Principles of antibodies with ultralong complementarity-determining regions and picobodies

Marcel Passon, Stefaan De Smedt, Hristo L. Svilenov*

Faculty of Pharmaceutical Sciences, Ghent University, Ottergemsesteenweg 460, 9000 Gent

*To whom correspondence should be addressed:

hristo.svilenov@ugent.be

phone: +32 9 264 8075

ORCID: https://orcid.org/0000-0001-5863-9569

Abstract

In contrast to other species, cattle possess exceptional antibodies with ultra-long complementarity-determining regions (ulCDRs) that can consist of 40-70 amino acids. The bovine ulCDR is folded into a stalk and a disulfide-rich knob domain. The binding to the antigen is via the 3-6 kDa knob. There exists an immense sequence and structural diversity in the knob that enables binding to different antigens. Here we summarize the current knowledge of the ulCDR structure and provide an overview of the approaches to discover ulCDRs against novel antigens. Furthermore, we outline protein engineering approaches inspired by the natural ulCDRs. Finally, we discuss the enormous potential of using isolated bovine knobs, also named picobodies, as the smallest antigen-binding domains derived from natural antibodies.

1. Introduction

Antibodies are an essential part of the adaptive immune system in vertebrates. Conventional antibodies consist of heavy (HCs) and light chains (LCs). Each HC and LC has constant and variable domains[1]. The variable domains form the antigen-binding site via short hypervariable loops called complementarity-determining regions (CDRs). There are three CDRs in the HC and three CDRs in the LC. The CDR-H3 is the most variable.

The hypervariable nature of the CDRs is fundamental for antibody diversity and robust adaptive immune response. Diverse antigen-binding sites in antibodies are generated via the randomized assembly of variable (V), diversity (D) and joining (J) genes (known as V(D)J recombination) and somatic hypermutation (SHM)[2–4]. Interestingly, some species (e.g., chickens) also employ gene conversion to diversify their antibody repertoire[5,6].

The sequence and structural features of the CDRs are essential for the binding and physicochemical properties of antibodies[7]. Interestingly, in most species the CDRs are limited in length. For example, the most variable CDR-H3 has on average only 6 to 20 amino acids in humans[8,9]. In therapeutic antibodies, the average CDR-H3 length is even shorter compared to repertoire human antibodies[10]. In contrast, cows have a subset of antibodies with an ultra-long CDR-H3 (ulCDR-H3), consisting of about 40 to 70 amino acids[11–13].

The genetic origins of the ulCDRs are fascinating[14,15]. Bovines have only a limited number of functional V, D and J genes (12 IGHV, 16 IGHD and 4 IGHJ)[16,17] compared to 57 IGHV, 23 IGHD and 6 IGHJ gene segments in humans[18]. Due to this limited germline diversity, cows exploit other mechanisms to expand their antibody repertoire. For example, one of the bovine DH segments (D8-2) is 148 bp, four times longer than the longest human DH gene[14]. Interestingly, a specific VDJ combination (VH1-7, D8-2, J2-4)[19,20] and exclusive HC pairing with a subset of lambda LCs (V30 segment)[17,21] results in antibodies with ulCDR-H3s. Cattle also rely heavily on SHM, which already starts before antigen exposure, while in humans SHM is extensive only after contact with the antigen[17,19,22–24]. Remarkably, the SHM in the ulCDR-H3 leads to unusually high usage of cysteines that can form disulfide bonds contributing further to paratope diversification[16,25,26]. The bias towards cysteine usage in the ulCDR-H3 is programmed in the bovine germline – about 80 % of the residues in the ultra-long DH segment can be changed to cysteine via a single nucleotide mutation[24]. In addition, the ultra-long DH contains multiple RGYW hotspots (recognition motifs for the activation-induced cytidine deaminase which is an enzyme playing a key role in SHM)[24]. The ulCDR-H3s are abundant in cattle – they are found in about 10 % of the bovine antibody repertoire and in all five immunoglobulin isotypes[11–13,27].

Recent studies have demonstrated that the antibodies with ulCDRs hold enormous potential for therapeutic applications and rational engineering of new target-binding proteins, for example, as antiviral[28,29], anticancer[30,31] or immunomodulatory[32,33] agents. Here we review the exceptional features of antibodies with ulCDRs. We begin with the unique structure of these antibodies, followed by a summary of the current strategies to discover antigen-specific ulCDRs. We continue by discussing the immense sequence diversity in ulCDRs and protein engineering opportunities. Finally, we present the prospect of picobodies – individual disulfide-rich antigen-binding peptides isolated from bovine ulCDRs.

2. Structure of the ultralong CDR

Several crystal structures of bovine ulCDR-H3s have been reported[20,24,34,35]. In all of them, the ulCDR-H3 folds into two distinct elements – a β-ribbon stalk that protrudes from the antibody framework, and a disulfide-bonded knob sitting on top of the stalk (Fig. 1a). The remaining CDRs are conserved and some of them appear to have a structural function via stabilizing interactions with the base of the stalk[20,24,28]. The bovine ulCDR-H3s are in stark contrast to CDR-H3s in conventional antibodies (Fig. 1b) and nanobodies (Fig. 1c).

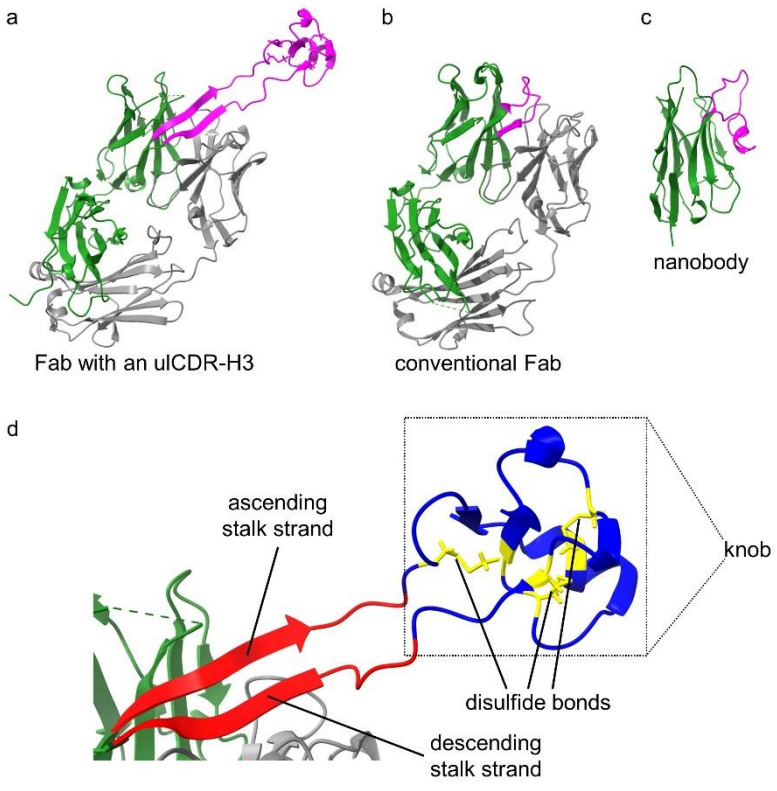

Figure 1. The structure of the ulCDR-H3. a *Bovine Fab with an ulCDR-H3 (PDB: 6OO0);* ***b*** *human Fab (PDB: 4NYL);* ***c*** *camelid-derived nanobody (PDB: 1MEL). In* ***a, b*** *and* ***c,*** *the heavy chains are marked in green, the light chains are in gray. The CDR-H3s are in magenta.* ***d*** *Magnified view of the ulCDR-H3 of a bovine antibody (PDB: 6OO0). The stalk is marked in red. The knob is marked in blue with the disulfide bridges marked in yellow.*

2.1. Stalk

The *stalk* consists of two antiparallel β-strands forming a β-hairpin "bridge" between the knob and the immunoglobulin scaffold (Fig. 1d)[24]. The ascending strand starts with a conserved motif (often TTVHQ) followed by a variable stretch of mostly hydrophilic residues[17]. The TTVHQ motif is encoded by the VH1-7 gene segment (almost exclusively used in the antibodies with ulCDR-H3s) and forms a key structural part of the ascending stalk strand[19]. The descending strand typically contains alternating aromatic residues that can form stabilizing stacking interactions[19,20,24,34]. Interestingly, the distance between the knob and the antibody framework is relatively conserved (~20 Å), but slight variations can occur in the stalk length, curvature and tilt angle that can change the relative position of the knob[34]. The notion that the stalk has an important structural function was supported by recent studies. Svilenov et al. analyzed mutants of a model bovine antibody, NC-Cow1, and showed that deleting or replacing the stalk with flexible linkers reduces the conformational stability and impairs the secretion by mammalian cells[36]. Despite the lower stability, the NC-Cow1 mutant with truncated stalk still binds to the antigen[36]. Alanine scanning in the ulCDR-H3 of NC-Cow1 revealed that mutations in the stalk do not affect the binding to the antigen measured in ELISA, but slightly reduce the virus neutralization efficiency[35]. Overall, it appears that the stalk conveys structural stability and favorable display of the antigen-binding knob away from the antibody framework. However, the knob can fold independently of the stalk to determine the antigen specificity.

