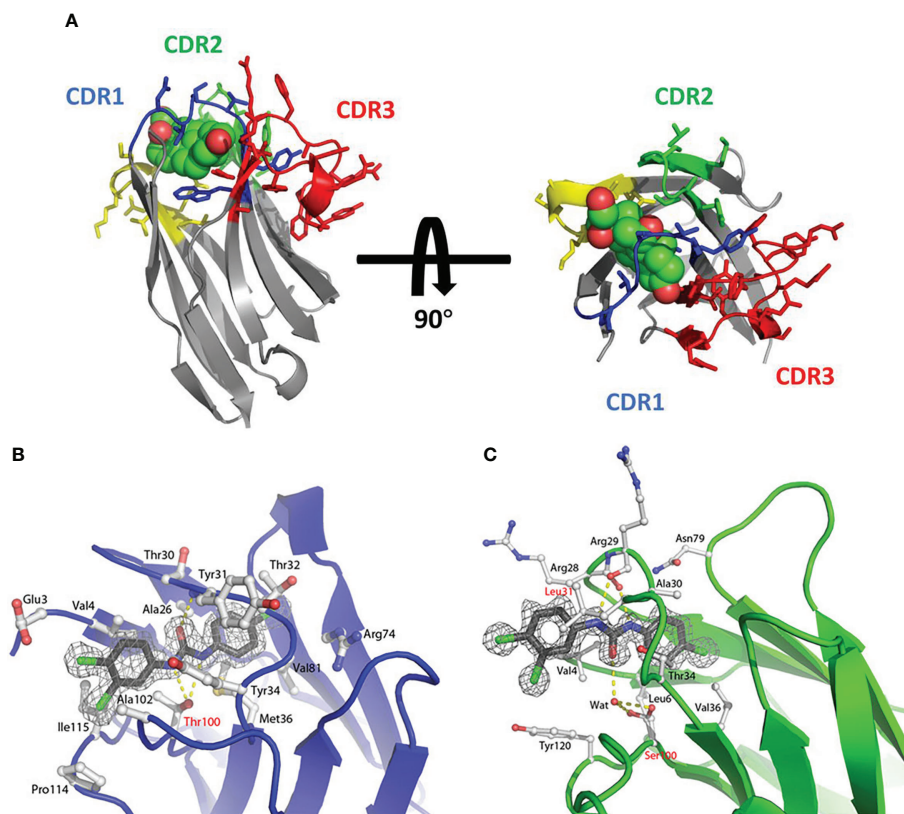


FIGURE 3

(A) The side view and top view of the cortisol in the complex with sdAb called NbCor, where the side chains of CDRs 1-3 in NbCor are shown in blue, green and red, respectively. The cortisol molecules are represented by spheres (85), Copyright © 2019 John Wiley & Sons. (B) The stereo view of the structure of TCC (gray carbons) in the complex with sdAb called T9 (blue main chain and white carbons) (86), Copyright © 2018 John Wiley & Sons. (C) The stereo view of the structure of TCC (gray carbons) in the complex with sdAb called T10 (green main chain and white carbons) (86), Copyright © 2018 John Wiley & Sons.

investigated (86). Due to the hydrophobicity of trilocarban, the interaction between TCC and sdAbs was a hydrophobic force. In this binding case, the binding mainly occurs in the large and hydrophobic “tunnel” structure formed by the CDR1 region. The more extensive binding interface leads to tight binding with the dissociation affinity of nanomolar. Similarly, in Ding’s research, the CDR1 region of the VHH domain also played an essential role in interacting with the target antigen cortisol. The hydrophobic part of cortisol inserts into the hydrophobic pocket formed by the CDR1 region to achieve the binding with sdAb (85) (Figures 3B, C).

3.3 Numerous CDR regions contribute to the antigen binding

3.3.1 CDR1 & CDR3

In the research of Decanniere, a VHH domain called cAb-RN05 was found that it had the specific binding ability to bovine

ribonuclease A (RNase A) (87). In this binding case, the CDR2 loop is not near the epitope and does not participate in antigen binding. Residues 27Tyr, 31Tyr, 32Ile, 33Tyr, 95Gly, 96Gly, 97Tyr, 100Arg, Thr100c, and 101Gly as the key amino acids to form the paratope of sdAbs. The main interaction force of this binding is the hydrogen bond formed by 27Tyr of paratope and the amide atom in 62Asn of RNase A. However, in conventional antibodies, the sizeable hydrophobic side chain of Tyr27 packs deeply into the internal hydrophobic pocket, so it is therefore unavailable to interact with the target antigen (Figure 4A). In Similar research from Koide (61), a VHH domain was screened that it has a specific binding to RNase A. In this binding case, after ALA scanning and analysis, CDR1 and CDR3 contribute nearly equal to the antigen binding. The key amino acids for antigen binding were identified as 27Tyr, 31Tyr, and 32Ile from the CDR1 region and 95Gly, 96Gly, 99Leu, and 100dTyr in the CDR3 region (Figure 4B).

The binding of sdAbs with an azo-dye called RR6 is the earliest study on the binding of sdAbs to the hapten (67). RR6 is

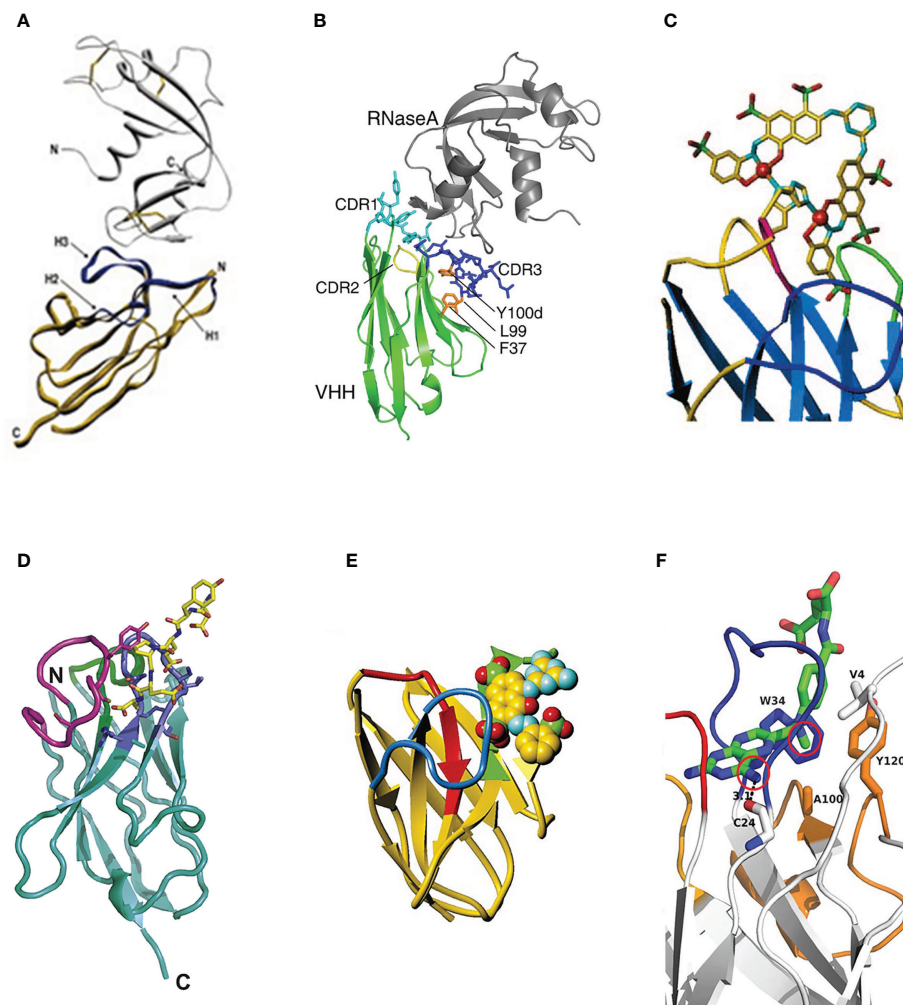


FIGURE 4

(A) The stereo view of the sdAb called cAb-RN05 (yellow) in the complex with the binding antigen RNase A (grey) (87), Copyright © 1999 Elsevier. (B) The stereo view of the crystal structure of sdAb (green) in the complex with RNase A (gray) (61), Copyright © 2007 Elsevier. (C) The stereo view of the haptin RR6 molecule in the complex with the sdAb fragment (67), Copyright © 2000 American Chemical Society. (D) The stereo view of the crystal structure of sdAb called NbSyn2 in the complex with a synthetic peptide called NGYQDYPEA-C (88), Copyright © 2010 Elsevier. (E) The stereo view of the haptin RR1 molecule in the complex with the sdAb fragment (66), Copyright © 2001 Academic Press. (F) The stereo view of haptin MTX molecule in the complex with VHH (89), Copyright © 2011 The Protein Society.

an unusual haptin with two copper ions, several aromatic rings, and numerous charged groups. Despite lacking the adjacent VL domain, the VHH domain has an excellent binding ability to the haptin. On the one hand, the CDR3 region can form a sizeable binding pocket to accommodate the haptin epitope through its massive topology, which has good surface complementarity with the haptin. On the other hand, residues 32aHis and 32cHis can bind to the copper ion in the haptin, resulting in the dissociation of sdAbs to the haptin RR6 is only 20 nM (67) (Figure 4C).

