



OPEN

## Development and characterization of a camelid derived antibody targeting a linear epitope in the hinge domain of human PCSK9 protein

Xinyang Li<sup>1,4</sup>, Jun Hong<sup>1,4</sup>, Xiaoyan Gao<sup>2</sup>, Meiniang Wang<sup>2</sup> & Naibo Yang<sup>2,3</sup>

PCSK9 is an effective target for lowering LDL-c. Previously, a camelid-human chimeric heavy chain antibody VHH-B11-Fc targeting human PCSK9 was designed. It had a potent hypolipidemic effect. However, the nanobody VHH-B11 interacts with PCSK9 at low affinity, while camelid VHH exhibits some immunogenicity. Moreover, the interacting epitope is yet to be identified, although VHH-B11 was shown to have distinct hPCSK9-binding epitopes for Evolocumab. This might impede the molecule's progress from bench to bedside. In the present study, we designed various configurations to improve the affinity of VHH-B11 with hPCSK9 (<10 nM) that in turn enhanced the druggability of VHH-B11-Fc. Then, 17 amino acids were specifically mutated to increase the degree of humanization of the nanobody VHH-B11. Using phage display and sequencing technology, the linear epitope "STHGAGW" (amino acids 447–452) was identified in the hinge region of PCSK9 as the interacting site between VHH-B11-Fc and hPCSK9. Unlike the interaction epitope of Evolocumab, located in the catalytic region of PCSK9, the binding epitope of VHH-B11 is located in the hinge region of PCSK9, which is rarely reported. These findings indicated that a specific mechanism underlying this interaction needs to be explored.

### Abbreviations

(G4S)3-linker	GlyGlyGlyGlySerGlyGlyGlyGlySerGlyGlyGlyGlySer
Antibody	Ab
Antigen	Ag
BSA	Bovine serum albumin
DNA	Deoxyribonucleic acid
<i>E. coli</i>	<i>Escherichia coli</i>
ELISA	Enzyme linked immunosorbent assay
Fc	The constant region fragment of the human immunoglobulin gamma
FR	Framework region
HcAb	Heavy chain antibody
HEK293F	Human embryonic kidney 293F cells
hPCSK9	Human proprotein convertase subtilisin/kexin type 9
HRP	Horseradish peroxidase
Hu-B11	Humanized VHH-B11
IgG	Immunoglobulin gamma
IPTG	Isopropyl-β-d-thiogalactoside
kDa	Kilo dalton
LDL	Low density lipoprotein
LDL-c	Low density lipoprotein cholesterol

<sup>1</sup>College of Life Science and Engineering, Henan University of Urban Construction, Pingdingshan 467036, China. <sup>2</sup>BGI-Shenzhen, Shenzhen 518083, China. <sup>3</sup>Complete Genomics, Inc., 2904 Orchard Parkway, San Jose, CA 95134, USA. <sup>4</sup>These authors contributed equally: Xinyang Li and Jun Hong. ✉email: yipinyoulan521@163.com; yangnaibo@genomics.cn

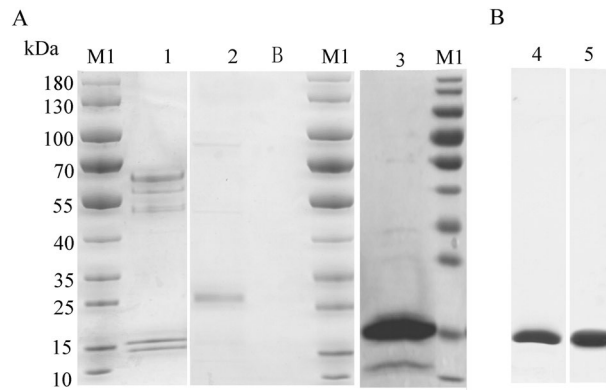
LDLR	Low density lipoprotein receptor
OD <sub>600</sub>	The optical density value at 600 nm
PCSK9	Proprotein convertase subtilisin/kexin type 9
PCR	Polymerase chain reaction
<i>P. pastoris</i>	<i>Pichia pastoris</i>
RFU	The relative fluorescence unit
RU	Response units
SD rat	Sprague Dawley rat
sdAb	Single domain antibody
SPR	Surface plasmon resonance
TES buffer	0.2 M Tris-HCl (pH 8.0), 0.5 mM EDTA, 0.5 M sucrose
VHH	High variable region of the heavy chain antibody
kDa	Kilo dalton
PBST	Phosphate buffer solution with 0.05% tween 20
PE	Pair end
Ph.D.-12 library	It is a combinatorial library of random 12-mer peptides fused to a minor coat protein (pIII) of M13 phage. The displayed peptide is expressed at the N-terminus of pIII.
RNA	Ribonucleic acid
YPD	1% Yeast extract, 2% peptone, 2% dextrose glucose

Cardiovascular and cerebrovascular diseases refer to the general term for some diseases, including coronary heart disease, myocardial infarction, cerebral apoplexy, myocarditis, and other ischemic or hemorrhagic diseases caused by the following symptoms: hyperlipidemia, atherosclerosis, and/or hypertension<sup>1</sup>. Despite adopting the most advanced treatments for cardiovascular and cerebrovascular diseases, > 50% of survivors of these diseases are unable to care for themselves and the quality of life declines severely. Previous studies have reported that cardiovascular disease has a mortality of up to 15 million people each year, surpassing cancer as the leading cause of death<sup>2–4</sup>. Presently, 3-hydroxy-3-methylglutaryl-coA reductase inhibitor is the first-line clinical lipid-lowering statin. However, recent studies have shown that some patients will develop drug tolerance if statin is administered for a prolonged time; then, the treatment needs to be discontinued, seeking other lipid-lowering approaches<sup>1,2,5</sup>.

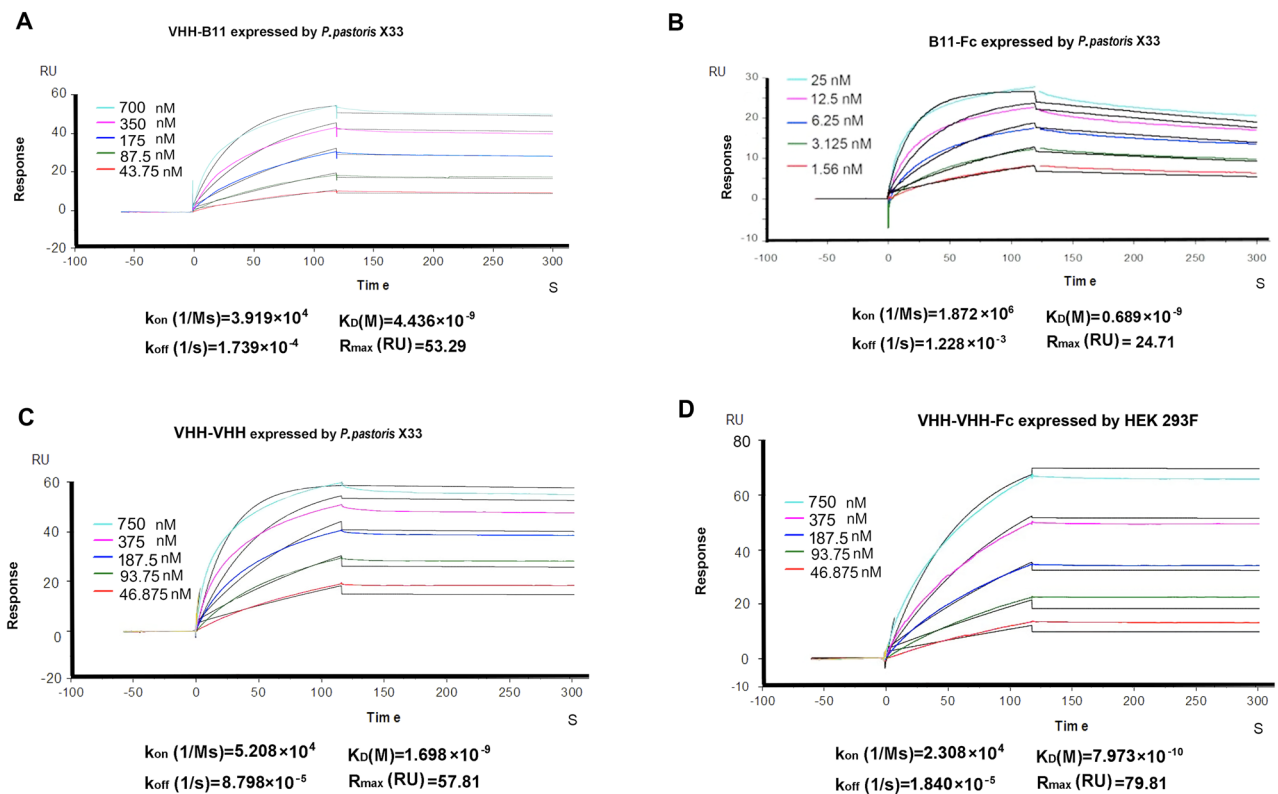
An effective and safe drug is the monoclonal antibody (Ab)<sup>6,7</sup>. In a previous study, we designed and developed a llama-human chimera heavy chain VHH-B11-Fc Ab, which showed similar lipid-lowering effects to the approved Evolocumab drug in hepatoma cells and human *proprotein convertase subtilisin/kexin 9* (hPCSK9) transgenic rat model<sup>7</sup>. However, the druggability of this Ab is not excellent, which is manifested as inadequate affinity (> 0.1 nM), immunogenicity (camelid variable region) and unknown epitopes. Therefore, the present study aimed to improve the druggability of this potential drug molecule in the following aspects. The affinity of therapeutic Ab should be ≤ 0.1 nM. Hence, we attempted to improve the affinity between the VHH-B11 Ab and PCSK9. For instance, the transformation of fragmentary Ab from monovalent to polyvalent effectively improves the affinity and function<sup>8–10</sup>. This is mainly achieved by short peptide ligation (“Cablivi” from Ablynx Inc.) and Fc fusion (“KN035” from Alphamab Oncology Inc.)<sup>11–13</sup>. Secondly, good therapeutic Abs should have low immunogenicity. The high immunogenicity may cause major side effects<sup>14</sup>. For instance, Pfizer had discontinued the global development of bococizumab because of its high immunogenicity in 2016<sup>15</sup>. In this study, the humanization degree of the VHH-B11 was improved by replacing some hot spot amino acids<sup>16</sup>. Finally, we identified the interaction epitope between the VHH-B11-Fc Ab and human PCSK9 (hPCSK9) based on sequencing, phage display, and random polypeptide library technologies. All in a word, the present study aimed to develop and characterize a camelid derived Ab (VHH-B11-Fc) reported previously.