2.2. Knob

With a size of only 3-6 kDa, the knob is the smallest known antigen-binding domain found in immunoglobulins (Fig. 1d)[24,33]. The knob sequence is encoded by the IGHD8-2 gene segment which also encodes the adjacent tyrosine ladder of the descending strand of the stalk[19]. The knob is diversified by variations in its length, the number of cysteines, different disulfide bond patterns and fold motifs[35,37]. Despite this vast diversity, the knob also has some conserved features (Fig. 1d)[34]. The knob is initiated by a type I β-turn that is typically formed by a CPDG motif where the first cysteine is conserved in almost all ulCDRH3s[17,37]. In contrast to larger globular proteins, the knobs do not have a pronounced hydrophobic core[34]. Instead, several knobs contain three short antiparallel β-strands connected by two hypervariable loops[20]. In addition, knobs can have different relative orientations to the antibody framework[24,34,35]. This is mainly attributed to variations in the knob/stalk hinge region, which may further diversify the ulCDRH3 repertoire[34].

The pivotal functional role of the knob has been elucidated by mutagenesis studies. Deletion of the 41-residue knob in the antibody NC-Cow1 abolishes binding to the antigen[36]. Similarly, removal of the knob in the H12 antibody leads to a complete loss of antigen binding[24]. To further differentiate on this observation, each third of the knob in H12 was mutated to an

irrelevant sequence resulting in three different mutants, which showed that the C-terminal third of this knob is more important for binding than the other two thirds[24]. A similar observation was made by applying an analogical mutational approach on a SARS-CoV-specific bovine ulCDR-H3[29]. Furthermore, alanine scanning in the knob of NC-Cow1 identified the importance of individual residues. For example, mutating certain residues such as TyrD16 and AspD35 leads to a complete loss of antigen binding, while mutation of other knob residues does not abolish activity, but reduces binding affinity and virus neutralization efficiency[35]. In contrast to the loss of antigen binding, the deletion of the knobs in two bovine antibodies had no negative effect on the thermal stability and secretion by mammalian cells[36].

The disulfide bonds in the knob are essential for the antigen-binding properties. Site-directed mutagenesis in the NC-Cow1 knob showed that replacing any pair of cysteines that form a disulfide bond abolishes binding to the antigen[36]. However, even exchanging all six cysteines from the knob of NC-Cow1 for serines resulted in a well-secreted and thermally stable antibody[36]. In addition, molecular dynamics (MD) simulations demonstrated that the disulfide bonds provide rigidity in the knob of NC-Cow1 which seems to be essential for the antigen-binding properties but not for the overall conformational stability of the antibody[36].

3. Discovery of ultralong CDRs with desired antigen specificities

The current approaches to isolate antigen-specific ulCDR-H3s with high binding affinity are based on immunization of cows (Fig. 2).

The first bovine antibodies with an ulCDR-H3 specific to a target antigen were obtained via vaccination of cattle (Fig. 2a)[24]. A Holstein steer was immunized with an inactivated bovine viral diarrhea virus (BVDV). After three boosters separated by one-month intervals, blood was collected to isolate lymphocyte mRNA. The cDNA of the isolated V_H sequences was synthesized and fused to bovine C_H1 and human Fc. The obtained HC sequences were cloned into the pFUSE expression vector for *E. coli* transformant selection and 132 unique HC sequences were identified. The isolated HCs were co-transfected with an invariant LC to create a small spatially addressed library. From this library, several binders to BVDV antigens were isolated, including the H12 antibody that has an ulCDR-H3 responsible for the antigen specificity[24].

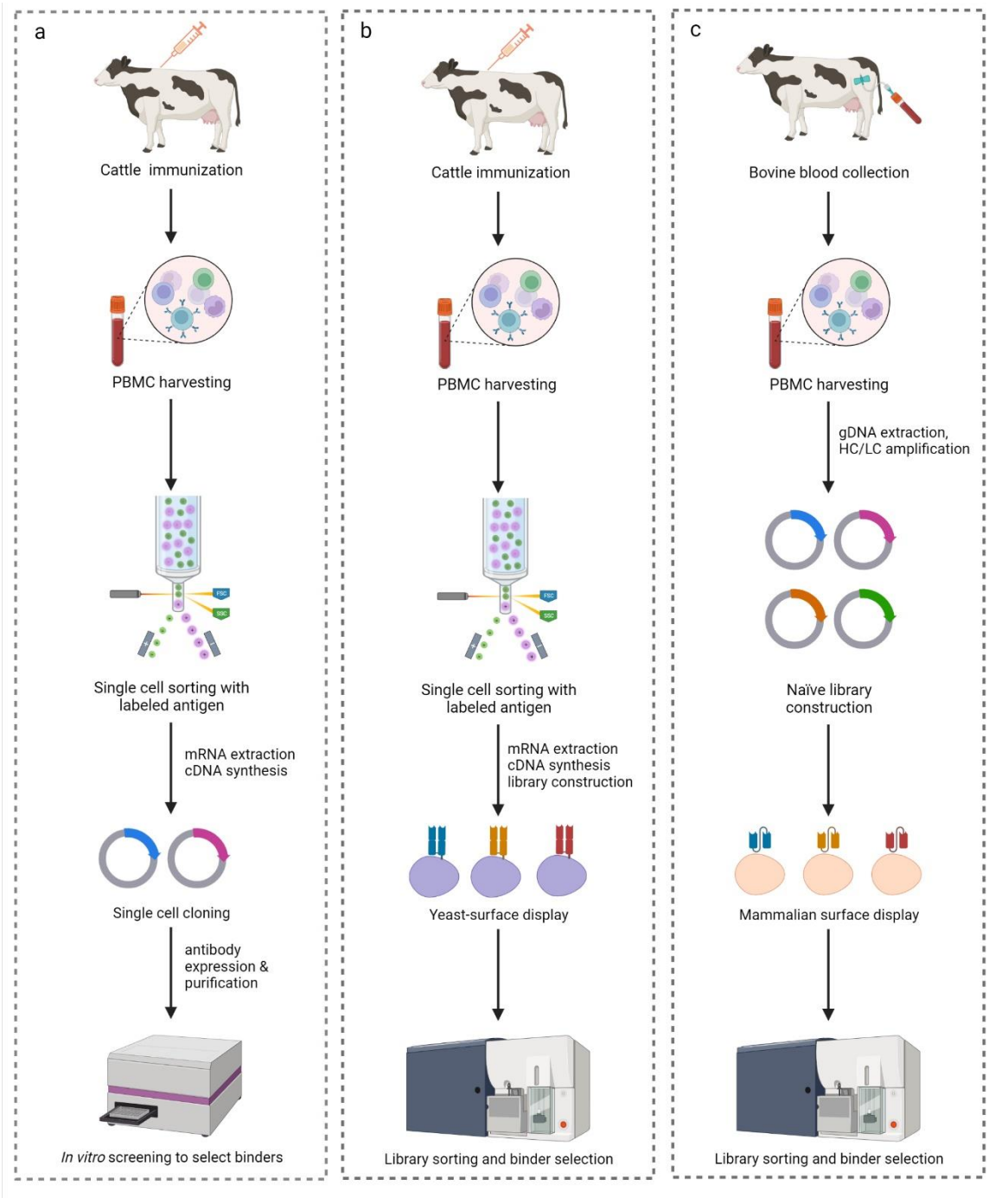

Figure 2. Workflow for the discovery of ulCDR-H3s with desired antigen specificities. ***a*** *Cattle immunization followed by isolation of peripheral blood mononuclear cells (PBMCs) and single cell sorting with labeled antigen using fluorescence-activated cell sorting (FACS). Subsequent extraction of mRNA, cDNA synthesis with RT-PCR and single cell cloning into expression vector. After antibody expression and purification, selection of binders with in vitro screening assays (e.g., ELISA).* ***b*** *as in* ***a****, but with construction and screening of a yeast-surface display library. Library sorting and binder selection with FACS.* ***c*** *Collection of bovine blood followed by isolation of PBMCs. gDNA extraction, HC/LC amplification and construction of a naïve library for mammalian surface display. Subsequent library sorting and binder selection with FACS.*