In this research (44), two VNAR domains that target apical membrane antigen 1 of *plasmodium falciparum* malarial parasites, called 12Y-1 and 12Y-2, were investigated. The antigen binding mainly contributed to the C-H... π

interactions by 29Phe from the CDR1 region and 87Tyr, 100Phe from the CDR3 region. The side chains and adjacent water molecules formed by charged and polar amino acid residues of 37Tyr, 46Glu, 82Lys, 84Gln, 101Arg, and 104Lys formed a charged pocket with hydrogen bonds, which are used for antigen binding. This research also proved that the extended β -hairpin structure in the VNAR domain could form unique paratopes to penetrate the hidden epitopes of the antigen.

3.3.2 CDR2 & CDR3

Numerous studies have shown that α -synuclein plays an essential role in the occurrence and development of Parkinson's disease, Alzheimer's disease, dementia with Lewy bodies, multiple

system atrophy, and other related neurological diseases (90). In Genst's research (88), a VHH domain selected by phage display technology, called NbSyn2, showed excellent binding affinity to α -synuclein with nanomolar affinity by NMR spectroscopy and X-ray crystallography. Detailed analysis of VHH- α -synuclein crystal structure indicated that the paratope formed by 50Arg, 52Asn, 58Lys of CDR2 region and 105Tyr, 107Gly, 113Phe, 116Trp of CDR3 region makes the contacts with the residues 136Tyr, 137Glu, 138Pro, 139Glu and 140Ala of α -synuclein. This binding case is mediated primarily by side chain interactions, which are essentially electrostatic (Figure 4D). In Spinelli's research (66), a VHH domain called VHH-52 was selected and confirmed the binding ability to the hapten azo-dye RR1. RR1 is an aromatic molecule containing three sulfate groups. In this binding case, the interaction of VHH-SO₃²⁻ dominates the binding pattern. The CDR2 and CDR3 regions form an approximately 8 Å deep crevice, acting as the antigen's paratope and leading to a strong interaction with two SO₃²⁻ groups of the hapten (Figure 4E).

3.4 Non-CDR or CDR4 of sdAbs in antigen binding

A growing number of binding cases indicated that the amino acid residues in the non-CDR region could also act as paratopes to participate in and mediate the binding to antigens (14). In a previous statistical study of VHH domains, nearly 16% are established by the residues from non-CDR interactions (91). In the research of Desmyter (92), three VHH domains called AMB7, AMD9, and AMD10 were identified to have the binding ability to pancreatic α -amylase. In this binding case, it is found that nearly 25-40% of framework residues participated in antigen recognition. A more plausible explanation is that the amino acid residues of the non-CDR region can compensate for the deficiencies of light chains of sdAbs, and this compensation mechanism can provide an equivalent or larger antigen contact surface than that of classical immunoglobulins. In Fanning's research (89), the loop structure composed of amino acid residues at positions 74-82 in the non-CDR region was defined as the CDR4 region. The residues of 74Arg, 79Asn, and 80Thr in the CDR4 region can form a hydrophobic pocket with the residues 4Val, 34Trp, 36Met, 100Ala, and 120Tyr and directly interact with the hydrophobic part of MTX. In this binding case, the CDR4 region plays an essential role in the affinity and specificity of binding to the hapten. Future research is needed to illustrate further the role of binding to antigens in this region (Figure 4F).

3.5 Other binding cases

3.5.1 CDR2 region in antigen binding cases

Xi used the phage display technology and obtained the VHH domains called AS33595 and AS32611 specific to the essential

and effective tumor target epidermal growth factor receptor (EGFR) (93). Homology modeling, molecular docking, COCOMAPS web application server, and other technical analysis methods were used to predict and confirm the paratope site. Results showed that the paratope of clone AS32611 was mainly on the surface of the CDR2 region. Residues 52Asn, 56Trp, and 58Asn act as H bond acceptors and primarily interact with 434Asp, and 436Asp of EGFR. Near the CDR2 region, residue 64Glu interacts with the 463Lys by forming a salt bridge (Figure 5A).

3.5.2 HV2 of VNAR domains as the paratope

The HV2 region is the unique structure of IgNAR-derived sdAbs. For the type I VNAR domain, the CDR3 region is held tightly in the direction of HV2. Consistent with this structure feature, the mutations in this particular type are favored, indicating that the HV2 region may act as the paratope in antigen binding (65, 95). Zielonka screened and engineered the HV2 region in the VNAR domain targeting epithelial cell adhesion molecule (EpcAM) by using the method of yeast surface display (96), which can bind to the target antigen independently of the traditional antigen binding sites composed of CDR1, CDR3 or HV4. Results indicated that the HV2 region could be used as a potential autonomous paratope that exclusively facilitates the target antigen binding without destroying the functional integrity or structure of VNAR scaffolds. This study provides a theoretical basis for developing bispecific sdAbs (97).

3.5.3 Dimer-VHH binding

In Lesne's research (94), the VHH domain combined with the target hapten caffeine with an unexpected 2:1 ratio. The caffeine molecule is sandwiched between two CDR3 regions from their VHH domains, stacking on an extremely dimer interface. Residues 34Tyr from the CDR1 region and 104Tyr from the CDR3 region are in direct contact with caffeine by hydrogen bond and π - π stacking, respectively (Figure 5B).

4 Systematic maturation of sdAbs

The study of the structure-activity relationship ultimately improves antibodies' systemic maturation, including the affinity, stability, solubility, specificity, and other properties of sdAbs. After defining and confirming the key amino acids, which contribute primarily to antigen binding, and the binding modes to the antigens, different strategies, such as the combination of mutagenesis and screening procedure, can be adopted at the molecular level to systematically mature the sdAbs (98). At present, strategies to improve the systematic maturation of antibodies *in vitro* are mainly divided into three strategies: random mutagenesis, target mutagenesis, and *in silico* mutagenesis. This section summarizes and classifies the existing

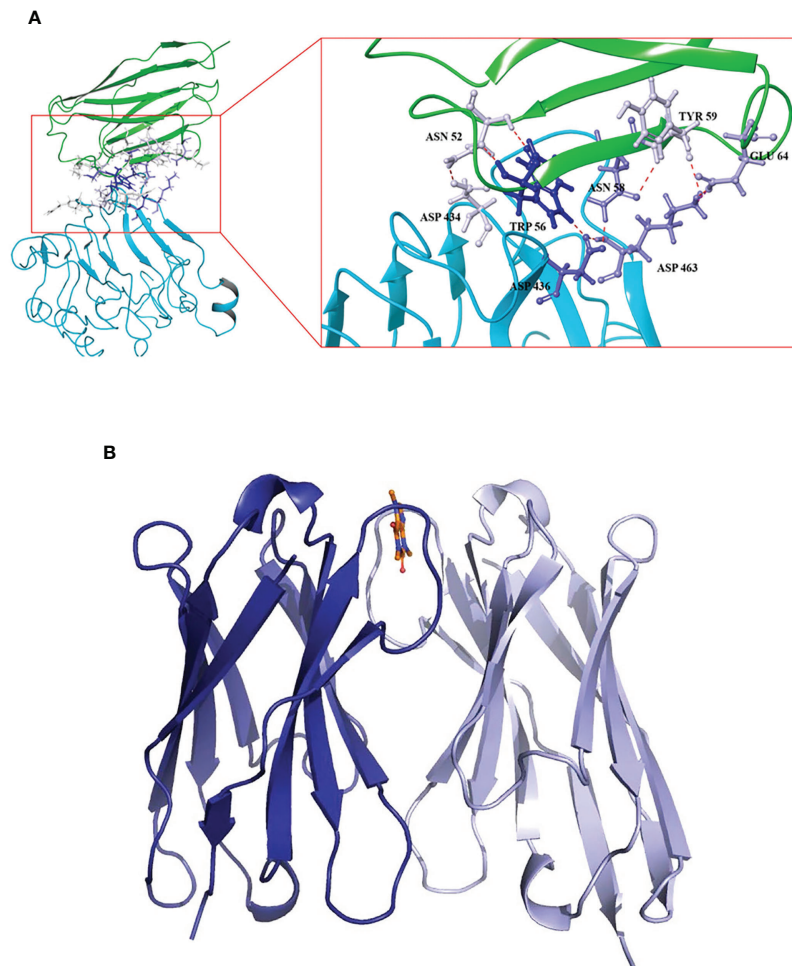


FIGURE 5

(A) The stereo view of the interaction between the binding sites of VHH called AS32611 and the domain III of EGFR (93), Copyright © 2020 Elsevier.
 (B) The stereo view of VHH dimers in the complex with caffeine, where the caffeine is represented as sticks (94), Copyright © 2019 Springer.

systematic maturation researches on sdAbs and compares the advantages and disadvantages of these methods (Table 3).