## Results

**Monovalent nanobody VHH-B11 expression in different systems.** Multiple systems were adopted to express the VHH-B11 Ab to seek a cost-effective Ab preparation method, preferably with one-step affinity purification for later preparation. Briefly, VHH was expressed in three expression systems (*Escherichia coli* (*E. coli*), *Pichia pastoris* (*P. pastoris*) and HEK293F mammalian cells). As shown in Fig. 1 and Figs. S1–S5, the VHH-B11 Ab was expressed successfully in the three different systems. VHH-B11 was expressed in BL21 *E. coli* (lane 1: *pET28a-VHH-B11*), TG1 *E. coli* (lane 2: *pMECS-VHH-B11*), and HB2151 *E. coli* (lane 3: *pMECS-VHH-B11*) (Fig. 1A and Figs. S1–S3). Due to the intracellular and low soluble expression in BL21 *E. coli* (*pET28a*, lane 1), adequate results by only one step Ni<sup>2+</sup> affinity purification is difficult. Also, ultrasonication resulted in partial degradation of the VHH-B11 Ab. In the case of TG1 *E. coli* (lane 2: *pMECS-VHH-B11*), the undesired protein band (~ 30 kDa) appeared on the gel. Lane B is the blank control. In lane 3, HB2151 *E. coli* (*pMECS-VHH-B11*) also seems to be not very suitable for VHH protein expression. Because one-step Ni<sup>2+</sup>-affinity purification provided an unsatisfactory eluate. As shown in Fig. 1B and Figs. S4, S5, the monovalent nanobody VHH-B11 was expressed in *P. pastoris* (lane 4) and HEK293F mammalian cells (lane 5). Because of the secretory expression, a favorable purification was achieved in a single Ni<sup>2+</sup> affinity purification step without lysis. Due to the amount of heteroprotein bands that may require multi-stage purification, the three *E. coli* expression systems are not our preferred choice. Besides, considering the competitive advantages of low cost and convenient purification, yeast secretion seems to be a more optimal VHH expression system than that of mammalian cells. Furthermore, the affinity between PCSK9 and VHH-B11 expressed by *P. pastoris* was determined by surface plasmon resonance (SPR) technology. As shown in Fig. 2A, VHH-B11 had an interaction affinity with PCSK9 at 4.436 nM.

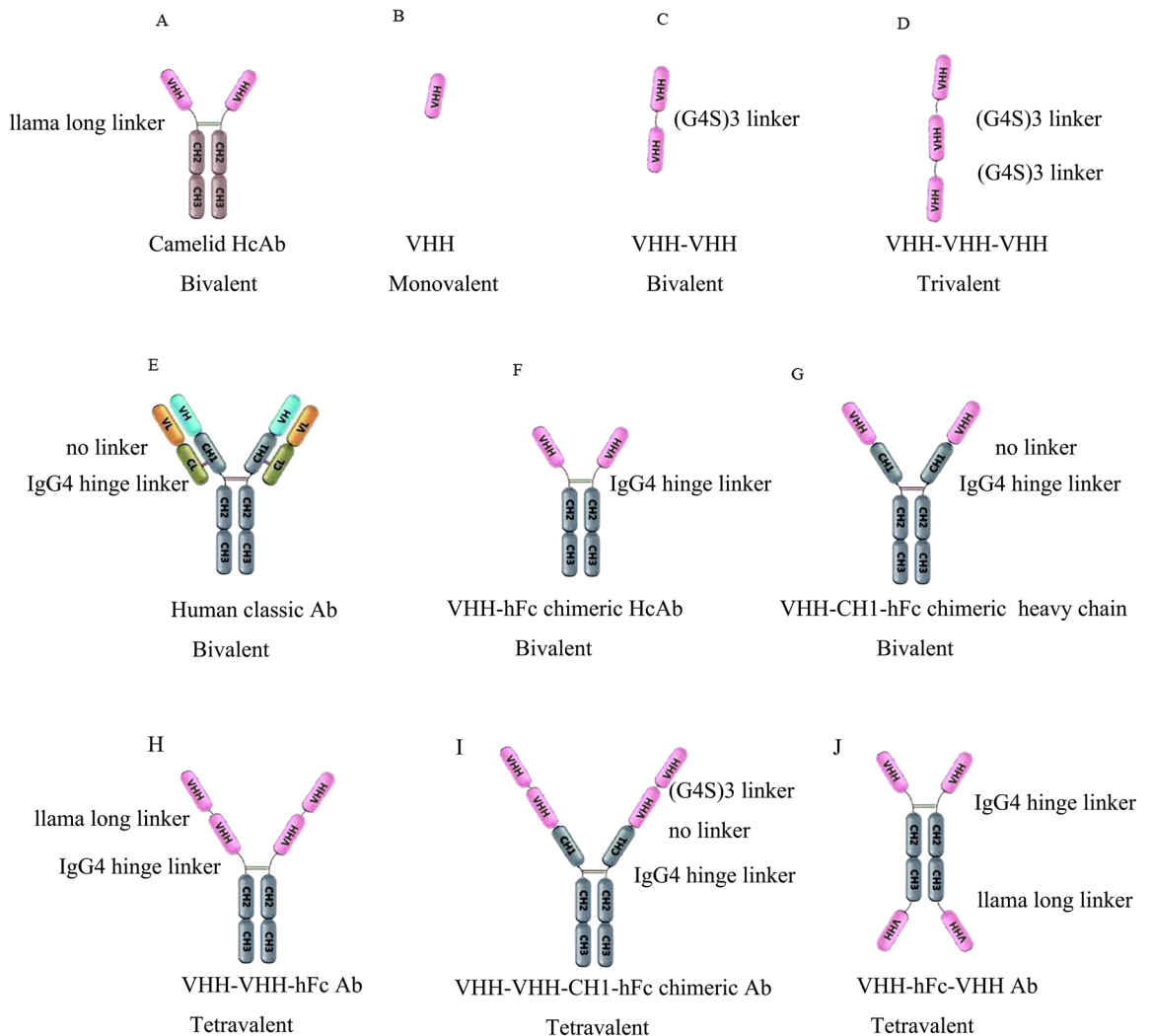


**Figure 1.** The SDS-PAGE gel of the VHH-B11. M1: the protein marker (Fermentas, USA). (A) Lane 1: the VHH-B11 antibody expressed by BL21 *E. coli* (*pET28a*); lane 2: the VHH-B11 expressed by TG1 *E. coli* (*pMECS*); lane B: blank control; lane 3: the VHH-B11 expressed by HB2151 *E. coli* (*pMECS*); (B) lane 4: the VHH-B11 expressed by *P. pastoris* X33 (*pPICZα*); lane 5: the VHH-B11 expressed by HEK293F (*pCDNA3.4*).



**Figure 2.** Affinity tests between the VHH-B11 or its derived antibodies and human PCSK9. The affinity was determined by SPR technology. Taking VHH-B11 for example, two-fold dilution of the VHH-B11 Ab (five concentrations in nM) expressed by *P. pastoris* X33 were loaded to bind the hPCSK9 coated on the CM5 chip. Each colored line represents a concentration (in nM) of the VHH-B11 Ab. The black line was the automatic fitting curve by the built-in evaluation software of Biacore T200. The horizontal axis represents the injection time, and the vertical axis represents the response unit (RU) value of the VHH-B11 interaction with hPCSK9. The VHH-B11 injection time point was set as 0 s by the evaluation software. (A) VHH-B11; (B) VHH-B11-Fc; (C) VHH-VHH; (D) VHH-VHH-Fc.

