

## Developing drug-like single-domain antibodies (VHH) from in vitro libraries

M. Frank Erasmus, Andre A. R. Teixeira, Esteban Molina, Luis Antonio Rodriguez Carnero, Jianquan Li, David Knight, Roberto Di Niro, Camila Leal-Lopes, Adeline Fanni, Hallie Troell, Ashley DeAgüero, Laura Spector, Sara D'Angelo, Fortunato Ferrara & Andrew R. M. Bradbury

To cite this article: M. Frank Erasmus, Andre A. R. Teixeira, Esteban Molina, Luis Antonio Rodriguez Carnero, Jianquan Li, David Knight, Roberto Di Niro, Camila Leal-Lopes, Adeline Fanni, Hallie Troell, Ashley DeAgüero, Laura Spector, Sara D'Angelo, Fortunato Ferrara & Andrew R. M. Bradbury (2025) Developing drug-like single-domain antibodies (VHH) from in vitro libraries, mAbs, 17:1, 2516676, DOI: [10.1080/19420862.2025.2516676](https://doi.org/10.1080/19420862.2025.2516676)

To link to this article: <https://doi.org/10.1080/19420862.2025.2516676>



© 2025 The Author(s). Published with license by Taylor & Francis Group, LLC.



[View supplementary material](#)



Published online: 25 Jun 2025.



[Submit your article to this journal](#)



Article views: 5742



[View related articles](#)



[View Crossmark data](#)

REPORT



## Developing drug-like single-domain antibodies (VHH) from in vitro libraries

M. Frank Erasmus<sup>ID</sup><sup>a\*</sup>, Andre A. R. Teixeira<sup>ID</sup><sup>b\*#</sup>, Esteban Molina<sup>ID</sup><sup>a</sup>, Luis Antonio Rodriguez Carnero<sup>ID</sup><sup>a</sup>, Jianquan Li<sup>ID</sup><sup>a</sup>, David Knight<sup>a</sup>, Roberto Di Niro<sup>a</sup>, Camila Leal-Lopes<sup>ID</sup><sup>b†</sup>, Adeline Fanni<sup>ID</sup><sup>a</sup>, Hallie Troell<sup>a</sup>, Ashley DeAguiro<sup>ID</sup><sup>a</sup>, Laura Spector<sup>ID</sup><sup>a</sup>, Sara D'Angelo<sup>ID</sup><sup>a</sup>, Fortunato Ferrara<sup>ID</sup><sup>a</sup>, and Andrew R. M. Bradbury<sup>ID</sup><sup>a</sup>

<sup>a</sup>Specifica LLC, an IQVIA Business, Santa Fe, NM, USA; <sup>b</sup>New Mexico Consortium, Los Alamos, NM, USA

### ABSTRACT

Here, we describe a new VHH library for therapeutic discovery which optimizes humanness, stability, affinity, diversity, developability, and facile purification using protein A in the absence of an Fc domain. Four therapeutic humanized VHHs were used as scaffolds, into which we inserted human HCDR1s, HCDR2s and HCDR3s. The HCDR1 and HCDR2 sequences were derived from human VH3 family next-generation sequencing datasets informatically purged of sequence liabilities, synthesized as array-based oligonucleotides, cloned as single CDR libraries into each of the parental scaffolds and filtered for protein A binding by yeast display to ensure correct folding and display. After filtering, the CDR1 and CDR2 libraries were combined with amplified human HCDR3 from human CD19<sup>+</sup> IgM<sup>+</sup> B cells. This library was further improved by eliminating long consecutive stretches of tyrosines in CDR3 and enriching for CDR1–2 diversity with elevated tolerance to high temperatures. A broad diversity of high affinity (100 pM–10 nM), developable binders was directly isolated, with developability evaluated for most assays using the isolated VHHs, rather than fused to Fc, which is customary. This represents the first systematic developability assessment of isolated VHH molecules.

### ARTICLE HISTORY

Received 25 March 2025  
Revised 29 May 2025  
Accepted 30 May 2025

### KEYWORDS

Antibody; camelid; drug discovery; phage display; single domain; VHH; yeast display

## Introduction

Antibody therapeutics have witnessed remarkable advances in recent years, with a substantial expansion of the palette of therapeutic antibody formats explored, including traditional antibodies, antibody fragments, various bi- and multi-specifics, antibody-drug conjugates, and single-domain antibodies known as VHHs (variable domains of Heavy-chain-only antibodies, also known as nanobodies),<sup>1–4</sup> which may also be incorporated into any of these formats. VHHs were originally derived from the heavy chain-only antibodies naturally found in camelids.<sup>5</sup> Their V regions are characterized by small size (~15 kDa) and binding loops similar to variable single-domain (vNAR) antibodies,<sup>6</sup> with unique properties that include binding epitopes with flat surfaces,<sup>7</sup> formation of cavities important in anti-hapten binding,<sup>8</sup> and enzyme inhibition.<sup>9</sup> The small size of the molecules facilitates tissue penetration and access to sterically hindered and cryptic epitopes. Together, these desirable properties have sparked a growing interest in their use as therapeutics, with many applications in various disease areas, including cancer<sup>10–12</sup> immune disorders<sup>13–15</sup> infectious diseases<sup>16–23</sup> and cardiovascular disorders.<sup>24</sup> Although fused to Fcs in their natural forms, most therapeutic VHHs, such as sonelokimab and

vobarilizumab, are composed of humanized VHH domains connected by short amino acid linkers, with different VHH domains conferring different properties, such as half-life extension, without FcRn-dependent recycling.


In addition to their structural versatility, VHH domains offer several advantages over traditional IgG antibodies in therapeutic development. Their single-domain nature eliminates the need for heavy and light chain pairing, which is necessary for Fab and scFv, simplifying library design, allowing deep next-generation sequencing (NGS), and improving display efficiency, with more uniform expression and folding in heterologous systems. Natural VHHs (i.e., those derived from camelids) exhibit exceptional thermal and chemical stability under harsh conditions, including elevated temperatures, extreme pH, high pressure, and protease exposure.<sup>25</sup> They also have enhanced solubility and can be produced at high yields in microbial expression systems, significantly reducing production costs.<sup>26,27</sup> VHHs are highly amenable to modular engineering, making them ideal for bispecific or multispecific antibody formats. Their compact structure allows for tandem fusion of multiple VHHs or linkage to Fc or Fab domains without compromising biophysical properties, enabling constructs with improved manufacturability, tailored half-lives,

**CONTACT** M. Frank Erasmus ✉ [mfrank.erasmus@iqvia.com](mailto:mfrank.erasmus@iqvia.com); Andrew R. M. Bradbury ✉ [andrew.bradbury@iqvia.com](mailto:andrew.bradbury@iqvia.com) Specifica LLC, An IQVIA Business, Santa Fe, NM, USA

\*equal contributors.

<sup>#</sup>Current address: Institute for Protein Innovation, Boston, USA.

<sup>†</sup>Current address: Sanofi, Cambridge, USA.

 Supplemental data for this article can be accessed online at <https://doi.org/10.1080/19420862.2025.2516676>

© 2025 The Author(s). Published with license by Taylor & Francis Group, LLC.