Several outstanding challenges in the systematic maturation of sdAbs still need to be addressed. Firstly, saturation mutation can be used to evaluate every possible amino acid. However, it is not likely to result in significant gains in a certain ability by single mutations (112). Finding the actual binding sites of multiple CDRs mapping to the precise paratopes is not a trivial task. Secondly, due to the limited capacity of the library diversity, it is not realistic to test all combinations of single or multiple mutations in a single library (113). Lastly, for the systematic maturation of antibodies, the enhancement of one certain property, such as affinity, stability or specificity, often comes at the cost of weakening other properties. For example, an affinity increase often leads to a decrease in specificity (114). Notably, residues such as Arg and other aromatic amino acids

tend to be enriched at antibody affinity maturation but also increase the risk of non-specific antibody interactions (115, 116).

4.1 Random mutagenesis

In random mutagenesis, a certain length of the sequence of antibodies would be mutated randomly (117). This mutagenesis method will introduce the mutation into the corresponding region by employing error-prone PCR (99, 100), DNA shuffling (101, 102), etc. Mutagenesis in this way can obtain various sequences from native antibodies, which is an incredible and powerful tool, but obviously, this mutagenesis method lacks pertinence and low efficiency.

In Kobayashi's research (99), the error-prone PCR was used to introduce random mutants into the sdAbs genes fragment

TABLE 3 The advantages and disadvantages of main mutagenesis strategies.

Mutagenesis strategies	Different Forms	Advantages	Disadvantages	Refs
Random Mutagenesis	error-prone PCR, saturation mutagenesis, DNA shuffling	Huge library with various mutants	High proportion of non-relevant mutations, low efficiency	(84, 99, 100, 101, 102)
Target Mutagenesis	Ala scanning, site-directed mutation	More targeted, High efficiency	limited mutants, Point mutations usually have no effect.	(77, 78, 103, 104, 105, 106, 107, 108)
<i>In silico</i> mutagenesis	Protein analysis software combining different algorithms	Virtual huge library mutants; better structure and interaction analyze	Restrict by different algorithms; Lack of maturity and credibility	(73, 79, 109, 110, 111)

target to Anti-E2 (a kind of cortisol with therapeutic significance). A mutant library with a capacity of 10^5 members was constructed. After biopanning, the binding ratio of mutants (23.8) was significantly higher than that of wild types (11.6) at the same antibody titer (250 ng/well). In Yau's research (100), the random hotspot mutation was used to construct the mutant ribosome display library targeted to the parathyroid hormone. After the measurement of SPR, the affinity of the mutant could be increased about 30 times.

In Sheedy's research (101), DNA shuffling by staggered extension process (StEP) method was used to create more recombinant sdAbs parental genes, targeted to auxinic herbicides. After CDR shuffling of the best sdAb with other four sdAb clones by staggered extension process and pentamerization of shuffled clones, it was found that the affinity of shuffled sdAbs was similar to that of the parental clone, but their affinity for auxinic herbicides decreased. In Harmson's research (102), DNA shuffling and treatment of resultant library with gastric and jejunal fluid were used before biopanning. After biopanning, the most stable clone, called K922, was selected with more stability and higher affinity to the target antigen *E. coli* F4 fimbriae (a toxin that causes diarrhea in newborns).

A new *in vitro* display technique, called ribosome display, was established by Pluckthun in 1997 (118), which mainly involved first constructing a DNA template for ribosome display, followed by *in vitro* transcription, translation and affinity screening. The transcribed product mRNA does not contain any stop code. During the translation, ribose will stay at the 3' end of mRNA, so the target gene's translation product can be displayed on the ribosome's surface, forming a ternary complex of "mRNA-protein-ribosome." Elution buffer is used to dissociate the ribosome and release mRNA. After purification, the DNA was used as a template for RT-PCR, and various essential elements of ribosome display were reintroduced for the next cycle of enrichment and selection. The target DNA with high affinity was finally screened out. At present, ribosome display technology has been applied to the optimization of the affinity of shark-derived sdAbs (119). A type of RNA polymerase from Q β bacteriophage is one of the most commonly used to

generate the diversity and transfer the mutated mRNA templates to ribosomes for translation (120). In Kopsidas' research (84), this method was used to create a diverse library of VNAR domains. Coupling these randomly mutated mRNA templates directly to the translating ribosome allowed *in vitro* selection of affinity matured variants showing enhanced binding to the target AMA1 from *Plasmodium falciparum*. Two affinity-matured variants were isolated carrying the mutations: Pro90Leu mutant confers a powerful (10-fold) enhancement of antigen-binding capability compared to the wild type (83).

4.2 Target mutagenesis

In target mutagenesis, one or more selected residues in a certain region (usually the CDR region) would be mutated. Target mutagenesis includes Ala scanning, site-directed mutation, including site-saturation mutagenesis to make a relatively small library and select the specific mutants through related display techniques (121).

In Jiao's research (77), the sdAbs called A8 targeted Cry toxins of *Bacillus thuringiensis* by screening a humanized sdAbs library. Site-saturation of key amino acids 105Asp, 106Arg, 107Val, and 114Arg was carried out by overlap extension PCR. Two mutants capable of detecting at least six kinds of Cry1 toxins, called 2-C_1 & 2-C_9, were screened, which broadened the detection range of target antigens. In Qiu's research (78), the sdAb called N-28 targeted to the deoxynivalenol (DON) was screened from a naïve sdAbs library and indicated the residues 102Thr, 103Val, 104Pro, 105Tyr, and 106Ser are essential in antigen binding. After saturation mutagenesis of the above five residues, mutants with improved sensitivity were selected by monoclonal phage ELISA and sequencing. Results showed that three mutants with position 102 mutating to Tyr, position 103 mutating to Leu, and position 105 mutating to Phe showed the half-maximal inhibitory concentrations (IC₅₀) were 24.49 ± 1.0 ng/mL (3.2-fold), 51.83 ± 2.5 ng/mL (1.5-fold), and 35.65 ± 1.6 (2.2-fold) ng/mL, respectively. In Wang's research (103), the sdAb called Nb28, which was against the mycotoxin ochratoxin A (OTA)

from an alpaca immune library was selected. After homology, molecular docking, and Ala scanning verification, 53Gly, 79Met, 102Ser, and 149Leu were confirmed as key amino acids for antigen binding. In this study, a two-site saturated mutation library was used to construct a mutation library to determine the best mutation combination. After the biopanning and identification procedure, the mutant with position 53 mutants to Gln and position 102 mutants to Asp can reduce the IC₅₀ to 0.29 ng/mL (1.4-fold) and K_D value to 52 nM (1.36-fold), respectively. In Tiller's research, the sdAb called NbSyn2 specific to the C-terminus of α -synuclein from an immune library was obtained (104). After computational and alanine scanning, amount of 14 permissive sites, including 49Ala, 52bLeu, 53Gly, 55Val from CDR2 and 94Ala, 95Lys, 96Phe, 97Ser, 99Gly, 100bGly, 100cTyr, 100dSer, 100fSer, 100gAsn from CDR3. In this research, the mutant with position 52b mutants to Typ, position 53 mutants to Arg, position 96 mutants to Ser, and 100 mutants to Thr, called N2.12, achieved more than 7-fold affinity enhancement without compromising stability. The mutation of position 96 mutants to Ser contributes positively to affinity and stability, while position 53 mutants to Arg increases the affinity with the cost of stability.