In another study, four six-month-old calves were immunized with a soluble cleaved trimer of the HIV envelope glycoprotein (Env)[28]. Mononuclear cells were isolated and sorted for antigen-specific cells with fluorescence-activated cell sorting (FACS). cDNA was amplified with primers specific for cow immunoglobulin, heavy and light chain pairs were cloned into human antibody expression vectors and the plasmids were used for co-transfection of HEK cells. This resulted in the isolation of ten bovine antibodies that all had ulCDR-H3s and five of which showed broad virus neutralization efficiency. One of these antibodies, NC-Cow1, showed the highest neutralization breadth of 72 % on a 117-virus panel and a high potency[28].

In further work aiming to discover bovine antibodies against HIV, adult cows were immunized with different HIV-1 Env antigens in a cross-vaccination approach[38]. HIV-1 Env-specific B cells were isolated via FACS. The variable antibody domains were amplified and cloned into an expression vector with human antibody constant domains. Subsequently, 92 chimeric antibodies binding to an HIV-1 Env were identified, including two antibodies with ulCDR-H3s that exhibited a higher virus neutralization breadth compared to NC-Cow1[38].

Macpherson et al. immunized two adult Holstein cattle with the complement component C5[33]. Using FACS, antigen-specific B-cells were isolated from a draining lymph node of one immunized cow. Subsequently, the cDNA was amplified from the antigen-enriched pool of immune cells with primers specific for the CDR-H3. After a second PCR round, deep sequencing was performed on the CDR-H3 library and 154 ulCDR-H3s were isolated[33]. A subset of these ulCDR-H3s was then used for the construction of fusion proteins and isolated knobs.

The vaccination approaches to discover antigen-specific ulCDR-H3s can also be combined with *in vitro* display technologies. For example, Pekar et al. immunized cattle with recombinant epidermal growth factor receptor (EGFR) followed by mRNA isolation from blood and cDNA synthesis (Fig. 2b)[30]. The ulCDR-H3s were amplified from the cDNA using specific primers. Next, the yeast display library was designed in which the amplified ulCDR-H3s were displayed on chimeric Fabs composed of bovine variable domains and human constant domains. Only one light chain was used for the entire Fab library. Finally, FACS was utilized to isolate six unique ulCDR-H3s that bind EGFR with high affinity[30]. Using an analogical cattle immunization/yeast surface display approach, Klewinghaus et al. discovered 13 ulCDR-H3s with high binding affinity to the human natural cytotoxicity receptor NKp30[31].

Mammalian cell surface display has also been used to discover antigen-specific ulCDR-H3 (Fig. 2c)[29]. Burke et al. developed an scFv system for mammalian surface display of ulCDR-H3s employing a pBovShow vector where the heavy chains are paired with an invariant lambda light chain. The system was used to display a naïve ulCDR-H3 library with a low diversity (< 1

x 10^4). Strikingly, despite the small library size, the approach yielded a bovine ulCDR-H3 with a moderate binding affinity to the spike proteins of SARS-CoV and SARS-CoV-2[29].

In contrast to yeast and mammalian surface display, phage display has had less success in the discovery of ulCDR-H3s. Three studies employing phage display to isolate antigen-specific bovine antibodies yielded only binders with conventional CDR-H3s[39–41]. However, these studies did not employ a library design strategy to increase the success of discovering ulCDR-H3 binders.

Lastly, there were attempts to develop a transgenic mouse in which the endogenous murine D_H segments were replaced with bovine D_H genes[42]. Although the genetically engineered mice produced antibodies and showed humoral immune responses against several antigens, no ulCDR-H3s were found. This indicates that the long bovine D_H gene alone is not sufficient for the generation of ulCDR-H3s[42].

4. Epitopes targeted by ultralong CDRs

Conventional antibodies targeting protein antigens usually have large and flat antigen-binding surfaces that are not suited for occluded or concave epitopes[43]. In contrast, the knob in the ulCDR-H3 forms a smaller paratope displayed away from the antibody framework via the stalk. Therefore, the ulCDR-H3 appears fit for binding epitopes that are not accessible to conventional antibodies. This notion is supported by structural studies on bovine antibodies specific for the HIV-1 Env protein[35,38]. The HIV-1 Env is covered by a dense glycan shield which protects underlying epitopes, including the CD4 binding site (CD4bs). The broadly neutralizing bovine antibody NC-Cow1 employs its ulCDR-H3 to bind the CD4bs[28]. The binding mode of NC-Cow1 is not an exception because several other bovine antibodies with ulCDR-H3s also target the CD4bs[38].

The detailed structural basis of NC-Cow1 interaction with the HIV Env was elucidated recently (Fig. 3a)[35]. The crystal structure revealed a jaw-like arrangement of the antigen-binding site between two loops of the NC-Cow1 knob and the CD4 binding loop of the HIV Env. Despite the broad neutralization breadth and high potency of NC-Cow1, the buried molecular surface of the interaction with the antigen is small compared to human antibodies that bind to the CD4bs[35]. Interestingly, the cryo-EM data indicated flexibility in the region between the knob and the framework of NC-Cow1, which was also observed in MD simulations[35,36].

Burke et al. used hydrogen-deuterium exchange and site-directed mutagenesis to find that the knob of a bovine antibody binds to a glycan-shielded cryptic epitope in the spike protein of *Sarbecoviruses*[29]. The epitope is located on the inner face of the receptor-binding domain and becomes only transiently exposed upon structural movements which makes it a challenging target for conventional antibodies. However, the epitope is a lucrative drug target because it

corresponds to a conserved vulnerability site of *Sarbecoviruses*[44]. Remarkably, the bovine antibody neutralized pseudo-typed SARS-CoV, possibly by destabilizing the prefusion spike protein complex. Interestingly, only two conventional antibodies that target this site in *Sarbecoviruses* were discovered previously, and their isolation required multiple rounds of mice immunization[45].

Bovine ulCDR-H3 can also attain specificity for non-viral epitopes. For example, the K8 and K92 knobs isolated from ulCDR-H3s were crystallized in complex with their antigen, the complement component 5 (C5) (Figs 3b and c)[32,46]. The two knobs interact with distinct epitopes on the C5[32]. The K8 binds in the macroglobulin domain (MG) 8 in the alpha-chain, while the epitope of K92 is in a cleft between the MG1 and MG5 domains in the beta-chain. Excitingly, the bovine knobs induce allosteric effects on C5 with therapeutic potential[32].