This is an Open Access article distributed under the terms of the Creative Commons Attribution-NonCommercial License (<http://creativecommons.org/licenses/by-nc/4.0/>), which permits unrestricted non-commercial use, distribution, and reproduction in any medium, provided the original work is properly cited. The terms on which this article has been published allow the posting of the Accepted Manuscript in a repository by the author(s) or with their consent.

and reduced steric interference compared to traditional IgG-based designs.<sup>28</sup>

The most common method for generating therapeutic VHHs involves immunization, followed by humanization and further optimization to enhance affinity and developability.<sup>29</sup> However, many VHH advantages tend to be reduced or eliminated when VHHs are humanized by replacing key framework residues with corresponding amino acids from the most similar human germlines,<sup>30,31</sup> requiring extensive further optimization for VHHs used as therapeutics. As an alternative approach, naïve VHH display libraries have been created. These fall into two categories: 1) those derived from naturally occurring, non-immunized camelid VHHs<sup>32–35</sup> which also require subsequent humanization, and 2) more advanced synthetic libraries.<sup>36–41</sup> A primary challenge in using naturally occurring VHHs for therapeutic applications, whether from immunization or display libraries, lies in the need for downstream optimization after discovery. Despite their similarity to human antibodies, VH3–23 in particular, camelid VHHs are routinely humanized to ensure their sequences more closely resemble those of human molecules, with subsequent optimization to maintain solubility, affinity, stability, immunogenicity profiles and “VHH-like” characteristics.<sup>42–49</sup> In addition to humanization and aggregation challenges, downstream purification is another significant obstacle in therapeutic VHH development. Traditional methods for purifying antibodies, such as protein A affinity chromatography, rely on interactions with the Fc domain. However, these methods cannot be applied directly to isolated camelid VHHs as they bind to protein A only sporadically in the absence of an Fc domain, necessitating alternative purification strategies.<sup>50,51</sup>

Here, we present a new humanized VHH phage display library designed to develop VHH-based therapeutic antibodies directly, reducing the need for downstream affinity maturation, developability optimization or humanization. Inspired by our previously published scFv<sup>52</sup> and Fab/scFab<sup>53</sup> Generation 3 libraries, our approach involves harnessing the potential of existing clinical-stage humanized VHH scaffolds by introducing defined HCDR1 and HCDR2 (Note: CDRs are referred to as HCDRs when they are in, or from, human VH genes; once they have been inserted into VHHs they are referred to as CDRs) sequences, purged of sequence liabilities, from human VH3 germline gene family antibodies and combining them with a large diversity of human HCDR3 sequences ( $>10^8$ ), from CD19<sup>+</sup> IgM B cells, to generate the final library. IgM cells were chosen as an HCDR3 source since naïve IgM B cells have far greater diversity than IgG B cells, which are immune, and are far less effective when used as phage antibody library sources.<sup>54</sup>

The initial design (V1) was further improved (V2) by reducing poly-tyrosine stretches in the CDR3 and selecting CDR1 and CDR2 sequences with increased thermal stability. We conducted selections from the V1 and V2 libraries against interferon  $\alpha$ -2 (IFNa2), and a panel of other targets, using phage and yeast display,<sup>55</sup> yielding numerous binders with affinities in the subnanomolar and low nanomolar range, demonstrating our ability to discover high-affinity (under 10 nM) VHHs directly. The developability properties of isolated

VHH domains generated from the two libraries against IFNa2, were examined in depth. Both V1 and V2 yielded a significant number of developable molecules, with the V2 population showing enhanced properties compared to V1 in several developability assays. This work demonstrates that carefully designed VHH libraries can produce liability-free, humanized VHHs that can be easily purified via protein A without additional tags.

## Results

### Scaffold selection

To develop a semisynthetic library for therapeutic VHHs, we began by choosing suitable therapeutic scaffolds into which complementary-determining region (CDR) diversity could be grafted. Key considerations for creating a viable VHH library included: 1) prioritizing already humanized and optimized therapeutic VHH sequences to reduce immunogenicity, biophysical issues, and the need for downstream humanization; 2) ensuring selected framework and CDR sequences would allow binding to *Staphylococcus aureus* protein A (SpA) without relying on the Fc domain, simplifying purification without the need for additional tags; and 3) intentionally avoiding idiosyncratic sequences to enhance generalizability across the library, ensuring broader applicability and robustness.

We hypothesized that analyzing therapeutic VHHs, either approved for human therapy or in clinical trials, could help us refine our selection process, as these were expected to be already optimized for desired traits and we had successfully used this approach in our previous scFv<sup>52</sup> and Fab<sup>53</sup> libraries. We focused on a set of nine different VHHs: caplacizumab, envafolimab, gontivimab, isecarosmab\_1, ozoralizumab\_1, ozoralizumab\_2, sonelokimab\_1, sonelokimab\_3, and vobari-lizumab\_1. We first analyzed the sequences for the presence or absence of amino acid residues required for VH3-mediated SpA interaction as described previously in the literature (Supplementary Figure S1).<sup>50,51</sup> Of the nine VHH analyzed, seven were expected to bind SpA, likely due to intentional engineering for improved manufacturability or as a consequence of the humanization process, which closely resembles human VH3s known to bind SpA (Table 1). The two VHHs not binding SpA were excluded from further examination.

Next, we searched for non-canonical cysteine pairs often found in camelid VHH between CDR3 and either CDR1 or CDR2.<sup>56–58</sup> Since the grafted CDRs were designed to exclude cysteines (see next section on diversity design), we reasoned we should exclude scaffolds containing these additional stabilizing disulfide bridges, particularly since unpaired cysteines would likely compromise the CDR filtering process (as described in Figure 2a, see below). Notably, among the analyzed set, only envafolimab exhibited a non-canonical disulfide bond between CDR1 and CDR3 (Table 1; Supplementary Figure S1). As our plan was to obtain framework 4 sequences from human B cells (see the following section on diversity design), we ruled out VHHs with non-canonical sequences in this region (for instance, the ozoralizumab VHHs). Finally, we noted that all of these therapeutic VHHs, except envafolimab (which was

**Table 1.** List of therapeutic VHH domains analyzed for clinical status, sequence-based protein a binding prediction, presence of canonical cysteines only (upstream of CDR1 and CDR3), canonical framework 2 (arginine replaced for leucine to abolish light chain binding), and canonical framework 4 (presence of tryptophan at the first position).