**Polyvalent configuration design and comparison.** Figure 3A,E show the camelid heavy chain Ab (HcAb) and classic human Ab. The variable region of HcAb is also described as VHH (Fig. 3B). To increase the affinity and prolong the half-life of the VHH, we designed a series of VHH-based polyvalent Abs that can be divided into two classes of Ab configurations (Fig. 3). One was for non-IgG-like strategy (Fig. 3C,D) and the other for IgG-like strategy (Fig. 3F–J). However, both increased the affinity of VHH-B11 by > ten-fold or about



**Figure 3.** Schematic representation of Ab configurations. Schematic representation of different Abs' configurations. (A) HcAb has its llama hinge linker: bivalent; (B) Camelid VHH: monovalent; (C) VHH-(G4S)3-VHH: bivalent; (D) VHH-(G4S)3-VHH-(G4S)3-VHH: trivalent; (E) Human classic Ab with its hinge linker: bivalent; (F) Llama VHH-human IgG4 hinge linker-IgG4 Fc chimeric HcAb: bivalent; (G) Llama VHH-human CH1-IgG4 hinge linker-IgG4 Fc chimeric heavy chain: bivalent; (H) VHH-llama long linker-VHH-human IgG4 hinge linker-IgG4 Fc: tetraivalent; (I) VHH-(G4S)3 linker-VHH-human CH1-IgG4 hinge linker-IgG4 Fc chimeric Ab: tetraivalent; (J) VHH-human IgG4 hinge linker-IgG4 Fc-llama long linker-VHH chimeric Ab: tetraivalent. Thus, “(G4S)3” is a 15-peptide linker (amino acid sequence: GGGSGGGSGGGGS). “IgG4-hinge linker” means the hinge region between the human Fab and human IgG4 Fc (amino acid sequence: PPCSPAPEFLGGPS). “Llama long linker” means the hinge region between the llama Fab and llama IgG2 Fc (amino acid sequence: EPKIPQPQPKPQPQPQPKPQPKPEPECTCPKCP).

100-fold (Table 1). Peptide linker was used to form tandem bivalent or multivalent Abs for affinity increase. For instance, the GlyGlyGlySerGlyGlyGlySerGlyGlyGlySer-(G4S)3 is the most common polypeptide linker with joint flexibility (Fig. 3C,D,I)<sup>8,10</sup>. “AlaAlaAla” is a representative of rigid peptides<sup>11</sup> and the hinge region between Fab and Fc of the Ab is also considered a linker.

The principle of the first strategy is simple. Tandem VHHs were expressed in procaryotic or eukaryotic expression systems. The protein bands in lanes 1, 2, and 6 illustrated the intercomparison among the molecular weights (Fig. 4: lane 1, VHH: 17 kDa; lane 2, VHH-VHH: 35 kDa; lane 6, VHH-VHH-VHH: 51 kDa), and the medium supernatants and flow-through components are shown in lanes 4 and 5. The affinities between VHH-VHH (Fig. 2C) or VHH-VHH-VHH and PCSK9 are ~4 or 100-fold than that of VHH-B11 (Table 1).

The second type of IgG-like Abs mediate the augmentation of the biological half-life period because of Fc fusion that is expressed in the eukaryotic expression system, including *P. pastoris* and HEK293F cells. Five configurations were designed (Fig. 3F–J). Herein, VHH-Fc (Figs. 3F, 4, S6 lane 7: 75 kDa) has the hinge linker of the human IgG4 type, and the affinity increased ten-fold (VHH-Fc vs. VHH) (Fig. 2B, Table 1). VHH-CH1-Fc (Fig. 3G) is a recombinant heavy chain and has a very low expression due to lack of stability.

Name	<i>E. coli</i>	<i>P. pastoris</i>	HEK293F	Affinity	Linker1; linker2; linker3	Relative expression level	Configuration diagram	Gel photo
VHH	✓	✓	✓	~ 4.4 nM	NA	High	Figure 3B	Figure 4, lane1
VHH-VHH		✓	✓	~ 1.7 nM	(G4S)3	High	Figure 3C	Figure 4, lane2
VHH-VHH-VHH			✓	~ 0.01 nM	(G4S)3; (G4S)3	High	Figure 3D	Figure 4, lane6
VHH-Fc		✓	✓	~ 0.6 nM	IgG4-hinge linker	High	Figure 3F	Figure 4, lane7
VHH-CH1-Fc			✓	NA	NA; IgG4-hinge linker	Very low	Figure 3G	NA
VHH-VHH-Fc			✓	~ 0.8 nM	Llama long linker; IgG4-hinge linker	Medium	Figure 3H	Figure 4, lane8
VHH-VHH-CH1-Fc			✓	NA	(G4S)3; NA; hIgG4-hinge linker	Low	Figure 3I	Figure 4, lane9
VHH-Fc-VHH			✓	~ 0.6 nM	IgG4-hinge linker; Llama long linker	Medium	Figure 3J	Figure 4, lane10

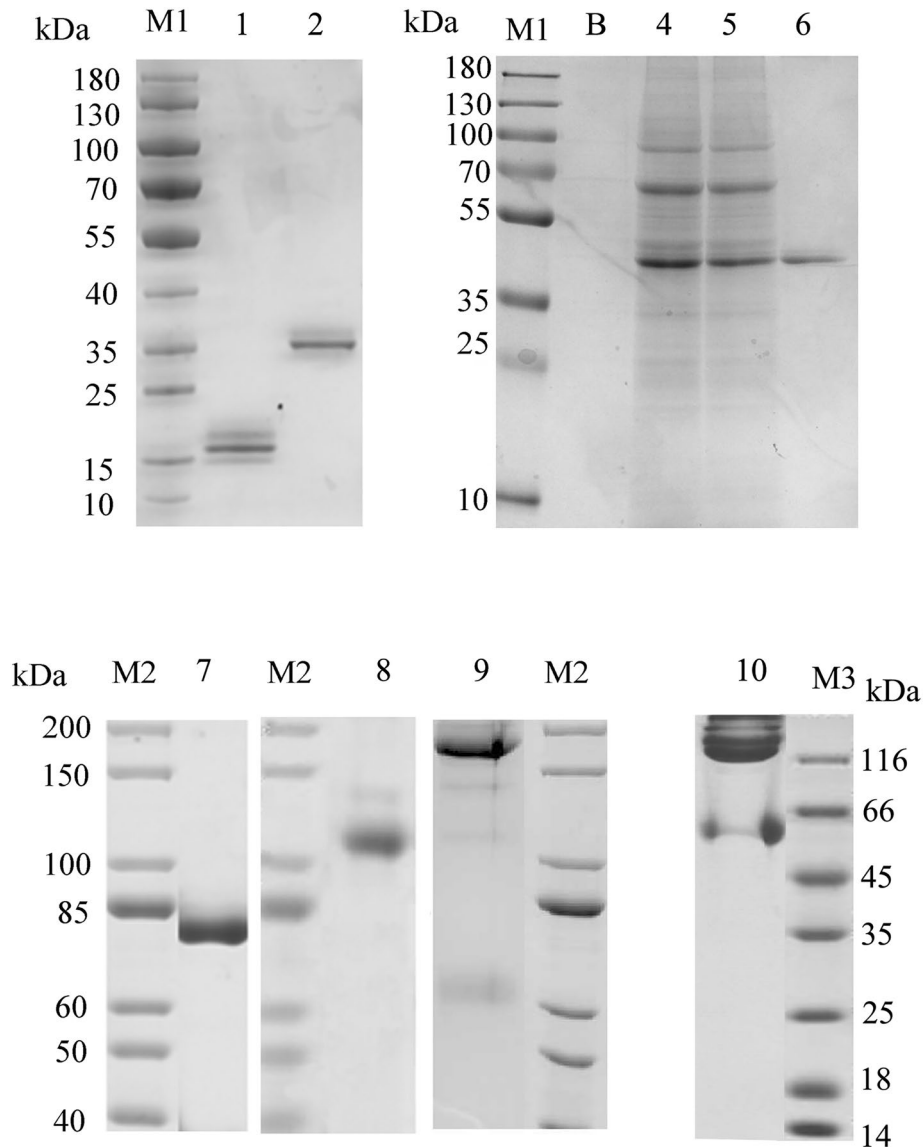
**Table 1.** Intercomparison of multiple VHH-B11-derived conformations. Note: “VHH” represents single domain Ab VHH-B11. “✓” indicates that the Ab is expressed in this expression system. “-” represents the linker between the domains (VHH, CH1 or Fc). “(G4S)3” is a 15-peptide linker (amino acid sequence: GGGGSGGGSGGGGS). “IgG4-hinge linker” indicates the hinge region between human Fab and human IgG4 Fc (amino acid sequence: PPCSPCAPEFLGGPS); “Llama long linker” represents the hinge region between llama Fab and llama IgG2 Fc (amino acid sequence: EPKIPQPQPKPQPQPQPKPQKPEPECTCKPCP). Relative expression level represents the expression status of each Ab in HEK293F cell expression system under same conditions. “High” means the purified Ab expression level > 20 mg/mL; “medium” means ~ 10 mg/mL; “low” means < 1 mg/mL; “very low” indicates that the Ab could be detected but cannot be purified easily; “NA” is not available.