Numerous studies have shown that targeted mutagenesis to Cys and the addition of an atypical disulfide bond can effectively improve the TM values (the temperature at which the antibody cannot maintain its original tertiary structures and result in denaturation of the antibody) of sdAbs. In Turner's research (105), a VHH specifically binding ricin was screened. Target mutagenesis of Arg at position 118 to Trp increased TM by 6 °C without decreasing the affinity of sdAbs. In Liu's research (106), a VHH was screened with the specific binding of recombinant Ebola virus GP protein with nM affinity using phage display technology. By target mutagenesis, Residues 54Ala and 78Ile were both mutated to Cys. After adding the disulfide bonds, the TM of the mutant can be increased by 15-17°C. In the research of Hagihara (107), the highly conserved residues 54Ala and 78Ile were mutated to Cys by adding atypical disulfide bonds, which could form a more stable tertiary structure of sdAbs, and the TM value of the mutant increased by about 10°C compared with the wild-type. In the research of Anderson (108), the FR1 region was reversed to hallmark amino acids by point mutations, which improved the TM by 2-6°C. Results showed that adding atypical disulfide bonds to the mutant could increase the TM by 9-13°C with nearly 100% of initial binding activity remaining.

4.3 *In silico* mutagenesis

The evolution of DNA sequencing techniques and 3D structural models using computational approaches have made remarkable achievements (122, 123). The method of *in silico* using computational prediction of antibody 3D structure for redesigning antibodies has gradually become a new method to

improve the maturity of antibodies (124, 125). These approaches are based on statistical models of different exhaustive algorithms, including Monte Carlo and Dead End Elimination (DEE), then quality filtration based on energy assessments such as solvent treatment and electrostatic interactions (126). In parallel, several mature protein analysis software could combine small fragments from different proteins and optimize tens of thousands of 3D protein models, which laid the foundation for model construction of sdAbs, such as Rosetta and I-Tasser (127-129). Very recently, machine learning and deep learning approaches, such as AlphaFold2 (130) and trRosetta (131), make antibody prediction and redesignation to a new level. These approaches are composed of multiple complex neural networks, which could combine very long-distance evolutionary searches and advanced local compositional proposals (75). These advances are due to the improvement of GPU computing power and better representations of mathematics in the past few years.

Using the database provided by online websites to predict and redesign sdAbs has become a new approach *in silico* mutagenesis. In Wilton's research (109), a database called sdAbs-DB (<http://www.sdab-db.ca/>) was constructed to provide free sdAbs and related proteins, which were summarized from NCBI and PDB databases, published articles and user-submitted contents. The sdAbs-DB is able to predict the protein structure and perform corresponding bioinformatics analysis. Similarly, the SABdab, including the SABdab-nano tracker database, was reported by Schneider in 2022 (<http://opig.stats.ox.ac.uk/webapps/newsabdab>) (110). This database is used to track the related research of sdAbs with weekly updates, providing more physicochemical properties of sdAbs.

The most crucial advantage of *in silico* maturation is that a virtual library with a capacity of 10⁴⁰ members could be constructed and better analyze the structure and interaction of antigen and antibody (121). Even if the crystal structure of the molecule has not been determined, the 3D structures of antigens and antibodies can be simulated using a considerable number of modeling and simulation software. While computational approaches are not generally considered a substitute for experimental verification, they can help generate testable assumptions through different algorithms.

In Sefid's research (73), a specific VHH against Bap antigen in *Acinetobacter baumannii* was selected by phage display. Later, structural prediction and docking of the Bap-VHH complex were used for designing and validation with a higher affinity of VHH. According to the VHH interfaces prediction and scores are given by model evaluation software, it is inferred that mutant 6 (Ile37Glu, Pro38Ser, Tyr43Asn, Ala82Arg, Asn84Asp, Phe89Ile, Tyr99Ser) and mutant 9 (Ala82Thr, Asn84Gln, Phe89Arg, Tyr99His) could significantly improve the binding ability to the antigen. In Nie's research, a novel humanized scaffold library was constructed by introducing degenerate primers (NNK/NNN) in 13-15AA to improve the diversity of the library (79). Finally, a VHH called NT-3 with an IC₅₀ of 0.804 μ M was screened, effectively inhibiting

the target antigen tumor necrosis factor TNF- α . In a recent study (111), the CDR3 region of sdAb was redesigned *in silico*, called DesAb-HAS-D3 and DesAb0. Enhancing-sampling molecular dynamics simulations were used to compare their free energy distribution. The results showed that although there are more sequences in the CDR3 region of DesAb-HAS-D3 and could theoretically generate more conformations, while its actual binding to antigen still shows strong structural complementarity. The design of DesAb0 reduces the rigidity of the CDR3 region and does not positively affect antigen binding.

4.4 Chemical mutagenesis

In Lindstedt's research (98), two rounds of chemical mutagenesis of sdAbs, called DesAb-Ab (3-9), by post-translationally installed synthetically versatile non-canonical amino acid dehydroalanine (Dha), to further inhibit the accumulation of Ab42 protein, which is closely related to Alzheimer's disease. In all residues of CDR3, 137Glu, 138Thr, and 139Leu are suitable for mutagenesis. The structural integrity after chemical mutagenesis was verified by LC-MS and circular dichroism. The results showed that five orders of magnitude could increase the inhibition rate of Ab42 protein by 138Thr-DHA without affecting its stability.

5 Conclusion and future perspective

Nowadays, in addition to the binding analysis based on the crystal structure of antigen-sdAbs, the continuous upgrading and optimization of next-generation sequencing, protein analysis software with different algorithms provide a platform for the research of binding of sdAbs and different antigens. A relatively accurate way of sdAbs binding to the antigens can also be obtained by protein modeling, energy optimization and molecular docking. In the future, the research on sdAbs based on protein analysis software with more accurate algorithms will become the mainstream, providing more accurate results for revealing the actual binding of sdAbs and different antigens.

In this paper, various binding modes between sdAbs and antigen molecules were reviewed. In addition to the traditional cognition, the CDR3 region acts as the main region of antigen binding, the different convex structure of CDR3 causes the different specificity and affinity in antigen binding. However, more studies have also shown that other regions, including the CDR1 region, the joint influence of multiple CDR regions, the framework regions, and the unique HV2 region of HCabs-derived sdAbs, can be used as the main binding regions and make a major contribution to the binding of antigen epitopes. The study of sdAbs' other regions in binding antigen epitopes is

beneficial to the subsequent target mutagenesis of sdAbs, designing and developing the bispecific of multi-specific sdAbs.

Studies have shown that the binding cases of sdAbs to haptens differ from that of macromolecular antigens. Haptens are more inclined to bind to the tunnel structure formed by the CDR1 region of sdAbs. At present, there are still few researches about the binding of sdAbs to haptens, which are limited to HCabs-derived sdAbs. It is of great significance to further clarify the binding mode of this region with haptens. On the one hand, the sdAbs could be targeted mutagenesis, to further improve the affinity and specificity of sdAbs binding to the haptens. On the other hand, it has potential research value for developing IgNAR-derived sdAbs with stronger tolerance of organic reagents and more suitable for detecting liposoluble haptens, such as organic pesticides.

The study of the binding modes and the related structure-activity relationship serves for the systematic maturity of sdAbs. The related antibody properties of sdAbs are restricted by many factors. Sometimes it is difficult to obtain the sdAbs that meet all requirements through traditional biopanning and heterologous expression, so the subsequent systematic maturation of sdAbs is particularly important. Nowadays, there are still unavoidable problems in different antibody systematic maturation methods. It is expected that protein analysis and docking software, relying on more accurate algorithms, could solve the issues of low diversity of mutants, low efficiency and false docking results. It will be an effective way to analyze the structure-activity relationship and systematic maturity of sdAbs in the future.

Author contributions

CL wrote the review, made the figures and tables under the supervision and revision of HL, LC, KW and JS. All authors contributed to the article and approved the submitted version.

Funding

This work was financially supported by the National Natural Science Foundation of China (No. 32072308), the National Key R&D Program of China (2018YFD0901005) and the Fundamental Research Funds for the Central Universities (No.202042011).