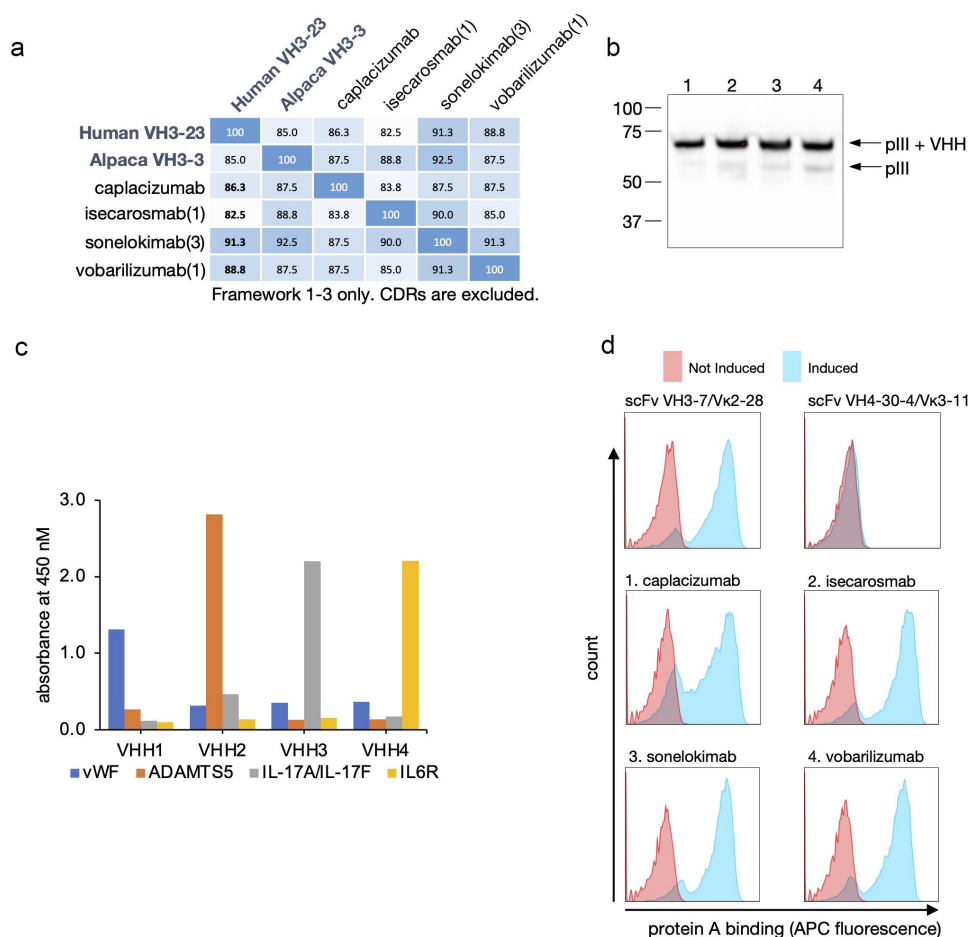
| Name             | Target          | Status               | Protein A binding | Canonical Cysteines | Canonical FWK2 | Canonical FWK4 |
|------------------|-----------------|----------------------|-------------------|---------------------|----------------|----------------|
| caplacizumab*    | vWF             | Approved             | Yes               | Yes                 | Yes            | Yes            |
| envafolimab      | PD-L1           | Phase-III - active   | Yes               | No                  | Yes            | No             |
| gontivimab       | F protein (RSV) | Phase II - abandoned | No                | Yes                 | Yes            | Yes            |
| isecarosmab_1*   | ADAMTSL5        | Phase I - active     | Yes               | Yes                 | Yes            | Yes            |
| ozoralizumab_1   | TNFr            | Phase III - active   | Yes               | Yes                 | No             | No             |
| ozoralizumab_2   | HSA             | Phase III - active   | Yes               | Yes                 | No             | No             |
| sonelokimab_1    | IL17F           | Phase II - active    | No                | Yes                 | Yes            | Yes            |
| sonelokimab_3*   | IL17A/F         | Phase II - active    | Yes               | Yes                 | Yes            | Yes            |
| vobarilizumab_1* | IL6R            | Phase II - active    | Yes               | Yes                 | Yes            | Yes            |

Domains isecarosmab (2), sonelokimab (2), and vobarilizumab (2) have identical sequences to domain ozoralizumab (2) and have been omitted from the table. Asterisks mark the domains chosen to serve as scaffolds.

excluded because of the non-canonical cysteine) had between zero and three amino acids corresponding to the classic “hallmark residues” (V37F, G44E, L45R and W47A) considered to enhance VHH stability in the absence of a light chain<sup>42,48,51</sup> (Supplementary Figure S1). The four therapeutic VHH domains with the greatest number of hallmark residues (caplacizumab, isecarosmab\_1, sonelokimab\_3, and vobarilizumab\_1) were consequently selected as scaffolds for the

library. For ease of reference, we will call these Lib1, Lib2, Lib3, and Lib4, respectively. Framework regions 1–3 of these VHHs exhibited an identity of 83.8% to 91.3% to one another, and 82.5% to 91.3% (Figure 1a) to the human VH3–23 germ-line, which is comparable to the 85% identity noted between VH3–23 and alpaca VH3–3.

To check the ability to display these VHH scaffolds in phage, their genes were inserted into pDAN5,<sup>59</sup> and phage



**Figure 1.** Scaffolds. (a) Percent identity comparisons between frameworks 1–4 across the four therapeutic VHH, human germline VH3–23, and alpaca germline VH3–3. (b) An immunoblot (anti-SV5) detecting display of the four selected therapeutic VHH on phage particles. The upper band corresponds to the phage pIII linked to the VHH domain, while the lower band represents a degradation product comprising pIII alone. 1-caplacizumab, 2-isecarosmab, 3-sonelokimab, 4-vobarilizumab. (c) Results from a direct ELISA where the four therapeutic phage-VHH were probed against their corresponding immobilized antigens. (d) Flow cytometric analysis of yeast displaying two control scFv, binding (VH3–7) and not binding (VH4–30) SpA, and the four therapeutic VHHs. Yeast cells (blue: induced for display; pink: not induced) were incubated with protein A conjugated to the APC fluorophore. A right shift in fluorescence along the X-axis reflects protein A binding.

particles generated using M13KO7 helper phage. Remarkably high display levels were detected in Western blots of the phage particles when probed with anti-SV5 tag antibodies (Figure 1b). Proper folding and functionality on phage were validated by confirming binding against the four targets recognized by the parental VHs (Figure 1c). Lastly, we displayed the four VHH molecules in yeast and confirmed SpA binding (Figure 1d).

### **CDR1 and CDR2 replicated diversity from the human VH3 family**

To develop therapeutic VHs, we aimed to enhance “human-ness” while minimizing the number of poorly behaved VHs caused by biophysical liabilities or CDRs incompatible with the VHH structure. Based on the high similarity between the VHH therapeutic frameworks (82.5% to 91.3%) and human VH3–23, we reasoned that inserting HCDR1 and HCDR2 sequences from the VH3 family repertoire into the VHH CDR1 and CDR2 sites should provide structural compatibility and maximize humanness, while eliminating CDR sequences with sequence liabilities (Table 2; Supplementary Figure S2) should minimize biophysical liabilities. Figure 2a illustrates the design strategy: liability-free, human VH3 family, HCDR1s and HCDR2s were inserted into the four clinical VHH frameworks, while HCDR3 was sourced directly from CD19+ B-cells, extracted from donor LeukoPaks via PCR using IgM-specific primers.

The CDR1 and CDR2 sequences originated from an internal NGS database of the naïve and memory B cells of 40 individuals, split into four datasets of 10 each. Each CDR used in the library had to be found in one of the four datasets, and two additional datasets, which could be either one of the remaining three internal datasets, or public datasets such as OAS.<sup>60</sup> This ensured the library was composed of commonly used CDRs, and excluded rare CDRs. We also discarded CDRs containing a known set of sequence liabilities (Table 2) which included arginines, because of their tendency to increase polyreactivity,<sup>61,62</sup> except at the last position of HCDR2 which facilitates protein A binding. CDR sequences were also removed if they exhibited high hydrophobic scores according to a previously published index<sup>63</sup> or contained three consecutive aromatic residues, to minimize the presence of exposed hydrophobic areas<sup>64,65</sup> (Table 2).

The final counts of HCDR1s and HCDR2s were 22,062 and 16,704, respectively. Logo plots (Figure 2b) displaying the IGHV3 diversity before (Figure 2b, top) and after liability removal (Figure 2b, middle) revealed notable similarities between these CDRs and the circulating repertoire of alpaca VHH<sup>56</sup> (Figure 2b, bottom). In most positions, the dominant amino acid in the alpaca repertoire aligned with the dominant amino acid found in the human CDRs. At positions 3 and 6 of CDR2, the most common alpaca amino acid ranked as the second most frequent in the designed CDRs. Some differences stemmed from the elimination of sequence liabilities,<sup>61</sup> including the absence of arginines at position 2 of CDR1 and tandem aspartates at positions 5 and 6 (relating to the aspartate isomerization motif).