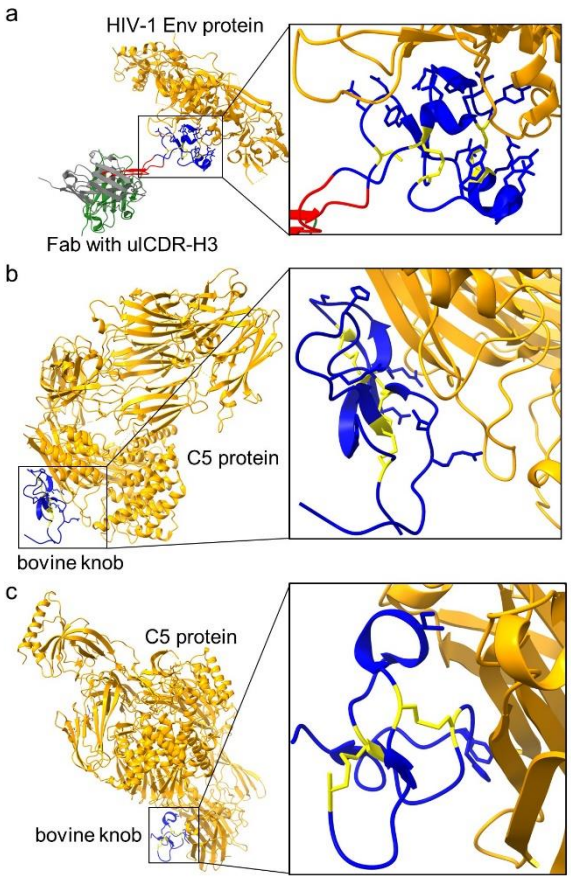

Figure 3 Crystal structures of knob-antigen complexes. a *Crystal structure of bovine Fab NC-Cow1 in complex with HIV-1 BG505 SOSIP.664 (PDB: 6OPA).* ***b*** *Crystal structure of human complement C5 in complex with the K8 bovine knob domain peptide (PDB: 7AD7).* ***c*** *Crystal structure of human complement C5 in complex with the K92 bovine knob domain peptide (PDB: 7AD6). Knobs are marked in blue with disulfide bridges in yellow and the interacting amino acids visualized. Fab heavy chains are marked in green, light chains in gray. Antigens are marked in orange gold.*

Cancer-related epitopes can also be targeted by ulCDR-H3s. For example, six antibodies with ulCDR-H3s exhibited a broad epitope coverage spread over three domains of the EGFR[30]. Noteworthily, the bovine antibodies did not compete with the natural ligand of EGFR or the conventional anti-EGFR antibodies cetuximab and matuzumab[30]. Klewinghaus et al. also reported ulCDR-H3 antibodies against another target protein with relevance in immuno-oncology – the NKp30[31]. Five of these antibodies competed with a natural ligand of NKp30[31].

Overall, the epitope space for ulCDR-H3s is diverse. There are exciting indications that the ul-CDR-H3s can bind to antigen sites that are more challenging for conventional antibodies. Despite the anticipated smaller interface of the knob-antigen interaction compared to conventional antibodies, ulCDR-H3 antibodies isolated via cattle immunization have nanomolar to picomolar binding to their antigens[28,30,31,33,36].

5. Diversity of the ultralong CDR-H3

Enormous sequence diversity of natural bovine ulCDR-H3s can be found in the antibody repertoire of cattle.[47] There are almost 100 unique ulCDR-H3 or knob sequences with known antigens or crystal structures[13,20,24,28–35,38,46,48,49]. We compiled an overview of these sequences and their basic characteristics (Table 1). A comparison based on sequence homology demonstrates that these sequences are very diverse (Fig. 4a). The 91 CDR-H3 sequences in our analysis have an average length of 58 amino acids and an average pI value of 5.9 (Figs 4b and c). All but two of them (A01 and F103) have a conserved first cysteine in the beginning of the knob region. Most sequences contain 4, 6 or 8 cysteines, with two exceptions - E03, which is also the shortest of all CDR-H3s, has 2 cysteines and BOV-5 has an uneven number of 5 cysteines (Fig. 4d).

Nr. of sequences	*CDR-H3 length*	*Nr. Cys*	*Antigen*	*pI*	*MW [Da]*	*Isolation strategy*	*PDB*	*Ref.*
1[a]	51	6	SARS-CoV and SARS-CoV-2	6.0	5858	Mammalian surface display (SARS-CoV-2-naïve heavy-chain library)		[29]
11[b]	53 - 61	6 - 8	Gp140 HIV-1	5.4 - 7.7	5907 - 7008	Cattle immunization + B-cell isolation and sequencing		[38]
17[c]	53 - 63	4 - 8	NKp30	4.6 - 7.8	5628 - 7250	Cattle immunization + Yeast surface display		[31]
15[d]	56 - 62	4 - 8	EGFR (cocktail approach)	4.6 - 7.8	6133 - 7263	Cattle immunization + Yeast surface display		[30]
13[e]	46 - 65	4 - 8	C5	4.6 - 7.9	5152 - 7414	Cattle immunization + B-cell isolation and sequencing	7AD6, 7AD7, 7OP0f	[32,33,46]
7[g]	60 - 61	6	HIV Env trimer	4.47 - 6.06	6858 - 7343	Cattle immunization + B-cell isolation and sequencing	6OO0, 6OPA	[28,35]
14[h]	48 - 62	4 - 8	FMDV	4.5 - 7.8	5498 - 7229	Cattle immunization + B-cell isolation and sequencing		[49]
7	50 - 63	4 - 8	unknown	4.17 - 8.26	5680 - 7256	SMRT sequencing of a cDNA library	6E9G-K, 6E9Q, 6E9U-V	[34]
3	44 - 63	2 - 8	unknown	4.2 - 4.71	5110 - 7003	As in [24]	5IHU, 5IJV, 5ILT	[20]
1	63	6	BVDV	7.8	6973	Cattle immunization + heavy-chain library construction for small spatially addressed library		[24]
2	56 - 61	6	unknown	6.04 - 6.27	6360 - 7021	Hetero-hybridoma from PBL of a BLV-infected cow	4K3E, 4K3D	[13,24,48]

Table 1. *Overview of published CDR-H3 sequences with known antigens or crystal structures. For determining the starting and ending positions of the CDR-H3s, the Kabat numbering scheme was used.*

[a] *Isolated as a functional scFv*

[b] *Six were functional as chimeric Abs in a virus neutralization assay.*

[c] *Thirteen were functional in the monovalent chimeric antibody format.*

[d] *Six were functional in chimeric IgG format in in vitro binding assays.*

[e] *52 were produced as scFc fusion, 14 of which had binding in antigen-specific ELISA. Of the latter, six were produced as fusions with PGT121, five of which had the desired binding in surface plasmon resonance.*

[f] *These crystal structures are of isolated knobs bound to their antigen and do not include the stalk regions.*

[g] *Ten functional antibodies were isolated, with some of them sharing the same CDR-H3 sequence.*