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated

organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

References

- Hamers-Casterman C, Atarhouch T, Muyldermans S, Robinson G, Hammers C, Songa E, et al. Naturally occurring antibodies devoid of light chains. *Nature* (1993) 363:446–8. doi: 10.1038/363446a0
- Greenberg A, Avila D, Hughes M, Hughes A, McKinney E, Flajnik M. A new antigen receptor gene family that undergoes rearrangement and extensive somatic diversification in sharks. *Nature* (1995) 374:168–73. doi: 10.1038/374168a0
- Kohler G, Milstein C. Continuous cultures of fused cells secreting antibody of predefined specificity. *Nature* (1975) 256:495–7. doi: 10.1038/256495a0
- Skerra A, Pluckthun A. Assembly of a functional immunoglobulin fv fragment in *Escherichia coli*. *Science* (1988) 240:1038–41. doi: 10.1126/science.3285470
- Ghahroudi M, Desmyter A, Wyns L, Hamers R, Muyldermans S. Selection and identification of single domain antibody fragments from camel heavy-chain antibodies. *FEBS Lett* (1997) 414:521–6. doi: 10.1016/S0014-5793(97)01062-4
- Mishra A, Mariuzza R. Insights into the structural basis of antibody affinity maturation from next-generation sequencing. *Front Immunol* (2018) 9:117. doi: 10.3389/fimmu.2018.00117
- Gaylord S, Dinh T, Goldman E, Anderson G, Ngan K, Walt D. Ultrasensitive detection of ricin toxin in multiple sample matrixes using single-domain antibodies. *Analyt Chem* (2015) 87:6570–7. doi: 10.1021/acs.analchem.5b00322
- Liu X, Tang Z, Duan Z, He Z, Shu M, Wang X, et al. Nanobody-based enzyme immunoassay for ochratoxin A in cereal with high resistance to matrix interference. *Talanta* (2017) 164:154–8. doi: 10.1016/j.talanta.2016.11.039
- He T, Wang Y, Li P, Zhang Q, Lei J, Zhang Z, et al. Nanobody-based enzyme immunoassay for aflatoxin in agro-products with high tolerance to cosolvent methanol. *Analyt Chem* (2014) 86:8873–80. doi: 10.1021/ac502390c
- Cheong W, Leow C, Majeed A, Leow C. Diagnostic and therapeutic potential of shark variable new antigen receptor (VNAR) single domain antibody. *Int J Biol Macromol* (2020) 147:369–75. doi: 10.1016/j.jbiomac.2020.01.039
- Wang K, Vasylieva N, Wan D, Eads D, Yang J, Tretten T, et al. Quantitative detection of fipronil and fipronil-sulfone in sera of black-tailed prairie dogs and rats after oral exposure to fipronil by camel single-domain antibody-based immunoassays. *Analyt Chem* (2019) 91:1532–40. doi: 10.1021/acs.analchem.8b04653
- Shan G, Huang H, Stoutamire D. A sensitive class specific immunoassay for the detection of pyrethroid metabolites in human urine. *Chem Res Toxicol* (2004) 17:218–25. doi: 10.1021/tx034220c
- Doyle P, Arbabi-Ghahroudi M, Gaudette N, Furzer G, Savard M, Gleddie S, et al. Cloning, expression, and characterization of a single-domain antibody fragment with affinity for 15-acetyl-deoxyvalenol. *Mol Immunol* (2008) 45:3703–13. doi: 10.1016/j.molimm.2008.06.005
- Zavrtanik U, Lukan J, Loris R, Lah J, Hadzi S. Structural basis of epitope recognition by heavy-chain camelid antibodies. *J Mol Biol* (2018) 430:4369–86. doi: 10.1016/j.jmb.2018.09.002
- Zielonka S, Empting M, Grzeschik J, Konning D, Barelle C, Kolmar H. Structural insights and biomedical potential of IgNAR scaffolds from sharks. *mAbs* (2015) 7:15–25. doi: 10.4161/19420862.2015.989032
- Muyldermans S. Nanobodies: Natural single-domain antibodies. *Annu Rev Biochem* (2013) 82:775–97. doi: 10.1146/annurev-biochem-063011-092449
- Juma S, Gong X, Hu S, Lv Z, Shao J, Liu L, et al. Shark new antigen receptor (IgNAR): Structure, characteristics and potential biomedical applications. *Cells* (2021) 10:1140. doi: 10.3390/cells10051140
- Goldman E, Liu J, Zabetakis D, Anderson G. Enhancing stability of camelid and shark single domain antibodies: An overview. *Front Immunol* (2017) 8:865. doi: 10.3389/fimmu.2017.00865
- Hoey R, Eom H, Horn J. Structure and development of single domain antibodies as modules for therapeutics and diagnostics. *Exp Biol Med* (2019) 0:1–9. doi: 10.1177/1535370219881129
- Khalid Z, Chen Y, Yu D, Abbas M, Huan M, Naz Z, et al. IgNAR antibody: Structural features, diversity and applications. *Fish Shellfish Immunol* (2022) 121:467–77. doi: 10.1016/j.fsi.2022.01.027
- Leow C, Fischer K, Leow C, Cheng Q, Chuah C, McCarthy J. Single domain antibodies as new biomarker detectors. *Diagnostics* (2017) 7:52. doi: 10.3390/diagnostics7040052
- Rabia L, Desai A, Jhaji H, Tessier P. Understanding and overcoming trade-offs between antibody affinity, specificity, stability and solubility. *Biochem Eng J* (2018) 137:365–74. doi: 10.1016/j.bej.2018.06.003
- Steeland S, Vandembroucke R, Libert C. Nanobodies as therapeutics: big opportunities for small antibodies. *Drug Discovery Today* (2016) 21:1076–113. doi: 10.1016/j.drudis.2016.04.003
- DeVlieger D, Ballegeer M, Rossey I, Schepens B, Saelens X. Single-domain antibodies and their formatting to combat viral infections. *Antibodies* (2018) 8:1. doi: 10.3390/antib8010001
- Konning D, Zielonka S, Grzeschik J, Empting M, Valldorf B, Krah S, et al. Camelid and shark single domain antibodies: structural features and therapeutic potential. *Curr Opin Struct Biol* (2017) 45:10–6. doi: 10.1016/j.sbi.2016.10.019
- Muyldermans S, Atarhouch T, Saldanha J, Barbosa J, Hamers R. Sequence and structure of V-h domain from naturally-occurring camel heavy-chain immunoglobulins lacking light-chains. *Protein Eng* (1994) 7:1129–35. doi: 10.1093/protein/7.9.1129
- Vu K, Ghahroudi M, Wyns L, Muyldermans S. Comparison of llama V-h sequences from conventional and heavy chain antibodies. *Mol Immunol* (1997) 34:1121–31. doi: 10.1016/S0161-5890(97)00146-6
- Harmen M, Ruuls R, Nijman I, Niewold T, Frenken L, deGeus B. Llama heavy-chain V regions consist of at least four distinct subfamilies revealing novel sequence features. *Mol Immunol* (2000) 37:579–90. doi: 10.1016/S0161-5890(00)00081-X
- Maass D, Sepulveda J, Pernthaner A, Shoemaker C. Alpaca (Lama pacos) as a convenient source of recombinant camelid heavy chain antibodies (VHHs). *J Immunol Methods* (2007) 324:13–25. doi: 10.1016/j.jim.2007.04.008
- Lefranc M, Ruiz C, Giudicelli V, Foulquier E, Truong L, Thouvenin-Contet V, et al. IMGT unique numbering for immunoglobulin and T cell receptor variable domains and ig superfamily V-like domains. *Dev Comp Immunol* (2003) 27:55–77. doi: 10.1016/S0145-305X(02)00039-3
- Nguyen V, Hamers R, Wyns L, Muyldermans S. Camel heavy-chain antibodies: diverse germline VHH and specific mechanisms enlarge the antigen-binding repertoire. *EMBO J* (2000) 19:921–30. doi: 10.1093/emboj/19.5.921
- Ebrahimizadeh W, Mousavi G, Javidan Z, Rajabibazl M. Production of novel VHH nanobody inhibiting angiogenesis by targeting binding site of VEGF. *Appl Biochem Biotechnol* (2015) 176:1985–95. doi: 10.1007/s12010-015-1695-y
- Stanfield R, Dooley H, Flajnik M, Wilson I. Crystal structure of a shark single-domain antibody V region in complex with lysozyme. *Science* (2004) 305:1770–3. doi: 10.1126/science.1101148
- Gonzalez-Sapienza G, Rossotti M, Rosa S. Single-domain antibodies as versatile affinity reagents for analytical and diagnostic applications. *Front Immunol* (2017) 8:977. doi: 10.3389/fimmu.2017.00977
- Zhang W, Qin L, Cai X, Juma S, Xu R, Wei L, et al. Sequence structure character of IgNAR sec in whitespotted bamboo shark (*Chiloscyllium plagiosum*). *Fish Shellfish Immunol* (2020) 102:140–4. doi: 10.1016/j.fsi.2020.04.037
- Zielonka S, Weber N, Becker S, Doerner A, Christmann A, Christmann C, et al. Shark attack: High affinity binding proteins derived from shark vNAR domains by stepwise *in vitro* affinity maturation. *J Biotechnol* (2014) 191:236–45. doi: 10.1016/j.jbiotec.2014.04.023
- Matz H, Munir D, Logue J, Dooley H. The immunoglobulins of cartilaginous fishes. *Dev Comp Immunol* (2021) 115:103873. doi: 10.1016/j.dci.2020.103873
- Dooley H, Flajnik M. Antibody repertoire development in cartilaginous fish. *Dev Comp Immunol* (2006) 30:43–56. doi: 10.1016/j.dci.2005.06.022