The designed CDR1 and CDR2s were synthesized as array-based oligo pools, and single-CDR yeast display libraries were created for each scaffold at the corresponding CDR position (Figure 2c). In these libraries, the scaffold and two CDRs are parental, while diversity is introduced at only one site (see Figure 2c). Each library had >10<sup>6</sup> transformants, providing extensive over-representation of the intended diversity. After induction, cells were purified using magnetic-assisted cell sorting (MACS) with SpA microbeads. Proper folding is necessary for SpA binding, as this interaction relies on conformational epitopes. This process helps eliminate poorly folded molecules and those with frameshifts, stop codons, or other deleterious mutations which may occur in oligo synthesis.

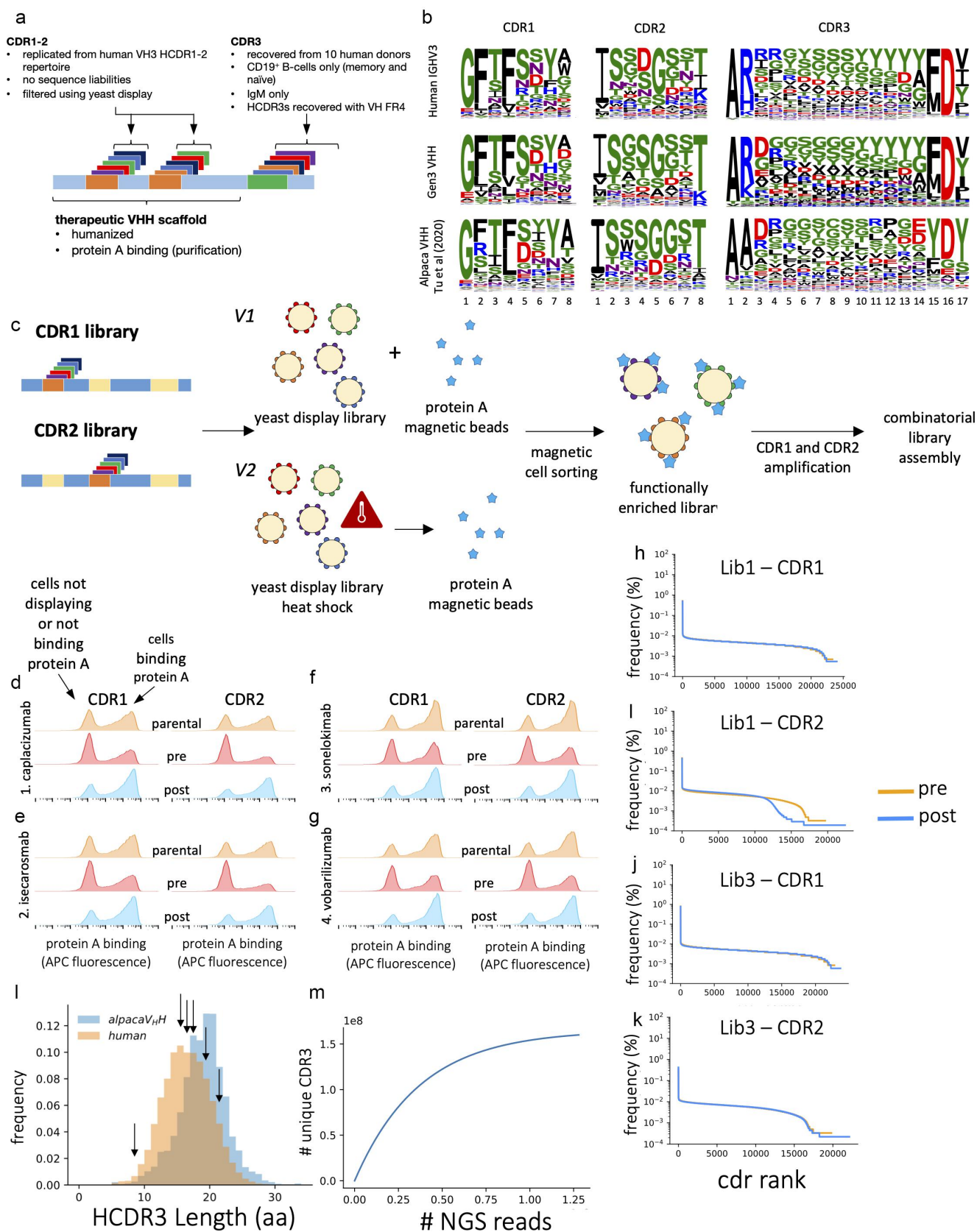
As anticipated, filtering the single-CDR libraries resulted in increased SpA binding relative to pre-SpA filtering, indicating enhanced protein A binding functionality for the repertoire (Figure 2d–g), in line with the therapeutic parental displayed on the yeast surface. To determine if filtering led to a loss or significant bias of diversity, we sequenced libraries 1 and 3 at the single CDR level. Representative plots of their behaviors are reported in Figure 2h–k. Notably, for Lib1 CDR1 (Figure 2h) and Lib3 CDR1–2 (Figure 2j–k), limited loss of diversity or bias was detected between pre- and post-filtering, and the filtering effectively removed nonfunctional molecules (frameshifted/stop codon). In contrast, for CDR2 of Lib1, a significant frequency drop was noted after approximately 13,000 CDRs, suggesting that roughly 3,000 designed CDRs were incompatible with this scaffold (Figure 2i). The diversity from each filtered library was recovered by PCR and used to assemble the final library along with HCDR3 for the CDR3. CDR1 and CDR2 diversities in the final combined library for this “V1” design (more on “V2” design below) can be seen in Supplementary Table S1.

### **CDR3 diversity and library assembly**

HCDR3 is the primary locus of antibody diversity and binding activity.<sup>66</sup> While it is possible to utilize array-based oligo synthesis to generate specific sequences<sup>61</sup> for HCDR1 and HCDR2, HCDR3 diversity far exceeds the current limit of ~10<sup>6</sup> unique sequences in this technology. Therefore, we opted to utilize natural human HCDR3 sequences including the IGHJ gene diversity of framework 4, obtained directly from >10<sup>9</sup> CD19+ IgM+ B cells collected from 10 donors. This approach maintains the benefits of being human-derived, offers a large diversity surpassing 10<sup>8</sup>, and significantly reduces the number of cysteines (see below) and other nonfunctional sequences.<sup>67</sup>

Differences exist between human HCDR3s and camelid CDR3s (Figure 2b,l). Notably, alpaca CDR3s exhibit a broader length distribution, with a peak at 19 amino acids, whereas human HCDR3s display a narrower distribution, peaking at 15 amino acids (see Figure 2l). The core of the sequences in both species primarily consists of polar amino acids. Over 50% of alpaca CDR3s contain a single cysteine that forms a disulfide bridge with external cysteines, mainly in other CDRs, a feature we aimed to exclude from our library. Additionally, an alanine in alpacas substitutes for