Next, three tetravalent Abs were designed. VHH-VHH-Fc (Figs. 3H, 4, S7 lane 8: ~ 110 kDa) has two hinge linkers; the llama long hinge peptide of the IgG2 type was used to link bivalent VHHs, and the hinge linker of the human IgG4 type was used to link human Fc. The Fc region of this tetravalent conformation is fully exposed. Also, its affinity increased ten-fold (VHH-VHH-Fc vs. VHH) (Fig. 2D, Table 1). As shown in Fig. 3I, VHH-VHH-CH1-Fc configuration has a low expression, and the protein band is ~ 155 kDa (Figs. 4, S8: lane 9). The last tetravalent Ab is VHH-Fc-VHH configuration (Figs. 3J, 4, S9: lane 10). The Fc region of this conformation acts as a linker. Similarly, its affinity to PCSK9 also increased ten-fold. A previous study reported an increased functionality of Fc-fused configuration<sup>10</sup>. Hence, the tetravalent Ab format VHH-VHH-Fc is relatively the best conformation based on the appropriate molecular weight, increased affinity, Fc exposure and medium expression level.

**Humanization and identification of key amino acids.** According to previous reports<sup>16</sup>, we summarized 17 hot spot amino acids of humanization in the three framework regions (FR 1–3) of VHH. To improve the degree of humanization of VHH-B11 (Fig. 5: lane1, we designed humanized VHH-B11 (Hu-B11) with all the 17 humanized amino acid mutation sites (Table 2). Surprisingly, Hu-B11 (Fig. 5 and Fig. S12: lane5) has a normal expression level similar to the wild-type VHH-B11 (~ 20 mg/L), but the affinity decreased by nine-fold such that Hu-B11 has an affinity interaction with hPCSK9 (4.4 nM vs. 39.7 nM) (Fig. 2A vs. Fig. S14A). However, each FR region has five or six humanization sites, and Hu-B11 has a high humanized degree; thus, affinity maturation is needed in the future. If not, it is highly likely that unsatisfactory pharmacodynamic evaluation effect will be obtained.

In order to identify which amino acid position plays a critical role, we designed and expressed another nine VHH-B11 humanized mutants named by VHH-Z1-Z9 in *P. pastoris* (~ 15 kDa, Fig. 5, Figs. S10–S13 and Table 2). Most of these differed from each other by only one amino acid. For instance, VHH-Z1 (Fig. 5 and Figs. S11: lane 2, S14C) and VHH-Z6, VHH-Z2 (Fig. 5 and Fig. S11: lane 3) and VHH-Z5 (Fig. 5 and Figs. S13: lane 7, S14D), and VHH-Z3 (Fig. 5 and Figs. S12: lane 4, S14B) and VHH-Z4 (Fig. 5 and Fig. S13: lane 6) all have only one amino acid difference between each other. Among these, Hu-B11, VHH-Z3 and VHH-Z5 are expressed normally (~ 20 mg/L) in yeast, but the others have very low expression levels (< 1 mg/L), and VHH-Z9 is not expressed. There is only a single amino acid difference that can cause high or low expression and affinity (Fig. 5, Figs. S10–S13 and Table 2), as well as failure. Also, it could be speculated that these amino acid positions play critical roles. For instance, compared to the normal expression of VHH-Z1 (80Pro) (Fig. 5 and Figs. S11: lane 2, S14C), one amino acid difference (80Pro vs. 80Arg) in FR3 region causes low expression of VHH-Z6 (80Arg). Another replacement of the 87Val in the FR3 region of VHH-Z5 (normal expression, Fig. 5 and Figs. S13: lane 7, S14D) also led to a low expression and affinity of VHH-Z2 (87Lys). Similarly, one amino acid difference in FR2 region led to normal expression and no decline in the affinity of VHH-Z3 (54Ala) (Fig. 5 and Figs. S12: lane 4, S14B) but the very low expression and affinity of VHH-Z4 (54Ser). This finding suggested that these amino acid positions (no. 54, no. 80, and no. 87) may be crucial to the humanization of VHH-B11 protein.

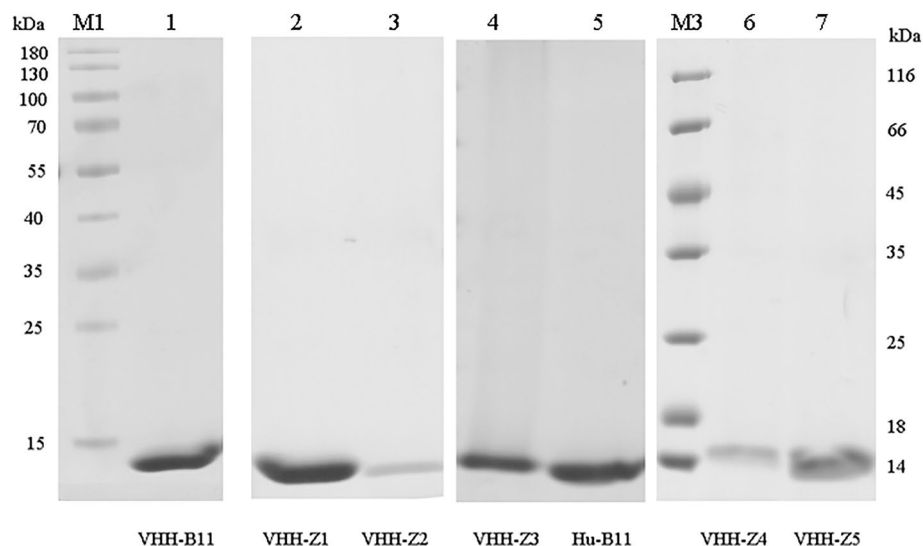
On the other hand, the first five amino acids, “QVQLQ” of VHH-B11 were not indispensable because the protein expression, structure, and avidity of VHH-B11 were not affected without these initial amino acids (data not shown). In this case, VHH-Z9 expression failure could be attributed to the replacement of 83Ser in the FR3



**Figure 4.** The SDS-PAGE gel of multiple VHH-B11 derived antibodies. The SDS-PAGE gel of multiple VHH-B11 derived antibodies. M1/M2/M3: the protein markers (Fermentas, USA). Lane 1: The VHH-B11 expressed by *P. pastoris* X33 (*pPICZα*); lane 2: The bivalent VHH-(G4S)<sup>3</sup>-VHH expressed by *Pichia pastoris* X33 (*pPICZα*); lane B: blank control; lane 4: the HEK293F culture medium superposition of the trivalent antibody VHH-(G4S)<sup>3</sup>-VHH-(G4S)<sup>3</sup>-VHH; lane 5: the Ni<sup>+</sup> affinity column flow-through of the trivalent antibody VHH-(G4S)<sup>3</sup>-VHH-(G4S)<sup>3</sup>-VHH; lane 6: the purified trivalent antibody VHH-(G4S)<sup>3</sup>-VHH-(G4S)<sup>3</sup>-VHH; lane 7: Llama VHH-human IgG4 hinge linker-IgG4 Fc chimeric HcAb (VHH-Fc); lane 8: the tetraivalent VHH-llama long linker-VHH-human IgG4 hinge linker-IgG4 Fc (VHH-VHH-Fc); lane 9: the tetraivalent VHH-(G4S)<sup>3</sup>-VHH-human CH1-IgG4 hinge linker-IgG4 Fc chimeric Ab (VHH-VHH-CH1-Fc); lane 10: the tetraivalent VHH-human IgG4 hinge linker-IgG4 Fc-llama long linker-VHH chimeric Ab (VHH-Fc-VHH).

region of VHH-Z3 (83Ala, normal expression, ~20 mg/L), indicating that 83Ala may also affect the normal formation of VHH-B11 Ab structure. When disregarding the initial five amino acids “QVQLQ”, VHH-Z2 and VHH-Z9 have only one amino acid difference at no. 54 (Ser vs. Ala). Among these, VHH-Z2 (Fig. 5 and Fig. S11: lane 3) has some expression, and Z9 has no expression, which further emphasizes the importance of no. 54 amino acid position on protein expression. It is worth noting that due to a poor fit, only three concentration gradients were showed the kinetic binding of VHH-Z1 and Z5 to the hPCSK9. (Fig. S14C,D). As for other humanized VHHs, they have poor or even no binding to PCSK9, so they are not shown.