39. Cabanillas-Bernal O, Duenas S, Ayala-Avila M, Rucavado A, Escalante T, Licea-Navarro A. Synthetic libraries of shark vNAR domains with different cysteine numbers within the CDR3. *PLoS One* (2019) 14:e0213394. doi: 10.1371/journal.pone.0213394
40. Wesolowski J, Alzogaray V, Reyelt J, Unger K, Juarez K, Urrutia M, et al. Single domain antibodies: Promising experimental and therapeutic tools in infection and immunity. *Med Microbiol Immunol* (2009) 198:157–74. doi: 10.1007/s00430-009-0116-7
41. English H, Hong J, Ho M. Ancient species offers contemporary therapeutics: An update on shark VNAR single domain antibody sequences, phage libraries and potential clinical applications. *Antibody Ther* (2020) 3:1–9. doi: 10.1093/abt/tbaa001
42. Kovaleva M, Ferguson L, Steven J, Porter A, Barelle C. Shark variable new antigen receptor biologics - a novel technology platform for therapeutic drug development. *Expert Opin Biol Ther* (2014) 14:1527–39. doi: 10.1517/14712598.2014.937701
43. Roux K, Greenberg A, Greene L, Strelets L, Avila D, McKinney C, et al. Structural analysis of the nurse shark (new) antigen receptor (NAR): Molecular convergence of NAR and unusual mammalian immunoglobulins. *PNAS* (1998) 95:11804–9. doi: 10.1073/pnas.95.20.11804
44. Streltsov V, Varghese J, Carmichael J, Irving R, Hudson P, Nuttall D, et al. Structural evidence for evolution of shark ig new antigen receptor variable domain antibodies from a cell-surface receptor. *PNAS* (2004) 101:12444–9. doi: 10.1073/pnas.0403509101
45. Streltsov V, Carmichael J, Nuttall S. Structure of a shark IgNAR antibody variable domain and modeling of an early-developmental isotype. *Protein Sci* (2005) 14:2901–9. doi: 10.1110/ps.051709505
46. Nuttall S, Krishnan U, Hattarki M, DeGori R, Irving R, Hudson P. Isolation of the new antigen receptor from wobbegong sharks, and use as a scaffold for the display of protein loop libraries. *Mol Immunol* (2001) 38:313–26. doi: 10.1016/S0161-5890(01)00057-8
47. Liu J, Anderson G, Delehanty J, Baumann R, Hayhurst A, Goldman E. Selection of cholera toxin specific IgNAR single-domain antibodies from a naïve shark library. *Mol Immunol* (2007) 44:1775–83. doi: 10.1016/j.molimm.2006.07.299
48. Crouch K, Smith L, Williams R, Cao W, Lee M, Jensen A, et al. Humoral immune response of the small-spotted catshark, *Scyliorhinus canicularis*. *Fish Shellfish Immunol* (2013) 34:1158–69. doi: 10.1016/j.fsi.2013.01.025
49. Diaz M, Velez J, Singh M, Cerny J, Flajnik M. Mutational pattern of the nurse shark antigen receptor gene (NAR) is similar to that of mammalian ig genes and to spontaneous mutations in evolution: The translation synthesis model of somatic hypermutation. *Int Immunol* (1999) 11:825–33. doi: 10.1093/intimm/11.5.825
50. Diaz M, Stanfield R, Greenberg A, Flajnik M. Structural analysis, selection, and ontogeny of the shark new antigen receptor (IgNAR): identification of a new locus preferentially expressed in early development. *Immunogenetics* (2002) 54:501–12. doi: 10.1007/s00251-002-0479-z
51. Feng M, Bian H, Wu X, Fu T, Hong J, Fleming B, et al. Construction and next-generation sequencing analysis of a large phage-displayed VNAR single-domain antibody library from six naïve nurse sharks. *Antibody Ther* (2019) 2:1–11. doi: 10.1093/abt/tby011
52. Chothia C, Lesk A. Canonical structures for the hypervariable regions of immunoglobulins. *J Mol Biol* (1987) 196:901–17. doi: 10.1016/0022-2836(87)90412-8
53. Stave J, Lindpaintner K. Antibody and antigen contact residues define epitope and paratope size and structure. *J Immunol* (2013) 191:1428–35. doi: 10.4049/jimmunol.1203198
54. Sundberg E, Mariuzza R. Molecular recognition in antibody-antigen complexes. *Adv Protein Chem* (2002) 61:119–80. doi: 10.1016/S0065-3233(02)61004-6
55. D'Angelo S, Ferrara F, Naranjo L, Erasmus M, Hrabec P, Bradbury A. Many routes to an antibody heavy-chain CDR3: necessary, yet insufficient, for specific binding. *Front Immunol* (2018) 9:395. doi: 10.3389/fimmu.2018.00395
56. Barelle C, Gill D, Charlton K. Shark novel antigen receptors - the next generation of biologic therapeutics? *Adv Exp Med Biol* (2009) 655:49–62. doi: 10.1007/978-1-4419-1132-2_6
57. Nuttall S, Walsh R. Display scaffolds: protein engineering for novel therapeutics. *Current Opin Pharmacol* (2008) 8:609–15. doi: 10.1016/j.coph.2008.06.007
58. Hinds K, Litman G. Major reorganization of immunoglobulin VH segmental elements during vertebrate evolution. *Nature* (1986) 320:546–51. doi: 10.1038/320546a0
59. Mordenti J, Cuthbertson R, Ferrara N, Thomsen K, Berleau L, Licko V, et al. Comparisons of the intraocular tissue distribution, pharmacokinetics, and safety of 125I-labeled full-length and Fab antibodies in rhesus monkeys following intravitreal administration. *Toxicol Pathol* (1999) 27:536–44. doi: 10.1177/019262339902700507
60. De Genst E, Silence K, Decanniere K, Conrath K, Loris R. Molecular basis for the preferential cleft recognition by dromedary heavy-chain antibodies. *PNAS* (2006) 103:4586–91. doi: 10.1073/pnas.0505379103
61. Koide A, Tereshko V, Uysal S, Margalef K, Kossiakoff A, Koide S. Exploring the capacity of minimalist protein interfaces: interface energetics and affinity maturation to picomolar K(D) of a single-domain antibody with a flat paratope. *J Mol Biol* (2007) 373:941–53. doi: 10.1016/j.jmb.2007.08.027
62. King M, Huh I, Shenai A, Brooks T, Brooks C. Structural basis of VHH-mediated neutralization of the food-borne pathogen *Listeria monocytogenes*. *J Biol Chem* (2018) 293:13626–35. doi: 10.1074/jbc.RA118.003888
63. Kubala M, Kovtun O, Alexandrov K, Collins B. Structural and thermodynamic analysis of the GFP: GFP-nanobody complex. *Protein Sci* (2010) 19:2389–401. doi: 10.1002/pro.519
64. Kirchofer A, Helma J, Schmidthals K, Frauer C, Cui S, Karcher A, et al. Modulation of protein properties in living cells using nanobodies. *Nat Struct Mol Biol* (2010) 17:133–8. doi: 10.1038/nsmb.1727
65. Flajnik M, Deschacht N, Muyldermans SA. Case of convergence: why did a simple alternative to canonical antibodies arise in sharks and camels? *PLoS Biol* (2011) 9:e1001120. doi: 10.1371/journal.pbio.1001120
66. Spinelli S, Tegoni M, Frenken L, van Vliet C, Cambillau C. Lateral recognition of a dye hapten by a llama VHH domain. *J Mol Biol* (2001) 311:123–9. doi: 10.1006/jmbi.2001.4856
67. Spinelli S, Frenken L, Hermans P, Verris T, Brown L. Camelid heavy-chain variable domains provide efficient combining sites to haptens. *Biochemistry* (2000) 39:1217–22. doi: 10.1021/bi991830w
68. Henry K, MacKenzie C. Antigen recognition by single-domain antibodies: Structural latitudes and constraints. *mAbs* (2018) 10:815–26. doi: 10.1080/19420862.2018.1489633
69. Mitchell L, Colwell L. Analysis of nanobody paratopes reveals greater diversity than classical antibodies. *Protein Eng Des Select* (2018) 31:267–75. doi: 10.1093/protein/gzy017
70. Mitchell L, Colwell L. Comparative analysis of nanobody sequence and structure data. *Proteins* (2018) 86:697–706. doi: 10.1002/prot.25497
71. Bond C, Marster Jr J, Sidhu S. Contributions of CDR3 to VHH domain stability and the design of monobody scaffolds for naïve antibody libraries. *J Mol Biol* (2003) 332:643–55. doi: 10.1016/S0022-2836(03)00967-7
72. Akiba H, Tamura H, Kiyoshi M, Yanaka S, Sugase K, Caaveiro J, et al. Structural and thermodynamic basis for the recognition of the substrate-binding cleft on hen egg lysozyme by a single-domain antibody. *Sci Rep* (2019) 9:15481. doi: 10.1038/s41598-019-50722-y
73. Sefid F, Payandeh Z, Azamirad G, Abdolhamidi R, Rasooli I. *In silico* engineering towards enhancement of bap-VHH monoclonal antibody binding affinity. *Int J Pept Res Ther* (2019) 25:273–87. doi: 10.1007/s10989-017-9670-9
74. Soler M, Medagli B, Semrau M, Storici P, Bajc G, Marco A, et al. A consensus protocol for the *in silico* optimization of antibody fragments. *Chem Commun* (2019) 55:14043–6. doi: 10.1039/C9CC06182G
75. Vishwakarma P, Vattekatte A, Shinada N, Diharce J, Martins C, Cadet F, et al. VHH structural modelling approaches: A critical review. *Int J Mol Sci* (2022) 23:3721. doi: 10.3390/ijms23073721
76. Desmyter A, Transue T, Ghahroudi M, Thi M, Poortmans F, Hamers R, et al. Crystal structure of a camel single-domain VH antibody fragment in complex with lysozyme. *Nat Struct Biol* (1996) 3:803–11. doi: 10.1038/nsb0996-803
77. Jiao L, Liu Y, Zhang X, Liu B, Zhang C, Liu X. Site-saturation mutagenesis library construction and screening for specific broad-spectrum single-domain antibodies against multiple Cry1 toxins. *Appl Microbiol Biotechnol* (2017) 101:6071–82. doi: 10.1007/s00253-017-8347-9
78. Qiu Y, He Q, Xu Y, Wang W, Liu Y. Modification of a deoxyribose-antigen-mimicking nanobody to improve immunoassay sensitivity by site-saturation mutagenesis. *Analyt Bioanal Chem* (2016) 408:895–903. doi: 10.1007/s00216-015-9181-5
79. Nie J, Ma X, Hu F, Miao H, Feng X, Zhang P, et al. Designing and constructing a phage display synthesized single domain antibodies library based on camel VHHs frame for screening and identifying humanized TNF- α -specific nanobody. *Biomed Pharmacother* (2021) 137:111328. doi: 10.1016/j.biopha.2021.111328
80. Ragland S, Criss A. From bacterial killing to immune modulation: Recent insights into the functions of lysozyme. *PLoS Pathog* (2017) 13:e1006512. doi: 10.1371/journal.ppat.1006512
81. De Genst E, Handelberg F, Meirhaeghe V. Chemical basis for the affinity maturation of a camel single domain antibody. *J Biol Chem* (2004) 279:53593–601. doi: 10.1074/jbc.M407843200