**Figure 2.** CDR diversity. (a) Schematic of the library design highlighting its three primary components: human diversity without sequence liabilities (CDR1–2), natural human diversity sourced from CD19<sup>+</sup> IgM<sup>+</sup> B-cells (CDR3), and humanized therapeutic VHH scaffold (frameworks 1–3) capable of binding protein A for easy purification. (b) Sequence logos representing natural human VH3 family HCDR diversity (top), human CDRs used to build the library (middle), and natural alpaca (*vicugna pacos*) VHH diversity (bottom). The height of each letter indicates the frequency of an amino acid at each position. For visual simplicity, only CDR3s with 17 amino acids are shown. (c) A schematic outline demonstrating the yeast filtering method applied to CDR1 and CDR2 in libraries v1 and v2 (heat shock), aimed at selecting CDR functionality and protein A binding in the libraries. (d–g) Histograms show the flow cytometry analysis of yeast cells featuring the four parental VHH (orange), CDR1, and CDR2 libraries prior to protein A filtering (red) and after filtering (blue). The left peaks indicate cells that either do not display the VHH or present molecules that cannot bind protein A. The right peak represents cells displaying a VHH that can bind protein A. (h–k) NGS sequencing results for single CDR libraries before (orange) and after (blue) protein A filtering. (l) Distribution of CDR3 lengths in the library and in the alpaca repertoire are depicted. Therapeutic VHH CDR3 lengths analyzed in the work are indicated with arrows. (m) Species accumulation plot from NovaSeq analysis reveals the diversity of HCDR3 in the cloned library.

**Table 2.** Liabilities purged from CDR1–2 oligo sets during gene synthesis.

| Liability          | Rule   |
|--------------------|--|
| Glycosylation      | NxS or NxT, where x is not P                   |
| Asn Deamidation    | NG, NN, NS, NT, GNF, GNY                       |
| Asp Isomerization  | DG, DS, DD                                     |
| Cysteine           | Presence of Cys                                |
| Hydrolysis         | DP   |
| Gln deamidation    | QG   |
| Aromatic Trimer    | Three aromatics in tandem (His, Phe, Trp, Tyr) |
| Polyreactivity     | GGG, RR, VV, VG, WW, WxW, YY                   |
| Positive Charge    | Charge >1 at pH 7                              |
| Arginine           | Presence of Arg (except R57)                   |
| Hydrophobic        | Hydropathy <0 (Parker scale)                   |
| Protein A Affinity | Position 57 is not T, K, or R                  |

**Table 3.** Number of transformants for each VHH sublibrary.

| Sublibrary                | #Transformants                         |
|---------------------------|--|
| Library 1 (caplacizumab)  | $4.5 \times 10^9$                      |
| Library 2 (isecaromab)    | $6.4 \times 10^9$                      |
| Library 3 (sonelokimab)   | $6.6 \times 10^9$                      |
| Library 4 (vobarilizumab) | $5.4 \times 10^9$                      |
| <b>Total</b>              | <b><math>2.3 \times 10^{10}</math></b> |

the typical arginine at position 2 in humans (as shown in Figure 2b). There is also a higher frequency of negatively charged amino acids at positions 3 and 12. While the presence of the extra disulfide contributes to structural integrity, it remains uncertain what additional roles, if any, these differences have in VHH structure and function.<sup>58</sup> Finally, some human antibody HCDR3s have long strings of consecutive tyrosines, encoded by JH6, which are absent from alpaca CDR3s.

The HCDR3 fragments from 10 donors were combined with the filtered CDR1 and CDR2 fragments by overlap PCR. These were then digested with BssHII and NheI, ligated into the phagemid vector pDAN5 digested with the same enzymes, and transformed into *E. coli* TG1 cells, resulting in a total of  $2.3 \times 10^{10}$  transformants (see Table 3). Using NovaSeq we extracted CDR3 sequences using consensus matching and determined a total unique CDR3 diversity of  $1.6 \times 10^8$  from  $1.3 \times 10^9$  usable reads (refer to Figure 2m).

## V2 design – tyrosine removal and thermal enhancement

### Reducing poly-tyrosine motifs in CDR3

One striking difference between human antibody HCDR3s and VHH CDR3s is the presence of strings of consecutive tyrosines in some antibody HCDR3s encoded by antibodies using the JH6 gene (Figure 2b). It has been proposed that high numbers of tyrosines may cause unwanted polyreactivity in VHH<sup>68</sup> and antibodies,<sup>69</sup> although high numbers of tyrosines are also found in specific antibodies.<sup>62,69</sup> To understand whether reducing consecutive tyrosines may improve polyreactivity, we developed a method (Figures 3a,b) to deplete HCDR3s containing poly-tyrosines encoded by the JH6 gene, based on annealing of JH6-specific oligo probes to HCDR3 mRNAs encoding poly-tyrosine and cleavage of the RNA/DNA hybrids with RNase H (see Supplementary Table S2). CH1 reverse transcription primers were subsequently used to generate cDNA from the

remaining poly-tyrosine-depleted HCDR3 diversity for subsequent PCR using framework 3 and JH primers. The method successfully reduced poly-tyrosines in the treated population as demonstrated by NGS (Figure 3c,d).

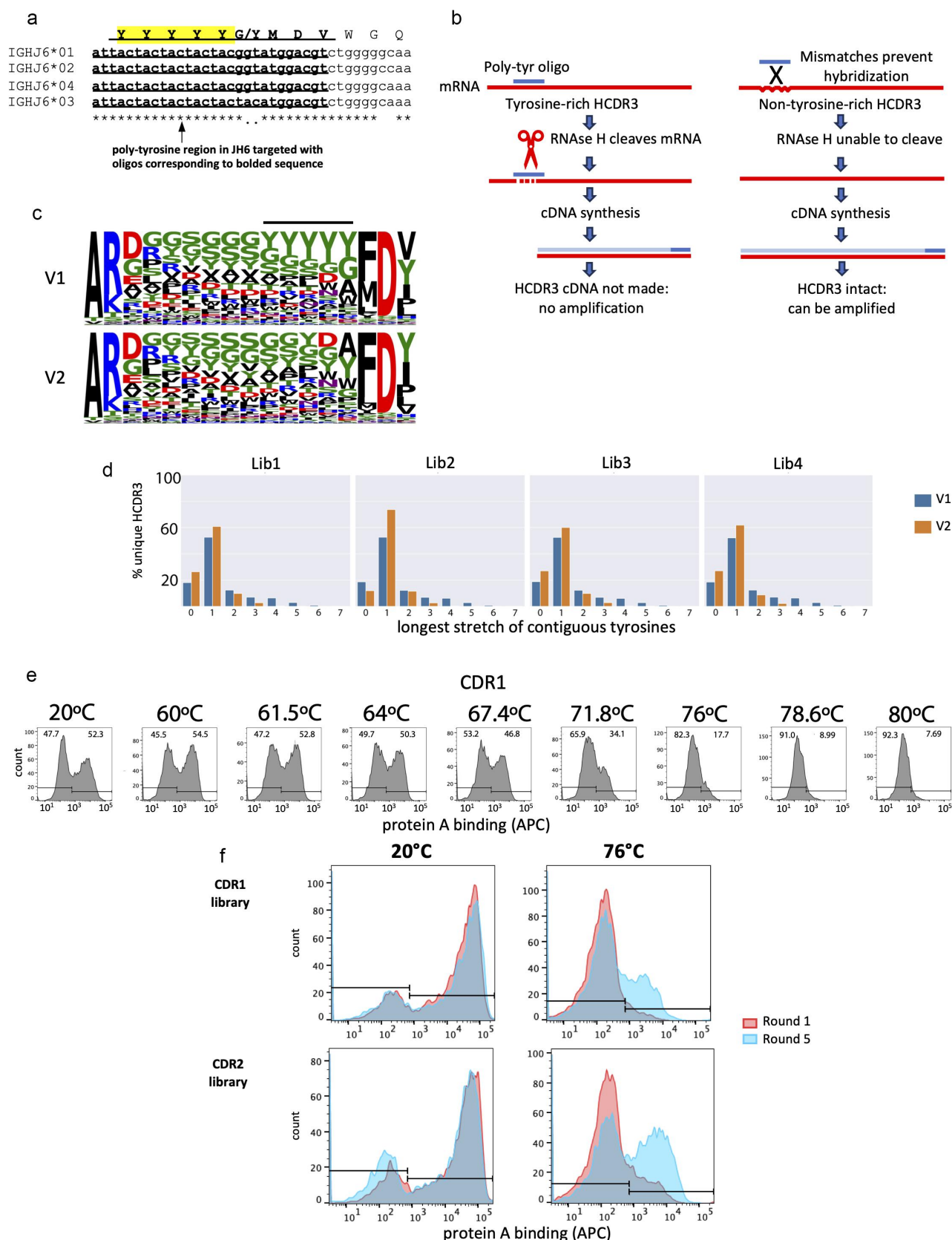
### Filtering for heat tolerant HCDR1–2 diversity