[h] *Nine were functional as bovine antibodies in a virus neutralization assay.*

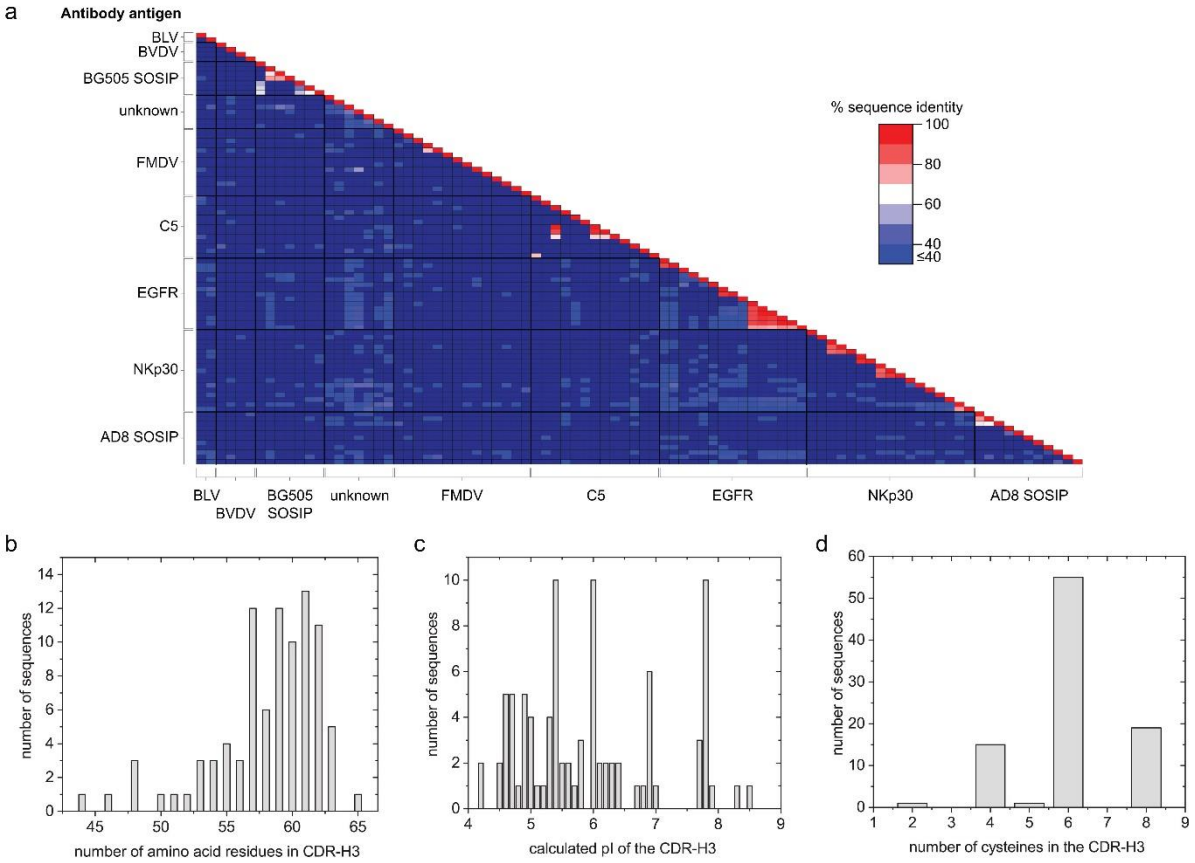

Figure 4 Sequence diversity of ulCDR-H3s with known antigens or crystal structures. ***a*** *Heat map showing percentage identity between ulCDR-H3 sequences. Pairwise alignment was done with Protein BLAST, using a cut-off of 0.01 for E values. The ulCDR-H3 sequences are clustered with their respective antigens. SARS-CoV-2 RBD-targeting ulCDR-H3 is not displayed as it has no sequence identity with any other ulCDR-H3.* ***b-d*** *Histograms showing distributions of the ulCDR-H3* ***a*** *length,* ***c*** *pI values, and* ***d*** *number of cysteines.*

6. Protein engineering opportunities

The peculiar structure of the ulCDR-H3 has inspired various protein engineering approaches (Fig. 5).

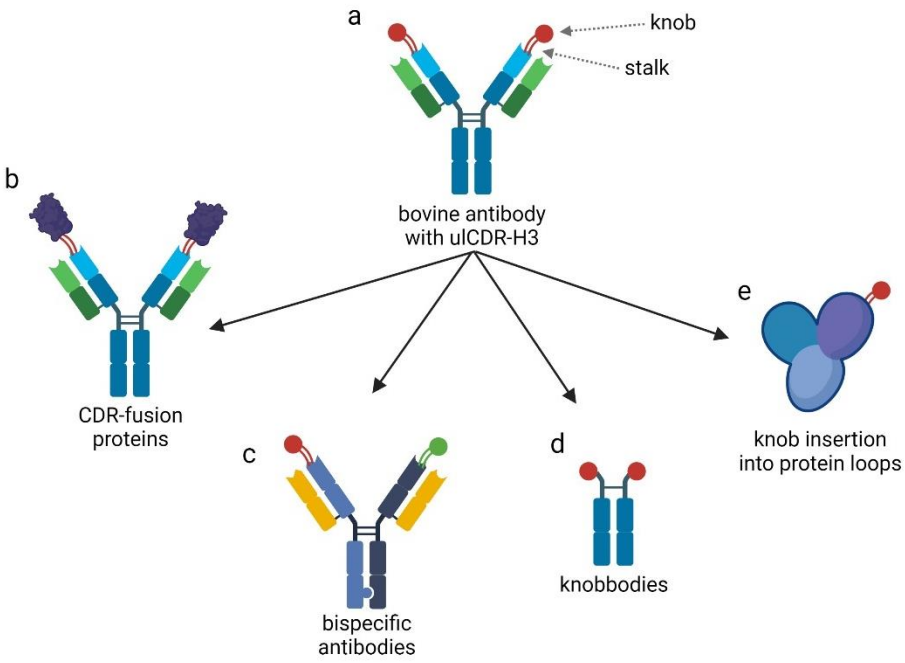

Figure 5. Protein engineering approaches inspired by antibodies with ulCDR-H3s. ***a*** *Bovine antibody with an ulCDR-H3.* ***b*** *Antibody fusion proteins with another protein (payload) inserted into the CDR-H3 or other CDRs.* ***c*** *Bispecific bovine antibodies with two different HCs and a common LC.* ***d*** *Knobbodies obtained by ligating the knob to an Fc segment.* ***e*** *Bovine knob insertion into loops of proteins (e.g., albumin).*

6.1. Antibody fusion proteins

Fusing proteins to immunoglobulins is a compelling approach to develop new biotherapeutics[50]. Antibody fusion proteins consist of an antibody-derived moiety to which one or more proteins, such as cytokines, toxins, enzymes, or peptides, are ligated. Engineered fusion proteins offer several benefits such as prolonged half-lives *in vivo*, fine-tuned biological activities and improved manufacturability[50,51]. However, there are challenges for developing novel antibody-fusions, for example, due to the need to introduce compatible junction sites between the fusion partners, the risk of immunogenicity[51], or structural and stability issues due to potential misfolding[52]. Therefore, there is a need for straightforward approaches for rational engineering of functional and stable antibody-fusion drug candidates.

The architecture of the ulCDR-H3s revealed that the knob is structurally isolated from the antibody via the stalk. This finding inspired protein engineering approaches to obtain antibody fusions by replacing the knob domain with proteins that have various biological activities[53]. For example, the knob of the bovine antibody BVL1H12 was replaced with bovine granulocyte-

colony stimulating factor (bGCSF), resulting in a chimeric antibody bearing two bGCSF moieties (Fig. 5b). This CDRH3-fusion protein exhibited a significantly prolonged half-life in mice of 8.8 – 9.5 d (compared to 2.9 h for isolated bGCSF) while maintaining the biological activity[54]. Using the same approach, the half-life of human erythropoietin (hEPO) could be extended from 3.6 h to 6.4 d when hEPO was fused to the ulCDR-H3 of BVL1H12. The biological activity of the antibody-hEPO fusion was similar to the free hEPO[55].

Besides exchanging the knob for other proteins, it is also possible to replace the entire ulCDR-H3 by engineered linkers such as a heterodimeric antiparallel coiled-coil motif (Fig. 5b)[56]. Such a stalk replacement did not negatively affect the thermodynamic stability of a bovine antibody. Furthermore, grafting bGCSF onto the coiled-coil stalk worked as well as grafting onto the natural bovine stalk[56].

Biologically active peptides can also be ligated to an ulCDR-H3. For example, a beta hairpin peptide antagonist of the chemokine receptor CXCR4 was successfully grafted onto an ulCDR-H3[57]. The resulting antibody-peptide fusion bound to CXCR4 and exhibited antagonist effector functions in cellular assays[57]. The fusion of proteins to CDR loops is not exclusive for CDR-H3. For example, CDR-H2 can also be used as a fusion site[57].

The success with the CDR-fusion experiments using bovine antibodies raised the question whether antibodies from other species are also amenable to similar engineering approaches. For example, trastuzumab was fused with the GLP-1 receptor agonist Exendin-4 (Ex-4)[58]. Ex-4 was placed into a coiled-coil stalk next to a Factor Xa cleavage site in CDR-H3. The resulting clipped trastuzumab-CDR-H3-coil-Ex-4 fusion exhibited a greatly improved serum half-life in mice (2.4 d compared to 1.5 h for Ex-4 alone) and showed sustained control of blood glucose levels[58]. A similar approach also worked with human growth hormone and human leptin[59]. Furthermore, dual-acting fusion proteins could be obtained by fusing hGCSF and hEPO to different CDRs of trastuzumab[60]. Other possibilities for dual-acting modalities were also reported, for example where human leptin is inserted into CDR-L3 while Ex-4 is at the N-terminus of the HC of trastuzumab[60,61]. Overall, these results show that the successful fusion of proteins and peptides to CDR loops is not exclusive to bovine antibody scaffolds.

Besides natural fusion partners, mini-proteins from *de novo* design can also be grafted onto CDRs. For example, parts of the de-novo designed helical peptide LCB1 that binds to the Spike protein of SARS-CoV-2 were used to create an artificial ulCDR-H3 in bovine and human antibodies. The resulting antibody-LCB1 fusions exhibited sub-nanomolar neutralizing activity against two variants of SARS-CoV-2[36]. This proof-of-principle study demonstrates the exciting possibilities for computational design of ulCDR-H3s with new specificities.