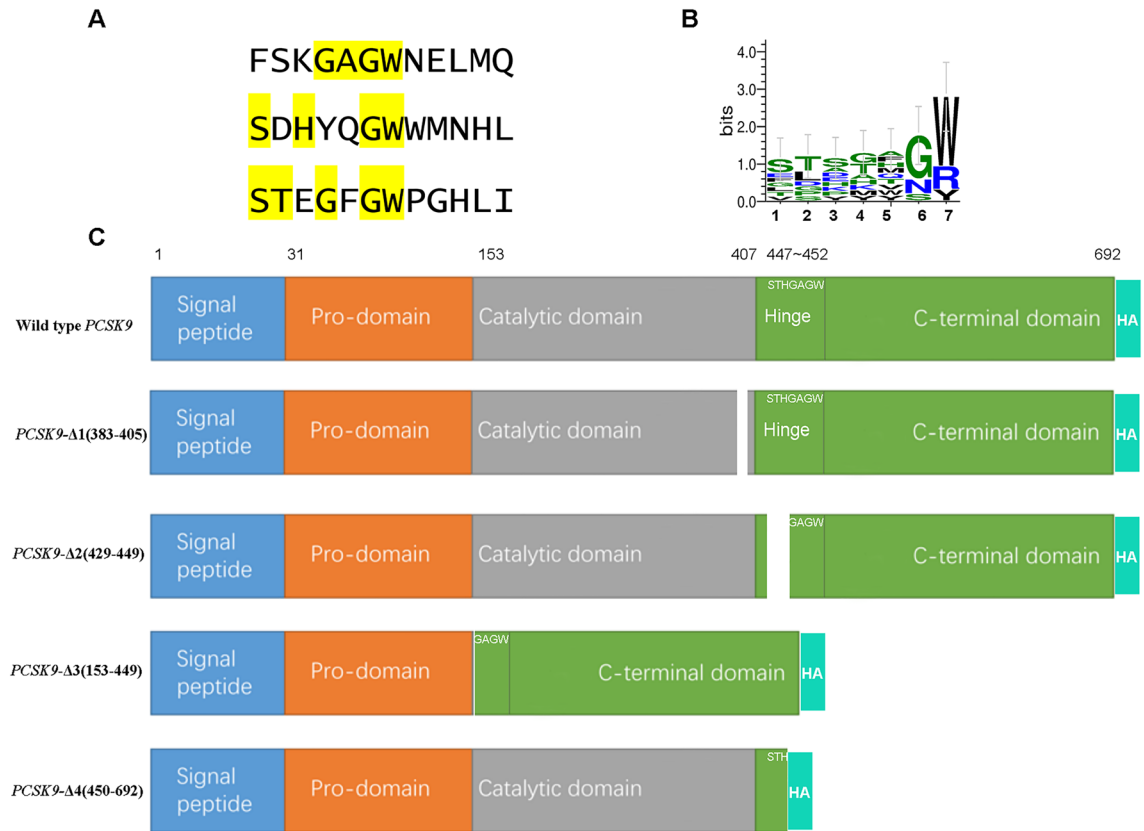
**Epitope speculation and determination.** In a previous study, we conducted SPR and sandwich ELISA (enzyme linked immunosorbent assay) assays and confirmed that the various PCSK9-binding epitope on VHH-B11 has different responses against Evolocumab<sup>7</sup>. However, the specific interaction epitope is yet unclear. To



**Figure 5.** The SDS-PAGE gel of the VHH-B11 humanized antibodies. The SDS-PAGE gel of multiple VHH-B11 humanized antibodies expressed by *P. pastoris* X33. M1/M3: the protein markers (Fermentas, USA). Lane 1: Wild type VHH-B11; lane 2: VHH-Z1; lane 3: VHH-Z2; lane 4: VHH-Z3; lane 5: Hu-B11; lane 6: VHH-Z4; lane 7: VHH-Z5.

Ab-ID	FR1	FR2	FR3	Hu-rate	K <sub>D</sub> (nM)
VHH-B11	QVQLQESGGGSVQAAGGSLRLSCTVS	WFRQAPGKEHEGVA RFTISQDNAKNTVYLQMNSLKPEDTAMYYCAV		0%	~5.013
VHH-Z1	EVQLLESGGGLVQPGGSLRLSCTVS	WFRQAPGKEHEGVS RFTISRDN SKNTLYLQMNSLRAEDTAVYYCAV		64.70%	~7.994
VHH-Z2	EVQLLESGGGLVQPGGSLRLSCAAS	WFRQAPGKEHEGVS RFTISQDNSKNTLYLQMNSLRAEDTAVYYCAV		70.59%	> 1000
VHH-Z3	EVQLLESGGGLVQPGGSLRLSCAAS	WFRQAPGKEHEGVA RFTISQDNAKNTLYLQMNSLRAEDTAVYYCAV		58.82%	~3.219
VHH-Z4	EVQLLESGGGLVQPGGSLRLSCAAS	WFRQAPGKEHEGVS RFTISQDNAKNTLYLQMNSLRAEDTAVYYCAV		64.70%	> 1000
VHH-Z5	EVQLLESGGGLVQPGGSLRLSCAAS	WFRQAPGKEHEGVS RFTISQDNSKNTVYLQMNSLRAEDTAVYYCAV		64.70%	~11.28
VHH-Z6	EVQLLESGGGLVQPGGSLRLSCTVS	WFRQAPGKEHEGVS RFTISQDNSKNTLYLQMNSLRAEDTAVYYCAV		64.70%	NA
VHH-Z7	EVQLLESGGGLVQPGGSLRLSCTVS	WFRQAPGKEHEGVS RFTISQDNAKNTLYLQMNSLRAEDTAVYYCAV		52.94%	NA
VHH-Z8	EVQLLESGGGLVQPGGSLRLSCAAS	WFRQAPGKEHEGVS RFTISQDNAKNTLYLQMNSLRAEDTAVYYCAV		64.70%	NA
VHH-Z9	QVQLVESGGGLVQPGGSLRLSCAAS	WFRQAPGKEHEGVA RFTISQDNSKNTLYLQMNSLRAEDTAVYYCAV		58.82%	NA
Hu-B11	EVQLLESGGGLVQPGGSLRLSCAAS	WVRQAPGKGLEWVS RFTISRDN SKNTLYLQMNSLRAEDTAVYYCAV		100%	~39.7

**Table 2.** VHH-B11 Ab and its humanized mutants. Ab ID is the name of VHH-B11 and each mutant. VHH-B11 is the wild type. The blue letters represent the locations of 17 “hotspot” humanization amino acids. Z1–Z9 are nine VHH-B11 humanized mutants, and each mutates at several of the hotspots. The mutated amino acids are labeled in red. Hu-B11 is a highly humanized mutant that mutates at all 17 hotspots. FR1–3 represents the three frameworks of VHH-B11. Hu-rate is short for the relative humanized rate (%). In this study, 17 amino acid mutations indicate 100% humanization rate. The Hu-rate is equal to the number of mutated amino acids divided by 17. K<sub>D</sub>(M) is the affinity between VHH-B11 mutant and hPCSK9 and is measured by the SPR technology at 25 °C. “NA” represents not available.



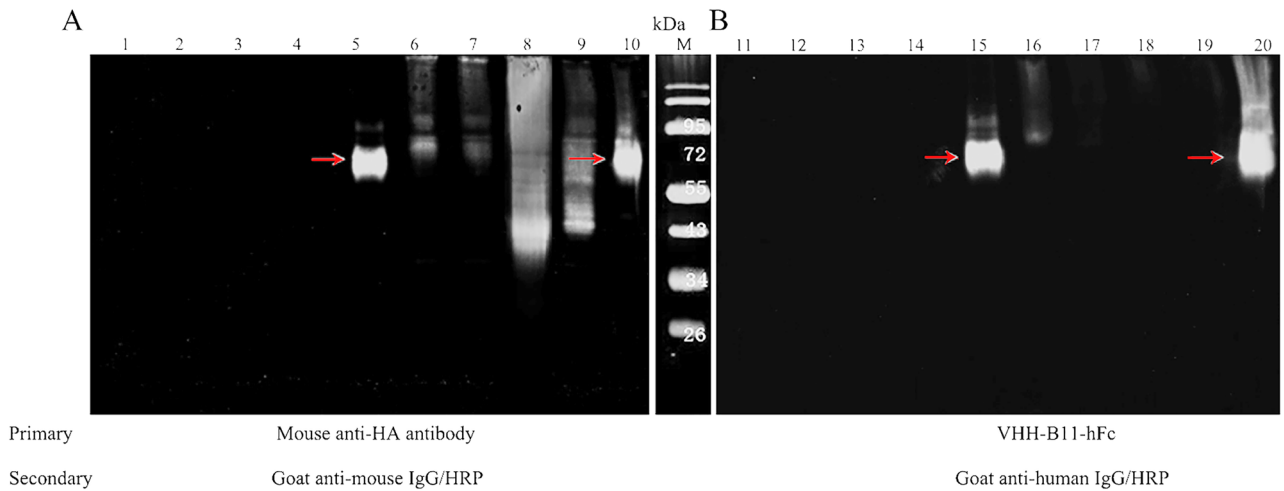
**Figure 6.** Epitope speculation and PCSK9 deletants design. **(A)** Epitope speculation. Many polypeptides with similar “GW” motif characteristics (for example,  $\times S \times GW$ ,  $\times T \times GW$ , or  $\times H \times GW$ .  $\times$  represents any amino acid) were identified from the random library (Table S2). Three peptides were listed as an illustration. **(B)** PCSK9 deletants design. Placing “GW” in the same amino acid position facilitated the identification of the possible common features (“STSGAGW”) of these short peptides after input in the Table S2 peptides by Weblogo 3 analysis tool. **(C)** Wild-type hPCSK9 and four hPCSK9-deletants lacked some modules: PCSK9-Δ1 (383–405), PCSK9-Δ2 (429–449), PCSK9-Δ3 (153–449), and PCSK9-Δ4 (450–692). hPCSK9 is composed of the signal peptide (amino acids 1–30), pro-domain (31–152), catalytic domain (153–407), hinge region (408–452), and C-terminal region (453–692). HA tag was fused for purification easily.

determine the interaction sites, we employed a random peptide library, phage display, and sequencing technology and characterized the amino acid of the binding epitope<sup>17,18</sup>.