To create libraries with enhanced heat tolerance, the single-CDR libraries containing CDR1 and CDR2 diversity underwent heat shock for 20 minutes (Figure 2c) at temperatures between 60°C and 80°C to evaluate their fold stability (Figure 3e and Supplementary Figure S3). Binding to protein A was used to select correctly folded molecules, reflecting the conformational epitope recognized. We determined optimal heat shock to be 76°C, which yielded the greatest SpA-binding diversity post-filtering, and conducted five rounds of filtering using protein A magnetic beads (MACS). A notable increase in thermal stability was observed at the end of the process (Figure 3f and Supplementary Figures S4, S5 for each of the sub-libraries), though with a sizable decline in CDR1–2 diversity (Supplementary Table S1, rightmost column).

In summary, the final library designs for versions 1 (V1) and 2 (V2) combined human VH3 germline matched CDR1–2 sequences free of sequence liabilities, with native human CDR3s. V2 was designed to reduce polyreactivity and improve thermostability by the depletion of HCDR3 poly-tyrosine stretches and enrichment of thermostable HCDR1–2 sequences.

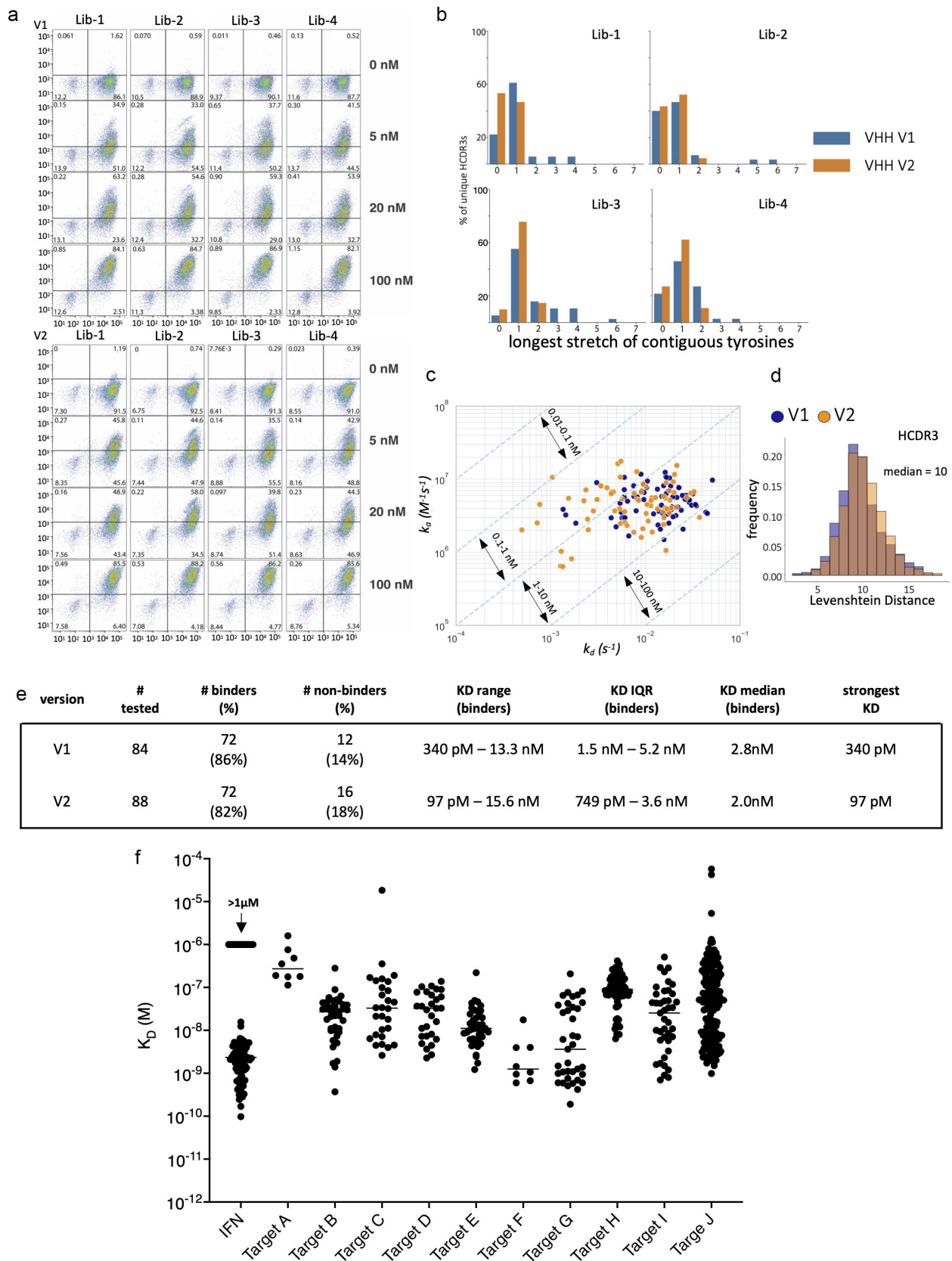
### Library performance

To assess the performance of the two VHH libraries from V1 and V2 iterations, we applied a selection strategy comprising two rounds of phage display followed by yeast display. Using a fixed amount of magnetic streptavidin beads coated with biotinylated antigen, we aimed to favor binder diversity over affinity in the phage display step, while fine-tuning of affinity was achieved during yeast sorting with decreasing amounts of target (100, 20, and 5 nM of biotinylated antigen). One target, IFNa2, was examined in depth, while performance against an additional 10 masked targets was limited to affinity. For the IFNa2 selection each of the eight sub-libraries was maintained separately and incubated with antigen-coated beads. After washing, the phage that remained bound to the beads were eluted using HCl and rescued with *E. coli*. Following two rounds of phage selection, VHH were PCR amplified and cloned into the yeast display system by homologous recombination. Yeast cells displaying VHH underwent selection rounds against gradually decreasing concentrations of biotinylated target. After the final yeast sorting round, we examined the population by flow cytometry for display (anti-SV5-tag PE) and antigen binding (streptavidin-Alexa 633) at different antigen concentrations (0, 5, 20, and 100 nM) for both the V1 and V2 final populations (Figure 4a). Binding was observed down to the lowest concentration tested (5 nM) and no signal was observed against the secondary reagents (0 nM). Sanger sequencing of the final sorted populations revealed the