6.2. Chimeric antibodies with human scaffolds and bovine knobs

A main challenge to using the bovine antibodies for therapeutic applications is the risk of immunogenicity. Therefore, it is compelling to develop antibodies with ulCDR-H3s that are more human-like. Sok et al. showed that it is possible to replace the CDR-H3 of the human PG9 antibody with the ulCDR-H3 from the bovine antibody NC-Cow1 that neutralizes HIV[28]. The chimeric antibody exhibited only slightly reduced virus neutralization breadth and potency compared to the parent bovine antibody, demonstrating the feasibility of this approach.

To minimize the size of the bovine sequence, only the knob can be grafted onto human antibodies instead of the entire ulCDR-H3. This was shown with the knob of NC-Cow1 which was inserted into the CDR-H3 loop of three human antibodies with different CDR-H3 lengths. It turned out that insertion of the knob into the short CDR-H3 of trastuzumab led to a more than five-fold reduction in binding affinity to the HIV-1 Env protein, while insertion of the knob into longer human CDR-H3s had minimal effect on the binding affinity[36]. This suggests that the distance of the knob to the V_H framework is critical for the functionality of the knob mini domain, and that the acceptor human antibody ideally has a stalk-like natural CDR that can support and display the bovine knob[36].

6.3. Bispecific antibodies

The conventional bispecific antibodies are made of two different HCs and two different LCs. During production, it is essential that each HC assembles with the corresponding LC. For that reason, there is significant interest in developing technologies for bispecific antibodies that circumvent the HC/LC pairing problem. Interestingly, the HCs of bovine antibodies with ulCDR-H3s pair with a restricted set of lambda LCs and the antigen binding is via the knob in the HC. Therefore, the bovine antibodies with ulCDR-H3s offer a natural platform for bispecific antibodies that have different HCs but share the same LC (Fig. 5c). This was demonstrated by pairing of 13 heavy chains encoding an ulCDR-H3 against NKp30 with 2 heavy chains encoding an ulCDR-H3 against EGFR, always in combination with the same light chain[31]. In total, 26 common light chain bispecific constructs were successfully produced and 24 of them exhibited high homogeneity (> 90 % monomer) after protein A purification. The bispecifics were able to bind both antigens and most of them induced NK-cell mediated killing of a cancer cell line that overexpresses EGFR[31].

Another exciting opportunity to engineer small bispecific molecules is to fuse the bovine knob to nanobodies. For example, a bovine knob specific for the C5 complement protein was inserted into a framework loop of a nanobody against the C3 complement component[62]. The insertion position was on the distal nanobody part relative to the three natural CDR loops. This bispecific entity could be produced and purified to homogeneity as a fusion to a single-chain

Fc tag. The dual specificity and functionality of the picobody-nanobody fusion for the two targets (C3 and C5) was confirmed in *in vitro* binding assays and a hemolysis assay[62].

6.4. Knobbodies

Bovine knobs can also be ligated to the hinge region of Fc segments to generate knob-Fc fusions called knobbodies (Fig. 5d)[30]. Pekar et al. fused anti-EGFR knobs to the Fc part of human IgG1[30]. Several of the resulting knobbodies could be produced and were functional. However, the knobbodies were prone to aggregate, but two constructs contained more than 80 % monomer after purification. Interestingly, a combination of non-reducing and reducing SDS-PAGE revealed that some of the knobbody aggregates are covalently bound, most likely via disulfide bonds[30].

6.5. Albumin-knob fusions

In addition to CDR loops, the bovine knobs can also be grafted into loops of other proteins (Fig. 5e). For example, four different positions were used for the insertion of picobodies into rat serum albumin (RSA)[62]. All four picobody-RSA fusions were well expressed in mammalian cells with yields around 100-200 mg/L and the secondary structure of the molecules was minimally affected as indicated by far-UV circular dichroism[62]. Furthermore, the knob-RSA fusion proteins retained their high potency for the knob target, the C5 protein, and bound to the neonatal Fc receptor (FcRn).

7. Picobodies as isolated target-binding moieties

The knob domain of the ulCDR-H3 is the smallest antibody-derived antigen-binding entity. It has been shown that knobs can be produced as isolated peptides[33,46]. Due to their small size and origin, the isolated knobs were named “picobodies” by the group around Vaughn Smider. A trademark on the name picobodies exists which draws parallels with the origin of the nanobody name used for camelid V_{HH}s[63]. In addition, the discovery of antibodies with ulCDR-H3s is offered commercially under the trademark OmniTaur™ (OmniAb, Inc).

The picobodies have structural similarities with other natural peptides of similar size and structure such as cystine-knot mini-proteins (knottins)[64], defensins[65], and snake- and spider-venom-derived toxins[66]. Interestingly, these cysteine-rich peptides have evolved to serve a wide range of functions such as protection against pathogens or other threats. The disulfide-stabilized mini-proteins typically exhibit high thermodynamic stability, rigid structure and potent biological activity which makes them lucrative for biomedical applications[64–66]. The picobodies are an important addition to the family of cysteine-rich mini-proteins with the essential advantage that picobodies against new pharmacological targets can be discovered quickly by immunization of cattle.

With their molecular weight of 3-6 kDa, the picobodies also have a similar size to non-immunoglobulin scaffolds, such as affibodies, that are used as high affinity binders[67,68]. These scaffolds are based on diverse structural templates and are typically isolated from *in vitro* designed synthetic libraries.

Picobodies can be produced in biological systems or via chemical synthesis[33,46]. It could be possible in principle to express isolated picobodies in mammalian cells, but the yields are low[33]. To solve this issue, the picobody can be produced in its natural form fused to the ulCDR-H3 of an antibody with protease sites in the ascending and descending strands of the stalks. Subsequently, the picobody can be enzymatically cleaved from the rest of the antibody. This approach was demonstrated with C5-specific picobodies fused to a human antibody scaffold with TEV protease cleavage sites[33]. Upon excision with TEV protease, the individual picobodies were purified by chromatography. Noteworthily, four out of the six isolated picobodies had excellent binding affinity to the antigen (C5) in surface plasmon resonance experiments[33]. Four of the picobodies target different epitopes and elicit a range of effects on the C5 antigen via allosteric modulation. Interestingly, the first picobody-antigen structure revealed binding to a previously unrecognized regulatory site on the C5 antigen, in line with the notion that picobodies can bind different epitopes compared to conventional antibodies[32]. Four of the C5-specific picobodies were also produced by solid-phase chemical synthesis, using two different techniques[46]. In the first, selected cysteines were protected to facilitate the formation of the right disulfide pattern. In the second technique, a free-energy approach was followed where the formation of the disulfide bonds occurred spontaneously under oxidizing conditions. Excitingly, the synthesized picobodies produced by the free energy method showed high binding affinity, comparable to the picobody produced by chemical synthesis with protected cysteines or via a cleavable scaffold secreted from mammalian cells[33,46]. Although the yields of active picobodies produced by this method were not reported, the free energy approach indicates that some picobodies can spontaneously fold and form the right disulfide pattern after chemical synthesis which will facilitate their production. Importantly, the crystal structures of the K92 picobody produced in mammalian cells or via chemical synthesis revealed almost identical folds (Fig. 6)[46].

During the chemical synthesis of C5-targeting picobodies, sequence modifications are also possible[46]. For example, head-to-tail cyclization was used to increase the stability against proteases while the biological activity was preserved for several C5-specific picobodies. Furthermore, non-natural amino acids were introduced into the K92 picobody with the aim to enhance the binding affinity, although the resulting variants did not have a higher binding compared to K92 wild-type. For prolonging the half-life and increasing the serum stability of picobodies, functionalization with fatty acids during chemical synthesis can be used. For

example, palmitoylation increased the stability of the picobody K57 in serum and prolonged the half-life of the picobody from 17 min to 1.6 h.