Sequencing revealed 321 Ab-binding peptides from groups A and B (Table S1). Further analyses showed similar motif characteristics ( $\times S \times GW$ ,  $\times T \times GW$  or  $\times H \times GW$ ) among several peptides; “ $\times$ ” represents one or more amino acids (Table S2). Three peptides are listed in Fig. 6A. If “GW” or “G” is placed in the same amino acid position, Weblogo 3 analysis tool identified the most common features (“STSGAGW”) of these short peptides after input of the peptides listed in Table S2 (Fig. 6B). Interestingly, the hinge region of hPCSK9, 447–452 amino acids (STHGAGW) met this characteristic. To verify this speculation, we designed and expressed wild-type hPCSK9 and four hPCSK9-deletants lacking some modules: PCSK9-Δ1(383–405), PCSK9-Δ2(429–449), PCSK9-Δ3(153–449), and PCSK9-Δ4(450–692) in HEK293F (Fig. 6C). The numbers in parentheses represent the deletion range of amino acids. For instance, PCSK9-Δ1(383–405) represents a PCSK9 deletant lacking 383–405 amino acids (“SGTSQAAAHVAGIAAMMLSAEPE”). The PCSK9-Δ1 deletant was regarded as the irrelevant contrast, and the other three deletants were experimental groups that the predicted epitope was broken by removing different modules.

An immunoblot assay was performed to verify this prediction. PCSK9 and its deletants were loaded in the following sequence in groups of five samples (supernatants and cell precipitation lysate): PCSK9-Δ1(383–405), PCSK9-Δ2(429–449), PCSK9-Δ3(153–449), PCSK9-Δ4(450–692), and wild-type PCSK9. As shown in Fig. 7A, anti-HA Ab was used as a primary Ab to assess whether PCSK9 was expressed intracellularly (Fig. 7A, lanes 1–5, supernatants) and extracellularly (Fig. 7A, lanes 6–10, cell lysate), while B11-Fc Ab was used as primary Ab to verify whether PCSK9 and the deletants could self-interact (Fig. 7B, lanes 11–15, supernatants and Fig. 7B, lanes 16–20, cell lysate). Anti-HA primary Ab group showed that wild-type PCSK9 was expressed extracellularly as well as intracellularly (~66 kDa & ~72 kDa, Fig. 7A, lanes 5 and 10), while all the PCSK9 deletants expressed only intracellularly (Fig. 7A, lanes 6–9). Furthermore, B11-Fc primary Ab group showed that wild-type PCSK9 (Fig. 7B, lanes 15 and 20) and the PCSK9-Δ1 deletant (the irrelevant contrast, Fig. 7B, lane 16) could still be detected by B11-Fc because of their intact epitope, but the other three deletants (PCSK9 Δ2-Δ4, Fig. 7B lanes





**Figure 7.** Immunoblot assay and epitope verification. Immunoblot assay was performed to verify the predicted epitope between PCSK9 and VHH-B11-Fc. PCSK9 and its deletants were loaded in the following sequence in groups of five samples (lanes 1–5 and 11–15: supernatants; lanes 6–10 and 16–20: cell precipitation lysate): PCSK9- $\Delta$ 1 (383–405), PCSK9- $\Delta$ 2 (429–449), PCSK9- $\Delta$ 3 (153–449), PCSK9- $\Delta$ 4 (450–692), and wild-type PCSK9 (red arrow). **(A)** The anti-HA Ab was used as primary Ab to check whether PCSK9 was expressed intracellularly (lanes 1–5) and extracellularly (lanes 6–10). The HRP-tagged goat anti-mouse IgG was used as the secondary antibody. **(B)** The B11-Fc Ab was used as a primary Ab to check whether PCSK9 and the deletants could still interact specifically with B11-Fc (lanes 11–20). The HRP-tagged goat anti-human IgG was used as the secondary antibody.

17–19) not. These findings suggested that the broken epitope “STHGAGW” (no. 447–452) of the PCSK9  $\Delta$ 2- $\Delta$ 4 deletants would lead to the failure of detection by the specific Ab B11-Fc, deeming that this linear epitope is the interacting site between B11-Fc and PCSK9.

## Discussion

The single-domain Ab of camelidae is fragmentary and a therapeutic Ab. Due to the small size (~15 kDa) and the fast metabolism (~30 min) in the organisms<sup>19–21</sup>, appropriate modifications are essential to improve the pharmacokinetic and pharmacodynamic performances of single domain Abs<sup>22,23</sup> that can be executed via VHH tandem (for example Cablivi: VHH-AlaAlaAla-VHH)<sup>11</sup>. However, human IgG Fc fusion is the general strategy for therapeutic VHH Abs<sup>13,24</sup>. For instance, KN035 Ab has VHH-Fc structure approved as an orphan drug by the US FDA in 2020<sup>12</sup>. Nevertheless, compared to conventional Abs, VHH tandem and/or Fc fusion is not optimal as it inevitably results in larger molecular weight, which in turn might alter the molecular structure and affect expression level and stability. Notably, human CH1 addition might also alter the structure of the camelid heavy chain Ab. For instance, VHH-Fc has a normal expression level, but VHH-CH1-Fc has a low expression level under the same conditions. Another example is VHH-VHH-Fc, with a normal expression level, but the quadrivalent VHH-VHH-CH1-Fc has a low expression level. These phenomena could be attributed to the lack of human CL1 domain coordination for CH1. Currently, researchers are seeking to maintain a balance of molecular weight, stability, pharmacodynamic efficacy and expression difficulty of the therapeutic Abs.

Ab humanization is crucial to reduce the immunogenicity of antigen (Ag)<sup>25</sup>. Low immunogenicity can reduce the occurrence of side effects<sup>26</sup>. In a previous study, we reported that human IgG4 Fc was fused to the llama VHH to form chimeras<sup>7</sup>. In the present study, we carried out 17 “hot spot” amino acid mutations to make it similar to the variable region of the human Ab gene<sup>16</sup>. However, the effect of immunogenicity reduction is to be evaluated in future studies<sup>27</sup>.

Epitope identification in the interaction between Ag and Ab is conducive to clarifying the mechanism of efficacy of the Ab<sup>28</sup>. Because of its unique advantages, camelid VHH Ab utilizes the conformation or charge complementarity for selective Ag binding, which mimics the conserved intracellular protein–protein interaction<sup>29</sup>. This characteristic is rarely exhibited by conventional IgG Abs. Reportedly, the interaction epitopes between the approved Evolocumab and PCSK9 are in the LDLR-binding site of the catalytic region<sup>30</sup>. In the present study, phage screening, random peptide library formation, and sequencing results proved that the VHH-B11-Fc binding epitope is in the hinge region of PCSK9. This phenomenon is rarely reported for anti-PCSK9 Ab. This method is economical, while the disadvantage is that only linear epitopes or linear parts of conformational epitopes can be identified. Thus, a better verification method, such as X-ray crystal diffraction, is required, especially for the conformational epitope. To obtain an ideal crystal of the Ag-Ab complex, some amino acid mutations on the molecular surface of protein<sup>31</sup> are required; for example, the hinge region mutation, which in turn affects the binding of VHH-B11-Fc and PCSK9. Thus, we opted for the random peptide library approach.

In this study, we developed and characterized a camelid VHH-B11 Ab. The druggability of VHH-B11 has been enhanced adequately. Both VHH tandem and Fc fusion have been used to form polyvalent Ab, such as tetravalent VHH-VHH-Fc. It has a high affinity to hPCSK9 (0.1 nM). Moreover, it has a relatively high expression

level (> 10 mg/L). Combined with the identified epitopes and good humanization results, this study provides experimental evidence for improved druggability of camelid-derived VHH Ab.

## Methods

**Monovalent nanobody expressed in the *E. coli* system.** VHH-B11 was expressed in three systems (*E. coli*, *P. pastoris* and HEK293F mammalian cells) to investigate the difference in the nanobodies. The BL21 *E. coli*, TG1 *E. coli*, and the *pET28a* vector were preserved in our laboratory. HB2151 *E. coli* and *pMECS* expression were obtained as a kind gift from the University of Brussels, Belgium<sup>32</sup>. After culture for 4–6 h in LB or TB medium (0.72 M  $\text{KH}_2\text{PO}_4$ , 170 mM  $\text{K}_2\text{HPO}_4$ , 1.2% peptone, 2.4% yeast extract, 0.4% glycerol, 100  $\mu\text{g}/\text{mL}$  ampicillin, 0.1% glucose, and 2 mM  $\text{MgCl}_2$ ), 1 mM IPTG was added when  $\text{OD}_{600}$  was 0.8–1.0 to induce VHH expression overnight. The cell pellet was collected by centrifugation at 10,000 rpm for 10 min. Ultrasonication was used to release the intracellular expression (BL21 *E. coli*). The TES buffer (0.2 M Tris–HCl pH 8.0, 0.5 mM EDTA, 0.5 M sucrose) was used for TG1 and HB2151 bacterial lysis<sup>32</sup>. The solvent was replaced with phosphate buffer for the immunoaffinity purification using ultrafiltration (3 kDa, Millipore, USA) for good purification results. The specific experimental procedures were carried out according to the manufacturer's instructions.