**Figure 3.** Iterative improvement of the Gen3 VHH design. (a) Poly-tyrosine regions in different human JH6 alleles. The highlighted area was targeted for degradation during cDNA synthesis. (b) A schematic illustration of targeted degradation for poly-tyrosine regions employing specific primers during HCDR3 retrieval from CD19+ cells. After poly-tyrosine encoding mRNAs are degraded by RNAse H, the remaining mRNAs are converted into cDNA using a CH1 primer and HCDR3s amplified from the intact cDNA. (c) Weblogo representations of HCDR3 for v1 (top) and v2, tyrosine depleted, (bottom) libraries. (d) NGS analysis of the naïve library from V1 (blue) and V2 (orange) populations showing the number of contiguous tyrosines. (e) Representative analysis of individual CDR libraries (e.g., CDR1 diversity with constant frameworks alongside other CDRs from library 1) that underwent heat shock and were evaluated across a range of temperatures, identifying an optimal temperature of 76°C. (f) Analysis from sublibrary 1 of the final cloned output shows pre-heat shock (round 1; red) and post-heat shock (round 5; blue) populations, assessed via flow cytometry for CDR1 (top) and CDR2 (bottom).





**Figure 4.** Discovery and affinity characterization of a broad panel of VHH-Fc from V1 and V2 libraries. (a) Flow cytometry analysis of the final sorted population of V1 (top panel) and V2 (bottom panel) binding to varying concentrations of IFN $\alpha$ 2 (top to bottom) for each of the different sublibraries (left to right). Display is detected using anti-SV5 conjugated to PE, while binding is detected with biotinylated IFN $\alpha$ 2 and streptavidin-alexa-633. (b) Sanger sequencing of clones selected against IFN $\alpha$ 2 plotting the number of contiguous tyrosines in the V2 library (orange) compared to V1 (blue). (c) An isoaffinity plot showing on-rate ( $M^{-1}s^{-1}$ ) on the y-axis and off-rate ( $s^{-1}$ ) on the x-axis, with affinities on the diagonal, indicating the wide range of affinities against IFN $\alpha$ 2 for VHH-Fc in both v1 (blue) and v2 (orange) populations. (d) Levenshtein distance of the HCDR3 amino acid sequence among the different binders of v1 (blue) and v2 (orange) populations. (e) Summary of binding statistics and affinities for v1 and v2 selected populations. (f) Affinity plots for IFN $\alpha$ 2 and ten other undisclosed targets.

decreased consecutive tyrosines trend in the V2 population (Figure 4b).

We examined the binding kinetics of VHHs selected against IFN $\alpha$ 2 from the V1 and V2 libraries using surface plasmon resonance (SPR) with the Biacore 8k+ system. We randomly picked 48 colonies from each sublibrary and population (V1/V2) and produced all VHHs with unique CDR3 sequences (see Supplementary Table S3 for population breakdown) based on 100% identity at amino acid level (84 unique CDR3 from a total of 84 unique full-length VHH in V1; 88 unique CDR3 from 88 total unique full-length VHH in V2) for biophysical characterization as dimeric VHH-Fcs or monomeric isolated VHH domains. The VHH-Fcs were immobilized on the SPR CM5 chip, which had a polyclonal anti-human Fc linked to its gold nanolayer. Then, IFN $\alpha$ 2 was injected over the VHH-Fc. We calculated the on-rate and off-rate constants via a first-order kinetic model for “one-to-one binding”. The isoaffinity plot (Figure 4c) indicates that the majority of VHH-Fcs from the V1 and V2 populations exhibited similar kinetics (2.8 nM vs. 2.0 nM median affinity, respectively). The V2 population yielded antibodies with the slowest off-rates and the highest affinity binders. Both populations resulted in 72 highly diverse (Figure 4d) binders ( $\leq 1 \mu\text{M}$ ), amounting to 82–86% of all tested VHH-Fc, with affinities ranging from 97 pM to 15.6 nM, as shown in Figure 4e. All four sublibraries produced binders, but their performance varied (see Supplementary Table S3–S5; and Supplementary Figures S6–S7 for SPR sensorgrams). VHHs were also selected against an additional 10 undisclosed targets, comprising soluble proteins or the external domains of membrane proteins. Results in Figure 4f show a broad range of affinities, with 90% of campaigns yielding at least one high affinity binder ( $< 10 \text{ nM}$ ), and 50% at least one subnanomolar binder.

### VHH and VHH-Fc developability for the V1 and V2 populations

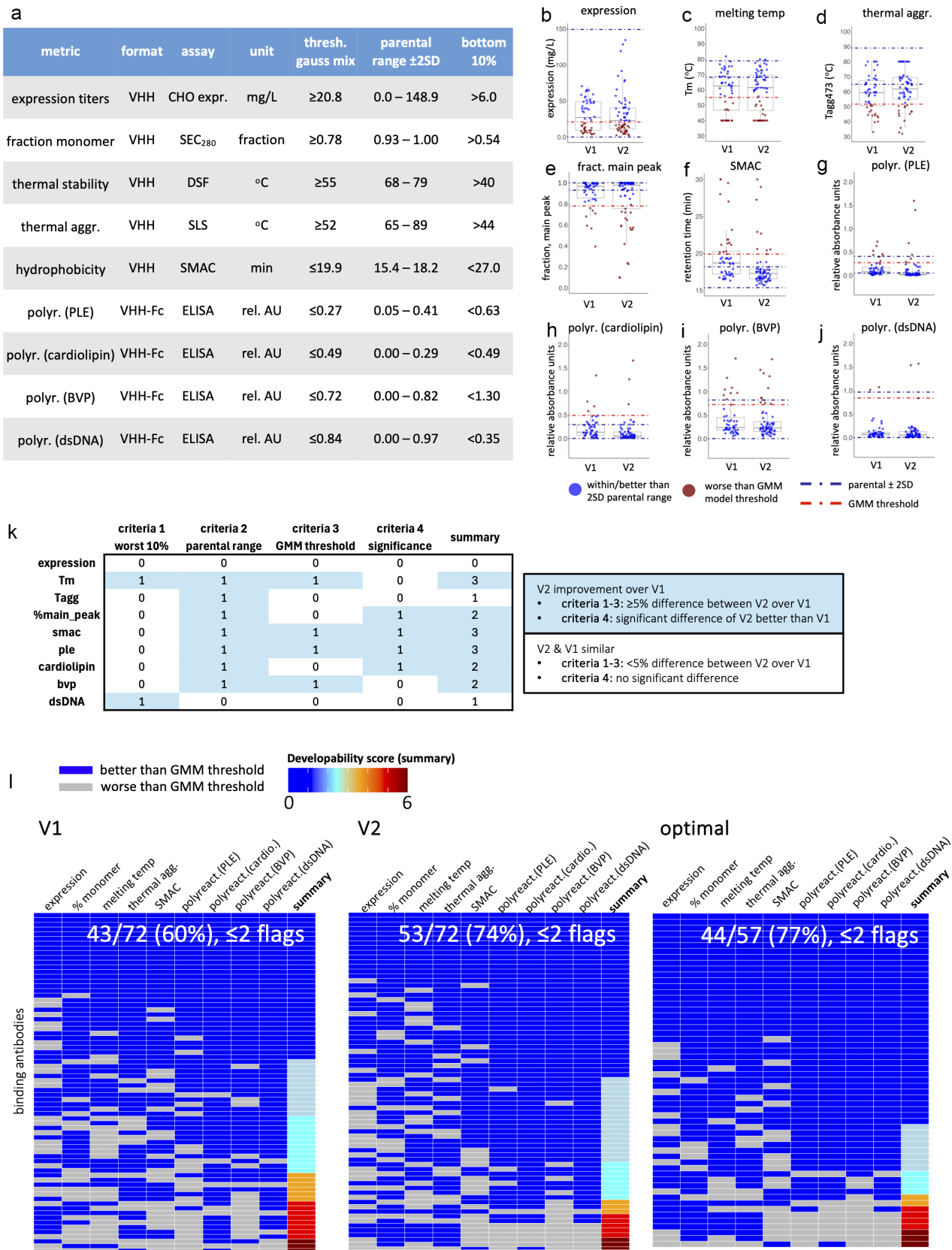
To ensure a comprehensive assessment of developability properties, we chose experimental assays that span a wide array of distinct (uncorrelated) characteristics (Supplementary Figure S8), including polyreactivity, thermal stability, hydrophobicity, and expression yield. While developability assessments usually focus on the VHH-Fc format, we reasoned the presence of the Fc domain may “buffer” intrinsic VHH developability properties, making them appear more developable than they would otherwise be as isolated VHH domains. Since most therapeutic VHHs to date do not have Fc domains, developability of the VHH alone, where possible, may be more clinically relevant. While many of the assays are amenable to isolated VHH domain testing, testing polyreactivity requires the detection of VHH binding, and the only way to detect binding of (protein A binding) VHH domains lacking a tag or Fc domain is with protein A. While protein A does not appear to prevent antigen binding by Fab domains,<sup>50</sup> it is not clear whether this is also true of the interaction with VHHs, particularly the humanized versions used here. Supplementary Figure S9 shows that protein A appears to interfere with specific binding (at least to IFN $\alpha$ 2) and so cannot be used to detect VHH polyreactive binding as it is expected to inhibit the interaction.