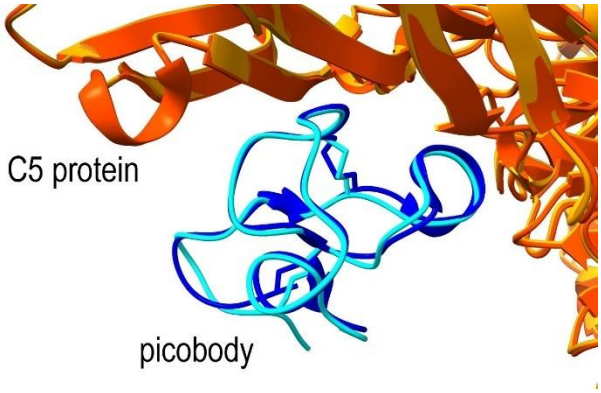

Figure 6. Superposition of biologically and chemically produced picobody. *The biologically produced K92 (PDB: 7AD6) is marked in blue with the C5 antigen in orange-red, the chemically produced K92 (PDB: 7OP0) is marked in cyan with the antigen in orange-gold.*

8. Outstanding challenges

The antibodies with ulCDR-H3s and the corresponding picobodies offer promising avenues for biomedical applications. However, to harness the full potential of the ulCDR-H3s and picobodies, several key challenges will have to be addressed. For example, little is known about the immunogenicity of bovine antibodies and picobodies in humans. The highly diverse ulCDR-H3 and the picobodies have no human analogs. However, due to the high sequence diversity, it is possible that some ulCDR-H3s are immunogenic while others are not. To circumvent this risk, the antibodies with ulCDR-H3 could be used for therapeutic applications where immunogenicity is less critical (e.g., intraocular, peroral administration). Overall, it will be important to understand the intrinsic immunogenicity of these proteins in different applications.

The full-length antibodies with ulCDR-H3 can be produced in biological systems, but the manufacturing of the picobodies is not yet straightforward. Chemical synthesis can be used, but this process is expensive and issues with the folding of the picobodies are likely if there are more than two disulfide bonds in the structure. Therefore, it will be essential to establish cost-effective methods for picobody production.

For biomedical applications, the antibodies with ulCDR-H3 and picobodies have to exhibit favorable biophysical properties such as good aqueous solubility and high stability. These features will have to be investigated in detail to reveal whether these proteins exhibit drug-like properties.

Finally, the big hope is that ulCDR-H3s and picobodies can expand the space of pharmacological targets beyond what is possible for conventional antibodies. The first studies

already hint that the bovine knobs often bind to novel or occluded epitopes[28,32,33,35,46], but the true potential of picobodies could become evident with more exciting targets such as G protein-coupled receptors (GPCRs).

References

1. Huber, R., Deisenhofer, J., Colman, P. M., Matsushima, M. & Palm, W. Crystallographic structure studies of an IgG molecule and an Fc fragment. *Nature* **264**, 415–420 (1976).

2. French, D. L., Laskov, R. & Scharff, M. D. The Role of Somatic Hypermutation in the Generation of Antibody Diversity. *Science* **244**, 1152–1157 (1989).

3. Papavasiliou, F. *et al.* V(D)J Recombination in Mature B Cells: A Mechanism for Altering Antibody Responses. *Science* **278**, 298–301 (1997).

4. Chi, X., Li, Y. & Qiu, X. V(D)J recombination, somatic hypermutation and class switch recombination of immunoglobulins: mechanism and regulation. *Immunology* **160**, 233–247 (2020).

5. Parvari, R. *et al.* Chicken immunoglobulin gamma-heavy chains: limited VH gene repertoire, combinatorial diversification by D gene segments and evolution of the heavy chain locus. *EMBO J* **7**, 739–744 (1988).

6. de los Rios, M., Criscitiello, M. F. & Smider, V. V. Structural and genetic diversity in antibody repertoires from diverse species. *Curr Opin Struct Biol* **33**, 27–41 (2015).

7. Jain, T. *et al.* Biophysical properties of the clinical-stage antibody landscape. *Proc Natl Acad Sci U S A* **114**, 944–949 (2017).

8. Johnson, G. & Wu, T. T. Preferred CDRH3 lengths for antibodies with defined specificities. *Int Immunol* **10**, 1801–1805 (1998).

9. Collis, A. V. J., Brouwer, A. P. & Martin, A. C. R. Analysis of the Antigen Combining Site: Correlations Between Length and Sequence Composition of the Hypervariable Loops and the Nature of the Antigen. *J Mol Biol* **325**, 337–354 (2003).

10. Raybould, M. I. J. *et al.* Five computational developability guidelines for therapeutic antibody profiling. *Proceedings of the National Academy of Sciences* **116**, 4025–4030 (2019).

11. Berens, S. J., Wylie, D. E. & Lopez, O. J. Use of a single VH family and long CDR3s in the variable region of cattle Ig heavy chains. *Int Immunol* **9**, 189–199 (1997).

12. Lopez, O., Perez, C. & Wylle, D. A single VH family and long CDR3s are the targets for hypermutation in bovine immunoglobulin heavy chains. *Immunol Rev* **162**, 55–66 (1998).

13. Saini, S. S., Allore, B., Jacobs, R. M. & Kaushik, A. Exceptionally long CDR3H region with multiple cysteine residues in functional bovine IgM antibodies. *Eur J Immunol* **29**, 2420–2426 (1999).

14. Burke, M. J., Stockley, P. G. & Boyes, J. Broadly Neutralizing Bovine Antibodies: Highly Effective New Tools against Evasive Pathogens? *Viruses* **12**, 473 (2020).

15. Haakenson, J. K., Huang, R. & Smider, V. v. Diversity in the cow ultralong CDR H3 antibody repertoire. *Front Immunol* **9**, 1262 (2018).

16. Ma, L. *et al.* Internal Duplications of DH, JH, and C Region Genes Create an Unusual IgH Gene Locus in Cattle. *The Journal of Immunology* **196**, 4358–4366 (2016).

17. Stanfield, R. L. *et al.* Chapter Four - The Unusual Genetics and Biochemistry of Bovine Immunoglobulins. in *Advances in Immunology* (ed. Alt, F.) vol. 137 135–164 (Academic Press, 2018).

18. Mikocziova, I., Greiff, V. & Sollid, L. M. Immunoglobulin germline gene variation and its impact on human disease. *Genes Immun* **22**, 205–217 (2021).

19. Deiss, T. C. *et al.* Immunogenetic factors driving formation of ultralong VH CDR3 in Bos taurus antibodies. *Cell Mol Immunol* **16**, 64–75 (2019).

20. Stanfield, R. L., Wilson, I. A. & Smider, V. V. Conservation and diversity in the ultralong third heavy-chain complementarity-determining region of bovine antibodies. *Sci Immunol* **1**, (2016).

21. Saini, S. S., Farrugia, W., Ramsland, P. A. & Kaushik, A. K. Bovine IgM antibodies with exceptionally long complementarity-determining region 3 of the heavy chain share unique structural properties conferring restricted V H + Vλ pairings. *Int Immunol* **15**, 845–853 (2003).

22. Kurosawa, Y. & Tonegawa, S. Organization, structure, and assembly of immunoglobulin heavy chain diversity DNA segments. *Journal of Experimental Medicine* **155**, 201–218 (1982).

23. Tonegawa, S. Somatic generation of antibody diversity. *Nature* **302**, 575–581 (1983).

24. Wang, F. *et al.* Reshaping Antibody Diversity. *Cell* **153**, 1379–1393 (2013).

25. Prabakaran, P. & Chowdhury, P. S. Landscape of Non-canonical Cysteines in Human VH Repertoire Revealed by Immunogenetic Analysis. *Cell Rep* **31**, 107831 (2020).

26. Haakenson, J. K. *et al.* A Broad Role for Cysteines in Bovine Antibody Diversity. *Immunohorizons* **3**, 478–487 (2019).

27. Walther, S., Czerny, C. P. & Diesterbeck, U. S. Exceptionally Long CDR3H Are Not Isotype Restricted in Bovine Immunoglobulins. *PLoS One* **8**, e64234 (2013).

28. Sok, D. *et al.* Rapid elicitation of broadly neutralizing antibodies to HIV by immunization in cows. *Nature* **548**, 108–111 (2017).

29. Burke, M. J. *et al.* A bovine antibody possessing an ultralong complementarity-determining region CDRH3 targets a highly conserved epitope in sarbecovirus spike proteins. *Journal of Biological Chemistry* **298**, 102624 (2022).