**Ab expression in *P. pastoris* system.** In this study, *P. pastoris* X33 and *pPICZ $\alpha$*  plasmids (Life Invitrogen, USA) were used to assess the secreted level of monovalent and partial multivalent VHH Abs, which improves the expression efficiency and facilitates purification without cell lysis. The recombinant *pPICZ $\alpha$ -VHH* vectors were linearized by single restriction endonuclease cleavage by *Sall* and purified by agarose gel electrophoresis. Subsequently, after electroporation, the bacterial suspension was spread on YPD solid medium plate (1% yeast extract, 2% peptone, 2% dextrose glucose, 2% agar). Then, zeocin was added in gradient (100, 200, 400, and 800 mg/mL) to the YPD medium for screening the stable expression of yeast strains. BMGY medium (1% yeast extract, 2% peptone, 3 g/L  $\text{K}_2\text{HPO}_4$ , 11.8 g/L  $\text{KH}_2\text{PO}_4$ , 0.02% biotin, 1% glycerol, and 1  $\times$  yeast nitrogen base source medium) was used to expand and propagate *P. pastoris* for 3 days. Methanol was added every 24 h to induce the Ab expression in the new BMMY medium (1% yeast extract, 2% peptone, 3 g/L  $\text{K}_2\text{HPO}_4$ , 11.8 g/L  $\text{KH}_2\text{PO}_4$ , 0.02% biotin, 0.5% methanol, and 1  $\times$  yeast nitrogen base source medium). After 5 days of inducing expression, the cell supernatant was collected by centrifugation at 12,000 rpm for 10 min. If the volume was large (> 2 L), ammonium sulfate precipitation method is recommended to precipitate the protein. For adequate immunoaffinity purification results, the protein precipitate needs to be solubilized in the equilibration buffer (such as phosphate buffer).

**Ab expression in HEK293F mammalian cell system.** To express Abs with advanced and complex structures, HEK293F mammalian cell (Sinobiological, China) expression system was selected for VHH-derived Ab secretory production. For easy purification, a special serum-free medium (SMM 293-TII, Sinobiological, China) was used to support the growth and transfection of HEK293F cells. The *pCDNA3.4* (purchased from Life Invitrogen) was used for the transient transfection (Sinofection, Sinobiological). Briefly, the freshly passaged cells were seeded at a density of  $0.5 \times 10^6/\text{mL}$  and incubated at 37 °C (175 rpm, 5%  $\text{CO}_2$ ) overnight. At the time of transfection, the suspension cells were in the log growth phase. Next, the endotoxin-free plasmid *pCDNA3.4* was diluted with 0.15 M NaCl solution at 20  $\mu\text{g}/\text{mL}$  (optimization recommended), and the transfection reagent (DNA: reagent = 1: 4) was diluted in the same volume. Then, the diluted DNA was combined with the diluted transfection reagent, incubated for 20 min at room temperature, and added drop-wise to the cells, followed by gentle shaking to ensure even distribution. Then, the cells were incubated at 37 °C in a  $\text{CO}_2$  shaker (175 rpm) for 48 h, and then the cell culture supplement (3.5%) was added every 48 h.

After 5 days of shaking cultivation, the cell supernatant was collected by centrifugation at 12,000 rpm for 10 min. The specific experimental procedures were carried out according to the manufacturer's instructions.

**Purification.** After culture and expression, the AKTA pure25 (GE Healthcare, USA) was used for automated purification after setting up the program. The 5 mL Hisrap FF purification column (GE Healthcare) was used for His-tag VHH protein purification. The 5 mL Protein A purification column (GE Healthcare) was used to purify the Fc-fused Abs. After purification, the solvent of the protein solution should be replaced with the phosphate buffer in the 15-mL ultrafiltration tube (Millipore, USA). The specific experimental procedures were according to the manufacturer's instructions.

**SPR assay.** SPR technology was adopted to determinate the affinities between the VHH-based Abs and hPCSK9 antigen. First, recombinant hPCSK9 (his-tagged, Sinobiological, China) was diluted to 10  $\mu\text{g}/\text{mL}$  with 10 mM sodium acetate (pH = 4.0). Then hPCSK9 protein was captured on the flow cell 2 of the CM5 chip (GE Healthcare, USA, Cat# 29104988) at 790 response unit (RU), with flow cell 1 as the blank. All the SPR assays were executed under 25 °C using the Biacore T200 (GE Healthcare, USA). The binding and dissociation time was set at 120 s and 300 s (or 350 s) respectively. A five (or three points), double fold dilution of different molar concentrations of these VHH-based proteins was injected and the sensorgrams were fitted with a floating  $R_{\text{max}}$  using the built-in evaluation software of the Biacore T200. The glycine (pH = 2.3) was used as regeneration buffer. The affinity ( $K_D$ , nM) is calculated as followings:  $K_D = k_{\text{off}}/k_{\text{on}}$ .  $k_{\text{off}}$  (1/s) is the dissociation constant and  $k_{\text{on}}$  (1/Ms) is the binding constant.

**Phage random peptide library screening.** Ph.D.-12 phage random peptide library (New England Biolabs) sequencing technologies were utilized to identify the interaction peptides<sup>18</sup>. This method was based on

expressing recombinant random peptides on the surface of the phage and selecting the potential interacting peptides bound to the solid phase. The initial experimental steps are similar to the phage-ELISA procedures<sup>18,32,33</sup>. Briefly, ELISA wells were coated with the VHH-B11-Fc Ab in duplicate at 4 °C overnight and blocked with 3% bovine serum albumin (BSA) after five washes with 1 × PBST, termed group A and group B, respectively. A non-coating well was set up as the control group. After blocking and washing, Ph.D.-12 phage random peptide library was added, and then the binding Ab-phage complex was reserved after five washes with 1 × PBST washing. With the complex washed off with triethylamine, the sequencing library was constructed from the DNA obtained by high-temperature lysis, as described previously<sup>18</sup>. Essentially, the DNA library was generated by two consecutive rounds of extension PCR using Q5 Hot Start High-Fidelity DNA Polymerase (New England Biolabs). After quality control, the barcoded DNA library was sequenced on Novaseq 6000 under the paired-end 2 × 150 sequencing (PE150) strategy.

**Data processing, analysis and epitope speculation.** To acquire high-quality reads from raw data, Trimmomatic software was used to filter out the low-quality reads<sup>34</sup>. The 150-bp PE reads with the exact same insertion sequence were considered clean data for subsequent analysis. A pair of 8-bp barcodes and 36-bp variable DNA inserts were considered corresponding to the 12-amino acid polypeptide for single reads. This pair of 8-bp barcodes distinguished the VHH-B11-Fc Ab sample from the control. Next, the peptide number of the Ab was counted and normalized to the total peptide number. If one peptide appeared in our sample and not the control, the count was set as 1 in control. The number of these peptides was not calculated for the total control sample counts. In addition, the arithmetic mean of each peptide was calculated for the control and added into the matrix as the control column. The resulting matrix was defined as the normalized count matrix. Subsequently, the Ab-binding peptides were determined by comparing between the Ab and the control. Since no Ab was added to the control, all the peptides were considered noise. In order to identify the binding peptides for the downstream analyses, the cutoff value was determined as described previously. Only the eligible peptides were recognized as binding peptides, according to a previous report<sup>18</sup>.

**Epitope determination by Western blot.** We designed four hPCSK9 deletants for epitope determination and disrupted (or removed) the putative epitopes of three to assess if the binding between the VHH-B11-Fc Ab and PCSK9 was affected<sup>30</sup>. For the remaining deletant, we removed an unrelated epitope as an irrelevant control to verify whether this affects the Ab-hPCSK9 interaction. Moreover, the wild-type hPCSK9 served as the positive control. Also, the HEK293F mammalian cell system was selected for hPCSK9 expression. hPCSK9 and all its deletants were expressed with HA tag. A volume of 10 µL medium supernatant and 1 µL of 1 mL cell precipitation lysate were used for Western blot. After non-reducing PAGE, the proteins were electrotransferred to a polyvinylidene difluoride membrane (PVDF; Millipore) at 400 mA for 70 min. The membranes were then blocked with 3% BSA and probed with two primary Abs (mouse anti-HA Ab or the VHH-B11-Fc Ab) at room temperature for 2 h, followed by two corresponding secondary Abs (horseradish peroxidase (HRP)-tagged goat anti-mouse IgG or HRP-tagged goat anti-human IgG) at 37 °C for 1 h to determine the PCSK9 protein levels. The immunoreactive bands were detected using the enhanced chemiluminescence system (BioLiquor, Nanjing, China). Herein, the mouse anti-HA primary Ab was used to evaluate whether the four PCSK9 deletants were expressed extracellularly or intracellularly. The VHH-B11-Fc primary Ab was used to check whether the binding between the Ab and the PCSK9 deletants was affected.

**Ethics approval and consent to participate.** Permission for this study was granted by the Bioethics and Biological Safety Review Committee of BGI-Shenzhen (NO. FT 17064). We confirm that all experiments were conducted in accordance with relevant guidelines and regulations.

### Data availability

The data that supports the findings of this study are available in the supplementary material of this article.