82. Dooley H, Stanfield R, Brady R, Flajnik M. First molecular and biochemical analysis of *in vivo* affinity maturation in an ectothermic vertebrate. *PNAS* (2006) 103:1846–51. doi: 10.1073/pnas.0508341103
83. Nuttall S, Humberstone K, Krishnan U, Carmichael J, Doughty L, Hattarki M, et al. Selection and affinity maturation of IgNAR variable domains targeting plasmodium falciparum AMA1. *Proteins* (2004) 55:187–97. doi: 10.1002/prot.20005
84. Kopsidas G, Roberts A, Coia G, Streltsov V, Nuttall S. *In vitro* improvement of a shark IgNAR antibody by q β replicase mutation and ribosome display mimics *in vivo* affinity maturation. *Immunol Lett* (2006) 107:163–8. doi: 10.1016/j.imlet.2006.09.004
85. Ding L, Wang Z, Zhong P, Jiang H, Zhao Z, Zang Y, et al. Structural insights into the mechanism of single domain VHH antibody binding to cortisol. *FEBS Lett* (2019) 593:1248–56. doi: 10.1002/1873-3468.13398
86. Rosa S, Wogulis L, Wogulis M, Gonzalez-Sapienza G, Wilson D. Structure and specificity of several tricolobin-binding single domain camelid antibody fragments. *J Mol Recognit* (2018) 32:e2755. doi: 10.1002/jmr.2755
87. Decanniere K, Desmyter A, Lauwereys M, Ghahrouti M, Muyldermans S, Wyns LA. Single-domain antibody fragment in complex with RNase a: non-canonical loop structures and nanomolar affinity using two CDR loops. *Structure* (1999) 7:361–70. doi: 10.1016/S0969-2126(99)80049-5
88. Genst E, Williams T, Wellens J, O'Day E, Waudby C, Meehan S, et al. Structure and properties of a complex of α -synuclein and a single-domain camelid antibody. *J Mol Biol* (2010) 402:326–43. doi: 10.1016/j.jmb.2010.07.001
89. Fanning S, Horn J. An anti-hapten camelid antibody reveals a cryptic binding site with significant energetic contributions from a nonhypervariable loop. *Protein Sci* (2011) 20:1196–207. doi: 10.1002/pro.648
90. Waxman E, Giasson B. Molecular mechanisms of alpha-synuclein neurodegeneration. *Biochim Biophys Acta* (2009) 1792:616–24. doi: 10.1016/j.bbdis.2008.09.013
91. Rubinstein N, Mayrose I, Halperin D, Yekutieli D, Gershoni J, Pupko T. Computational characterization of b-cell epitopes. *Mol Immunol* (2008) 45:3477–89. doi: 10.1016/j.molimm.2007.10.016
92. Desmyter A, Spinelli S, Payan F, Lauwereys M, Wyns L, Muyldermans S, et al. Three camelid VHH domains in complex with porcine pancreatic α -amylase. *J Biol Chem* (2002) 277:23645–50. doi: 10.1074/jbc.M202327200
93. Xi X, Sun W, Su H, Zhang X, Sun F. Identification of a novel anti-EGFR nanobody by phage display and its distinct paratope and epitope via homology modeling and molecular docking. *Mol Immunol* (2020) 128:165–74. doi: 10.1016/j.molimm.2020.10.014
94. Lesne J, Chang H, Visch A, Paloni M, Barther P, Guichou J, et al. Structural basis for chemically-induced homodimerization of a single domain antibody. *J Sci Rep* (2019) 9:1840. doi: 10.1038/s41598-019-38752-y
95. Stanfield R, Dooley H, Verdino P, Flajnik M, Wilson I. Maturation of shark single-domain (IgNAR) antibodies: evidence for induced-fit binding. *J Mol Biol* (2007) 367:358–72. doi: 10.1016/j.jmb.2006.12.045
96. Zielonka S, Empting M, Konning D, Grzeschik J, Krah S, Becker S, et al. The shark strikes twice: Hypervariable loop 2 of shark IgNAR antibody variable domains and its potential to function as an autonomous paratope. *Mar Biotechnol* (2015) 17:1–7. doi: 10.1007/s10126-015-9642-z
97. Kontermann R. Dual targeting strategies with bispecific antibodies. *mAbs* (2012) 4:182–97. doi: 10.4161/mabs.4.2.19000
98. Lindstedt P, Sormanni F, Rakoto R, Dobson C, Bernardes G, Vendruscolo M. Systematic activity maturation of a single-domain antibody with non-canonical amino acids through chemical mutagenesis. *Cell Chem Biol* (2021) 28:70–7. doi: 10.1016/j.chembiol.2020.11.002
99. Kobayashi N, Oyama H, Nakano M, Kanda T, Banzono E, Kato Y, et al. “Cleavable” hapten-biotin conjugates: Preparation and use for the generation of anti-steroid single-domain antibody fragments. *Analyt Biochem* (2009) 387:257–66. doi: 10.1016/j.ab.2009.01.004
100. Yau K, Dubuc G, Li S, Hiramata T, MacKenzie C, Jremutus L, et al. Affinity maturation of a VHH by mutational hotspot randomization. *J Immunol Methods* (2005) 297:213–24. doi: 10.1016/j.jim.2004.12.005
101. Sheedy C, Yau K, Hiramata T, MacKenzie C, Hall J. Selection, characterization, and CDR shuffling of naïve llama single-domain antibodies selected against auxin and their cross-reactivity with auxinic herbicides from four chemical families. *J Agric Food Chem* (2006) 54:3668–78. doi: 10.1021/jf060219i
102. Harmsen M, Solt C, Bommel A, Niewold T, Zijderveld F. Selection and optimization of proteolytically stable llama single-domain antibody fragments for oral immunotherapy. *Appl Microbiol Biotechnol* (2006) 72:544–51. doi: 10.1007/s00253-005-0300-7
103. Wang X, Chen Q, Sun Z, Wang Y, Su B, Zhang C, et al. Nanobody affinity improvement: Directed evolution of the anti-ochratoxin a single domain antibody. *Int J Biol Macromol* (2020) 151:312–21. doi: 10.1016/j.ijbiomac.2020.02.180
104. Tiller K, Chowdhury R, Li T, Ludwig S, Sen S, Maranas C, et al. Facile affinity maturation of antibody variable domains using natural diversity mutagenesis. *Front Immunol* (2017) 8:986. doi: 10.3389/fimmu.2017.00986