30. Pekar, L. *et al.* Milking the Cow: Cattle-Derived Chimeric Ultralong CDR-H3 Antibodies and Their Engineered CDR-H3-Only Knobbody Counterparts Targeting Epidermal Growth Factor Receptor Elicit Potent NK Cell-Mediated Cytotoxicity. *Front Immunol* **12**, 742418 (2021).

31. Klewinghaus, D. *et al.* Grabbing the Bull by Both Horns: Bovine Ultralong CDR-H3 Paratopes Enable Engineering of 'Almost Natural' Common Light Chain Bispecific Antibodies Suitable For Effector Cell Redirection. *Front Immunol* **12**, 5586 (2022).

32. Macpherson, A. *et al.* The allosteric modulation of complement c5 by knob domain peptides. *Elife* **10**, 1–49 (2021).

33. Macpherson, A. *et al.* Isolation of antigen-specific, disulphide-rich knob domain peptides from bovine antibodies. *PLoS Biol* **18**, e3000821 (2020).

34. Dong, J., Finn, J. A., Larsen, P. A., Smith, T. P. L. & Crowe, J. E. Structural diversity of ultralong CDRH3s in seven bovine antibody heavy chains. *Front Immunol* **10**, 558 (2019).

35. Stanfield, R. L. *et al.* Structural basis of broad HIV neutralization by a vaccine-induced cow antibody. *Sci Adv* **6**, eaba0468 (2020).

36. Svilenov, H. L., Sacherl, J., Protzer, U., Zacharias, M. & Buchner, J. Mechanistic principles of an ultra-long bovine CDR reveal strategies for antibody design. *Nat Commun* **12**, 6737 (2021).

37. Jenkins, G. W., Safonova, Y. & Smider, V. v. Germline-Encoded Positional Cysteine Polymorphisms Enhance Diversity in Antibody Ultralong CDR H3 Regions. *The Journal of Immunology* **209**, 2141–2148 (2022).

38. Heydarchi, B. *et al.* Broad and ultra-potent cross-clade neutralization of HIV-1 by a vaccine-induced CD4 binding site bovine antibody. *Cell Rep Med* **3**, 100635 (2022).

39. Jeong, S. *et al.* Phage display screening of bovine antibodies to foot-and-mouth disease virus and their application in a competitive elisa for serodiagnosis. *Int J Mol Sci* **22**, (2021).

40. Hosking, C. G. *et al.* Using the local immune response from the natural buffalo host to generate an antibody fragment library that binds the early larval stages of Schistosoma japonicum. *Int J Parasitol* **45**, 729–740 (2015).

41. Hosking, C. G. *et al.* Generation of a Novel Bacteriophage Library Displaying scFv Antibody Fragments from the Natural Buffalo Host to Identify Antigens from Adult Schistosoma japonicum for Diagnostic Development. *PLoS Negl Trop Dis* **9**, e0004280 (2015).

42. Di, Y. *et al.* Reshaping the murine immunoglobulin heavy chain repertoire with bovine DH genes. *Immunology* **165**, 74–87 (2022).

43. Webster, D. M., Henry, A. H. & Rees, A. R. Antibody-antigen interactions. *Curr Opin Struct Biol* **4**, 123–129 (1994).

44. Starr, T. N. *et al.* Deep Mutational Scanning of SARS-CoV-2 Receptor Binding Domain Reveals Constraints on Folding and ACE2 Binding. *Cell* **182**, 1295-1310.e20 (2020).

45. Li, T. *et al.* Cross-neutralizing antibodies bind a SARS-CoV-2 cryptic site and resist circulating variants. *Nat Commun* **12**, 5652 (2021).

46. Macpherson, A. *et al.* The Chemical Synthesis of Knob Domain Antibody Fragments. *ACS Chem Biol* **16**, 1757–1769 (2021).

47. Safonova, Y. *et al.* Variations in antibody repertoires correlate with vaccine responses. *Genome Res* **32**, 791–804 (2022).

48. Shojaei, F., Saini, S. S. & Kaushik, A. K. Unusually long germline DH genes contribute to large sized CDR3H in bovine antibodies. *Mol Immunol* **40**, 61–67 (2003).

49. Li, K. *et al.* Development of Foot-and-Mouth Disease Virus-Neutralizing Monoclonal Antibodies Derived From Plasmablasts of Infected Cattle and Their Germline Gene Usage. *Front Immunol* **10**, 2870 (2019).

50. Silver, A. B., Leonard, E. K., Gould, J. R. & Spangler, J. B. Engineered antibody fusion proteins for targeted disease therapy. *Trends Pharmacol Sci* **42**, 1064–1081 (2021).

51. Berger, S., Lowe, P. & Tesar, M. Fusion protein technologies for biopharmaceuticals: Applications and challenges: Editor Stefan R Schmidt. *MAbs* **7(3)**, 456–460 (2013).

52. Peng, Y. *et al.* A General Method for Insertion of Functional Proteins within Proteins via Combinatorial Selection of Permissive Junctions. *Chem Biol* **22**, 1134–1143 (2015).

53. Koti, M., Saini, S. S., Sachan, A. & Kaushik, A. K. Engineered Bovine Antibodies in the Development of Novel Therapeutics, Immunomodulators and Vaccines. *Antibodies* **3**, 205–214 (2014).

54. Zhang, Y. *et al.* Functional Antibody CDR3 Fusion Proteins with Enhanced Pharmacological Properties. *Angewandte Chemie* **125**, 8453–8456 (2013).

55. Zhang, Y. *et al.* An antibody CDR3-erythropoietin fusion protein. *ACS Chem Biol* **8**, 2117–2121 (2013).

56. Zhang, Y. *et al.* An antibody with a variable-region coiled-coil 'knob' domain. *Angewandte Chemie - International Edition* **53**, 132–135 (2014).

57. Liu, T. *et al.* Rational design of CXCR4 specific antibodies with elongated CDRs. *J Am Chem Soc* **136**, 10557–10560 (2014).

58. Zhang, Y. *et al.* Rational design of a humanized glucagon-like peptide-1 receptor agonist antibody. *Angewandte Chemie - International Edition* **54**, 2126–2130 (2015).

59. Liu, T. *et al.* Functional human antibody cdr fusions as long-acting therapeutic endocrine agonists. *Proc Natl Acad Sci U S A* **112**, 1356–1361 (2015).

60. Zhang, Y., Liu, Y., Wang, Y., Schultz, P. G. & Wang, F. Rational design of humanized dual-agonist antibodies. *J Am Chem Soc* **137**, 38–41 (2015).

61. Liu, Y. *et al.* Rational design of dual agonist-antibody fusions as long-acting therapeutic hormones. *ACS Chem Biol* **11**, 2991–2995 (2016).

62. Hawkins, A. *et al.* The proximity of the N- and C- termini of bovine knob domains enable engineering of target specificity into polypeptide chains. *MAbs* **14**, 2076295 (2022).

63. Huang, L., Muyldermans, S. & Saerens, D. Nanobodies®: proficient tools in diagnostics. *Expert Rev Mol Diagn* **10**, 777–785 (2010).

64. Kintzing, J. R. & Cochran, J. R. Engineered knottin peptides as diagnostics, therapeutics, and drug delivery vehicles. *Curr Opin Chem Biol* **34**, 143–150 (2016).

65. Gao, X. *et al.* Defensins: The natural peptide antibiotic. *Adv Drug Deliv Rev* **179**, 114008 (2021).

66. Lazarovici, P. Snake- and Spider-Venom-Derived Toxins as Lead Compounds for Drug Development. in *Methods in molecular biology (Clifton, N.J.)* vol. 2068 3–26 (2020).

67. Vazquez-Lombardi, R. *et al.* Challenges and opportunities for non-antibody scaffold drugs. *Drug Discov Today* **20**, 1271–1283 (2015).

68. Škrlec, K., Štrukelj, B. & Berlec, A. Non-immunoglobulin scaffolds: a focus on their targets. *Trends Biotechnol* **33**, 408–418 (2015).

Funding details

H.L.S. acknowledges the funding from the special research fund of Ghent University, grant number BOF/STA/202109/034.

Acknowledgements

Some images in the manuscript were created with BioRender.com.