Received: 8 February 2022; Accepted: 11 July 2022

Published online: 16 July 2022

### References

1. Cholesterol Treatment Trialists, C. Efficacy and safety of more intensive lowering of LDL cholesterol: A meta-analysis of data from 170 000 participants in 26 randomised trials. *The Lancet* **376**, 1670–1681. [https://doi.org/10.1016/S0140-6736\(10\)61350-5](https://doi.org/10.1016/S0140-6736(10)61350-5) (2010).
2. Kajinami, K. *et al.* Statin intolerance clinical guide 2018. *J. Atheroscler. Thromb.* **27**, 375–396. <https://doi.org/10.5551/jat.50948> (2020).
3. Flora, G. D. & Nayak, M. K. A brief review of cardiovascular diseases, associated risk factors and current treatment regimes. *Curr. Pharm. Des.* **25**, 4063–4084. <https://doi.org/10.2174/1381612825666190925163827> (2019).
4. Kloska, A., Malinowska, M., Gabig-Cimiaska, M. & Jakabkiewicz-Banecka, J. Lipids and lipid mediators associated with the risk and pathology of ischemic stroke. *Int. J. Mol. Sci.* **21**, 3618. <https://doi.org/10.3390/ijms21103618> (2020).
5. Scognamiglio, M., Costa, D., Sorriento, A. & Napoli, C. Current drugs and nutraceuticals for the treatment of patients with dyslipidemias. *Curr. Pharm. Des.* **25**, 85–95. <https://doi.org/10.2174/1381612825666190130101108> (2019).
6. Schmidt, A. F. *et al.* PCSK9 monoclonal antibodies for the primary and secondary prevention of cardiovascular disease. *Cochrane Database Syst. Rev.* **10**, CD011748–CD011748. <https://doi.org/10.1002/14651858.CD011748.pub3> (2020).
7. Li, X. *et al.* The novel llama-human chimeric antibody has potent effect in lowering LDL-c levels in hPCSK9 transgenic rats. *Clin. Transl. Med.* **9**, 16–16. <https://doi.org/10.1186/s40169-020-0265-2> (2020).
8. Xiang, Y. *et al.* Versatile and multivalent nanobodies efficiently neutralize SARS-CoV-2. *Science* **370**, 1479–1484. <https://doi.org/10.1126/science.abe4747> (2020).

9. De Greve, H., Virdi, V., Bakshi, S. & Depicker, A. Simplified monomeric VHH-Fc antibodies provide new opportunities for passive immunization. *Curr. Opin. Biotechnol.* **61**, 96–101. <https://doi.org/10.1016/j.copbio.2019.11.006> (2020).
10. Xu, J. *et al.* Nanobodies from camelid mice and llamas neutralize SARS-CoV-2 variants. *Nature* **595**, 278–282. <https://doi.org/10.1038/s41586-021-03676-z> (2021).
11. Duggan, S. Caplacizumab: First global approval. *Drugs* **78**, 1639–1642. <https://doi.org/10.1007/s40265-018-0989-0> (2018).
12. Zhang, F. *et al.* Structural basis of a novel PD-L1 nanobody for immune checkpoint blockade. *Cell Discov.* **3**, 17004–17004. <https://doi.org/10.1038/celldisc.2017.4> (2017).
13. Godakova, S. A. *et al.* Camelid VHHS fused to human fc fragments provide long term protection against botulinum neurotoxin A in mice. *Toxins* **11**, 464. <https://doi.org/10.3390/toxins11080464> (2019).
14. Rossotti, M. A., Bélanger, K., Henry, K. A. & Tanha, J. Immunogenicity and humanization of single-domain antibodies. *FEBS J.* <https://doi.org/10.1111/febs.15809> (2021).
15. Dyson, M. R. *et al.* Beyond affinity: Selection of antibody variants with optimal biophysical properties and reduced immunogenicity from mammalian display libraries. *MAbs* **12**(1), 335. <https://doi.org/10.1080/19420862.2020.1829335> (2020).
16. Vincke, C. C. *et al.* General strategy to humanize a camelid single-domain antibody and identification of a universal humanized nanobody scaffold. *J. Biol. Chem.* **284**, 3273–3284. <https://doi.org/10.1074/jbc.M806889200> (2009).
17. Li, Y. *et al.* Linear epitopes of SARS-CoV-2 spike protein elicit neutralizing antibodies in COVID-19 patients. *Cell. Mol. Immunol.* **17**, 1095–1097. <https://doi.org/10.1038/s41423-020-00523-5> (2020).
18. Qi, H. *et al.* Antibody binding epitope mapping (AbMap) of hundred antibodies in a single run. *Mol. Cell Proteomics* **20**, 100059–100059. <https://doi.org/10.1074/mcp.RA120.002314> (2021).
19. Coppieters, K. *et al.* Formatted anti-tumor necrosis factor  $\alpha$  VHH proteins derived from camelids show superior potency and targeting to inflamed joints in a murine model of collagen-induced arthritis. *Arthritis Rheum.* **54**, 1856–1866. <https://doi.org/10.1002/art.21827> (2006).
20. Peyron, I. *et al.* Camelid-derived single-chain antibodies in hemostasis: Mechanistic, diagnostic, and therapeutic applications. *Res. Pract. Thromb. Haemost.* **4**, 1087–1110. <https://doi.org/10.1002/rth2.12420> (2020).
21. Wesolowski, J. *et al.* Single domain antibodies: Promising experimental and therapeutic tools in infection and immunity. *Med. Microbiol. Immunol.* **198**, 157–174. <https://doi.org/10.1007/s00430-009-0116-7> (2009).
22. Holliger, P. & Hudson, P. J. Engineered antibody fragments and the rise of single domains. *Nat. Biotechnol.* **23**, 1126–1136. <https://doi.org/10.1038/nbt1142> (2005).
23. Asaadi, Y., Jouneghani, F. F., Janani, S. & Rahbarizadeh, F. A comprehensive comparison between camelid nanobodies and single chain variable fragments. *Biomark. Res.* **9**, 87–87. <https://doi.org/10.1186/s40364-021-00332-6> (2021).
24. Dong, J. *et al.* Development of humanized tri-specific nanobodies with potent neutralization for SARS-CoV-2. *Sci. Rep.* **10**, 17806–17806. <https://doi.org/10.1038/s41598-020-74761-y> (2020).
25. Safdari, Y., Farajnia, S., Asgharzadeh, M. & Khalili, M. Antibody humanization methods—a review and update. *Biotechnol. Genet. Eng. Rev.* **29**, 175–186. <https://doi.org/10.1080/02648725.2013.801235> (2013).
26. Nelson, A. L., Dhimolea, E. & Reichert, J. M. Development trends for human monoclonal antibody therapeutics. *Nat. Rev. Drug Discov.* **9**, 767–774. <https://doi.org/10.1038/nrd3229> (2010).
27. Liang, S. & Zhang, C. Prediction of immunogenicity for humanized and full human therapeutic antibodies. *PLoS ONE* **15**, e0238150–e0238150. <https://doi.org/10.1371/journal.pone.0238150> (2020).
28. Hua, C. K. *et al.* Computationally-driven identification of antibody epitopes. *Elife* **6**, e29023. <https://doi.org/10.7554/eLife.29023> (2017).
29. Li, Y. *et al.* Linear epitope landscape of the SARS-CoV-2 Spike protein constructed from 1,051 COVID-19 patients. *Cell Rep.* **34**, 108915–108915. <https://doi.org/10.1016/j.celrep.2021.108915> (2021).
30. Weider, E. *et al.* Proprotein convertase Subtilisin/Kexin Type 9 (PCSK9) single domain antibodies are potent inhibitors of low density lipoprotein receptor degradation. *J. Biol. Chem.* **291**, 16659–16671. <https://doi.org/10.1074/jbc.M116.717736> (2016).
31. Mizutani, H. *et al.* Systematic study on crystal-contact engineering of diphthine synthase: Influence of mutations at crystal-packing regions on X-ray diffraction quality. *Acta Crystallogr. Sect. D Biol. Crystallogr.* **64**, 1020–1033. <https://doi.org/10.1107/s0907444908023019> (2008).
32. Vincke, C. *et al.* in *Antibody Engineering: Methods and Protocols, Second Edition* (ed Patrick Chames) 145–176 (Humana Press, 2012).
33. Li, X. *et al.* Oral immunization with recombinant *Lactococcus lactis* delivering a multi-epitope antigen CTB-UE attenuates *Helicobacter pylori* infection in mice. *Pathogens Dis.* **72**, 78–86. <https://doi.org/10.1111/2049-632x.12173> (2014).
34. Bolger, A. M., Lohse, M. & Usadel, B. Trimmomatic: A flexible trimmer for Illumina sequence data. *Bioinformatics* **30**, 2114–2120. <https://doi.org/10.1093/bioinformatics/btu170> (2014).

## Acknowledgements

We thank the funding support from Science, Technology and Innovation Commission of Shenzhen Municipality under grant NO. JCY20170817145305022 and the Research Funding (Grant NO. K-Q2021020) support of Henan University of Urban Construction. We also thank Dr. Mingliang Ma and Prof. Shengce Tao from Shanghai Jiaotong University for their supports on phage random peptide library screening and sequencing.

## Author contributions

L.X.Y. carried out most of the experiments, performed most of the statistical analysis, and wrote the main manuscript text. G.X.Y. provided the amino acid mutation and humanization strategy. W.M.N. provided the basic data of Table 1 and Fig. 1. H.J. performed the statistical analysis and prepared Tables 1 and 2. L.X.Y., Y.N.B. and W.M.N. conceived of the study, and participated in its design and coordination. H.J. and Y.N.B. revised and improved the manuscript. All authors reviewed the manuscript.

## Competing interests

The authors declare no competing interests.

## Additional information

**Supplementary Information** The online version contains supplementary material available at <https://doi.org/10.1038/s41598-022-16453-3>.

**Correspondence** and requests for materials should be addressed to X.L. or N.Y.

**Reprints and permissions information** is available at [www.nature.com/reprints](http://www.nature.com/reprints).

**Publisher's note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2022