105. Turner K, Liu J, Zabetakis D, Lee A, Anderson G, Goldman E. Improving the biophysical properties of anti-ricin single-domain antibodies. *Biotechnol Rep* (2015) 6:27–35. doi: 10.1016/j.btre.2015.01.001
106. Liu J, Shriver-Lake L, Anderson G, Zabetakis D, Goldman E. Selection, characterization, and thermal stabilization of llama single domain antibodies towards Ebola virus glycoprotein. *Microbial Cell Fact* (2017) 16:223. doi: 10.1186/s12934-017-0837-z
107. Hagihara Y, Mine S, Uegaki K. Stabilization of an immunoglobulin fold domain by an engineered disulfide bond at the buried hydrophobic region. *J Biol Chem* (2007) 282:36489–95. doi: 10.1074/jbc.M707078200
108. Anderson G, Liu J, Zabetakis D, Liu J, Goldman E. Thermal stabilization of anti- α -cobratoxin single domain antibodies. *Toxicon* (2017) 129:68–73. doi: 10.1016/j.toxicon.2017.02.008
109. Wilton E, Opyr M, Kailasam S, Kothe R, Wieden H. SdAb-DB: The single domain antibody database. *ACS Synthetic Biol* (2018) 7:2480–4. doi: 10.1021/acssynbio.8b00407
110. Schneider C, Raybould M, Deane C. SABDab in the age of biotherapeutics: updates including SABDab-nano, the nanobody structure tracker. *Nucleic Acids Res* (2022) 50:1368–72. doi: 10.1093/nar/gkab1050
111. Lohr T, Sormanni P, Vendruscolo M. Conformational entropy as a potential liability of computationally designed antibodies. *Biomolecules* (2022) 12:718. doi: 10.3390/biom12050718
112. Lippow S, Witttrup K, Tidor B. Computational design of antibody-affinity improvement beyond *in vivo* maturation. *Nat Biotechnol* (2007) 25:1171–6. doi: 10.1038/nbt1336
113. Winter G, Griffiths A, Hawkins R, Hoogenboom H. Making antibodies by phage display technology. *Annu Rev Immunol* (1994) 12:433–55. doi: 10.1146/annurev.iy.12.040194.002245
114. Dobson CL, Devine P, Phillips J, Higazi D, Lloyd C, Popovic B. Engineering the surface properties of a human monoclonal antibody prevents self-association and rapid clearance *in vivo*. *Sci Rep* (2016) 6:38644. doi: 10.1038/srep38644
115. Chi SW, Maeng C, Kim S, Oh M, Ryu C, Kim S. Broadly neutralizing anti-hepatitis b virus antibody reveals a complementarity determining region H3 lid-opening mechanism. *Proc Natl Acad Sci USA* (2007) 104:9230–5. doi: 10.1073/pnas.0701279104
116. Swindells M, Porter C, Couch M, Hurst J, Abhinandan K, Nielsen J, et al. abYsis: integrated antibody sequence and structure-management, analysis, and prediction. *J Mol Biol* (2017) 429:356–64. doi: 10.1016/j.jmb.2016.08.019
117. Lim C, Choong Y, Lim T. Cognizance of molecular methods for the generation of mutagenic phage display antibody libraries for affinity maturation. *Int J Mol Sci* (2019) 20:1861. doi: 10.3390/ijms20081861
118. Hanes J, Pluckthun A. *In vitro* selection and evolution of functional proteins by using ribosome display. *PNAS* (1997) 94:4937–42. doi: 10.1073/pnas.94.10.4937
119. Fennell B, Sheenhan A, Hufton S, Calabro V, Wu L, Muller M, et al. Dissection of the IgNAR V domain: Molecular scanning and orthologue database mining define novel IgNAR hallmarks and affinity maturation mechanisms. *J Mol Biol* (2010) 400:155–70. doi: 10.1016/j.jmb.2010.04.061
120. Irving R, Coia G, Roberts A, Nuttall S, Hudson P. Ribosome display and affinity maturation: from antibodies to single V-domains and stepstowards cancer therapeutics. *J Immunol Methods* (2001) 248:31–45. doi: 10.1016/S0022-1759(00)00341-0
121. Tabasinezhad M, Talebkhan Y, Wenzel W, Rahimi H, Omidinia E, Mahboudi F. Trends in therapeutic antibody affinity maturation: From *in-vitro* towards next-generation sequencing approaches. *Immunol Lett* (2019) 212:106–13. doi: 10.1016/j.imlet.2019.06.009
122. Garnier J. Protein structure prediction. *Biochimie* (1990) 72:513–24. doi: 10.1016/0300-9084(90)90115-W
123. Baker D, Sali A. Protein structure prediction and structural genomics. *Science* (2001) 294:93–6. doi: 10.1126/science.1065659
124. Rosenberg M, Goldblum A. Computational protein design: a novel path to future protein drugs. *Curr Pharm Des* (2006) 12:3973–97. doi: 10.2174/138161206778743655
125. Lippow S, Tidor B. Progress in computational protein design. *Curr Opin Biotechnol* (2007) 18:305–11. doi: 10.1016/j.copbio.2007.04.009
126. Marshall S, Vizcarra C, Mayo S. One-and two-body decomposable poisson-Boltzmann methods for protein design calculations. *Protein Sci* (2005) 14:1293–304. doi: 10.1110/ps.041259105
127. Bystroff C, Shao Y. Fully automated ab initio protein structure prediction using i-sites, hmmstr and rosetta. *Bioinformatics* (2022) 18:54–61. doi: 10.1093/bioinformatics/18.suppl_1.S54

128. Rohl C, Strauss C, Misura K, Baker D. Protein structure prediction using Rosetta. *Methods Enzymol* (2004) 383:66–93. doi: 10.1016/S0076-6879(04)83004-0

129. Yang J, Yan R, Roy A, Xu D, Poisson J, Zhang Y. The I-TASSER suite: Protein structure and function prediction. *Nat Methods* (2015) 12:7–8. doi: 10.1038/nmeth.3213

130. Jumper J, Evans R, Prizel A, Green T, Figurnov M, Ronneberger O, et al. Highly accurate protein structure prediction with AlphaFold. *Nature* (2021) 596:583–9. doi: 10.1038/s41586-021-03819-2

131. Du Z, Su H, Wang W, Ye L, Wei H, Peng Z, et al. The TrRosetta server for fast and accurate protein structure prediction. *Nat Protocol* (2021) 16:5634–51. doi: 10.1038/s41596-021-00628-9