When IFN $\alpha$ 2 was immobilized, protein A-HRP failed to detect VHH binding, despite confirmed interaction in the VHH-Fc format by SPR. In contrast, an anti-IFN $\alpha$ 2 IgG control was detected, likely due to stronger Fc-mediated protein A-HRP binding compared to the VH region (Supplementary Figure S9A). Similarly, anti-VHH secondary antibodies showed limited detection (Supplementary Figure S9B). ELISA detection of specific binding was restored when using anti-Fc secondary probes in the VHH-Fc format (Supplementary Figure S9B, right side), as well as polyreactive signals for positive controls (C10, IgG controls; Supplementary Figure S10), while non-polyreactive antibodies (e.g., Ibalizumab) remained negative. Due to the challenges with detection using protein A or anti-VHH, we used the VHH-Fc fusion format with anti-Fc secondaries, to test polyreactivity (e.g., PLE, cardiolipin, BVP, dsDNA) covering a range of physicochemical properties,<sup>70,71</sup> but not those other developability assays that could be directly assessed on the isolated VHH domains.

Given the scarcity of established reference thresholds for VHH developability in existing literature, we formulated our own thresholds (Figure 5a) and utilized various statistical methods to compare selected VHHs to their therapeutic parental scaffolds, as well as VHHs from the V1 and V2 designs. The methods employed included: 1) the lowest 10% threshold based on the V1 design population, 2)  $\pm 2$  standard deviations of the parental mean (parental range), 3) Gaussian mixture model (GMM) for binary (e.g., “favorable” vs. “unfavorable” developability) classification, and 4) Mann-Whitney U significance test.

The thresholds outlined above allowed us to assess performance between the two design strategies and set cutoffs based on population distributions. Figure 5b–j (see Supplementary Figure S11 for individual libraries) shows the performance of VHHs from both V1 and V2 populations using the GMM model threshold and the parental range (mean  $\pm 2$  SD) across all developability assays. Using the bottom 10% of the V1 population to define an assay threshold we find VHHs from the V2 population have equal or better developability values than VHHs from the V1 population for all assays (Supplementary Figure S12A). For both populations, the majority of VHHs were within or better than the parental (therapeutic) range for expression titers and across the polyreactivity assays (Supplementary Figure S12B). A larger proportion of VHHs from the V2 population were within the parental range relative to V1 across most assays.

Most VHHs surpassed the GMM threshold across all the developability assays (Supplementary Figure S13A) with 60–97% of VHHs surpassing this threshold for both populations. The V2 population VHHs showed an elevated proportion of antibodies improved relative to V1 population for the GMM thresholds. The Mann-Whitney U test revealed significant improvement of V2 over V1 in 4 of 9 assays (Supplementary Figure S13B), including % main peak purity by HPLC-SEC, retention time with SMAC, polyreactivity with Poly-L-Glutamic acid and polyreactivity with cardiolipin. None of the significance tests showed improvements of V1 over the V2 population.



**Figure 5.** Developability of VHH and VHH-Fc from the V1 and V2 libraries. (a) Details of the format (VHH vs. VHH-Fc), assay, and thresholds used to carry out the analysis using 1) GMM, 2) parental range  $\pm 2$  SD of the mean, or 3) bottom 10% threshold. B–J) boxplot of VHH and VHH-Fc across the different developability assays from using parental range (b) Expression titers (mg/L), (c) Melting temperature ( $^{\circ}C$ ), (d) Thermal aggregation at SLS<sub>473</sub> ( $^{\circ}C$ ), (e) Fraction main peak at A<sub>280</sub> with SEC-HPLC, (f) standup monolayer chromatography retention time (min), and polyreactivity (relative absorbance units) with (g) PLE, (h) cardiolipin, (i) BVP, and (j) dsDNA. Boxplots show interquartile range (IQR) from quartile 1 to quartile 3 with the horizontal line showing the median, and whiskers showing the Q1–1.5\*IQR and Q3 + 1.5\*IQR. Each point represents a unique measurement and is color-coded dark red if worse than the GMM threshold or blue if better than GMM threshold. Horizontal lines represent the GMM threshold (red) or  $\pm 2$  SD from parental mean (dark blue). (k) summary of V2 over V1 improvement. If  $\geq 5\%$  antibodies show improvement of V2 population over V1 using criteria 1–3 or if there is significant difference of V2 better than V1 (using criteria 4), boxes are scored + 1 and colored light blue. No difference is indicated with 0 and colored white. Total summary of summed scores shown in last column with those  $\geq 2$  of V2 improvement over V1 colored light blue. l) attrition rate of binders (y-axis) for V1 (left panel), V2 (middle panel), and optimal combination [V1 (lib3)+V2 (lib1,2,4)] across the different developability assays (x-axis).

It is notable that we observed aberrant thermal profiles (Supplementary Figure S14A) for many of the VHHs for which we assigned a  $T_m$  value of 40°C. We speculated that these VHHs were highly aggregated; however, HPLC-SEC profiles showed no correlation between the fraction main peak and melting temperature (Supplementary Figure S14B). Further, many VHHs with aberrant profiles showed strong binding profiles with no correlation of aberrant thermal profiles to affinity (see Supplementary Figure S14C-D).

To evaluate whether V2 design improved over V1 across the assay, we incorporated a score which uses the four criteria outlined above: 1) 10% threshold; 2) parental range; 3) GMM threshold; and 4) statistical Mann–Whitney U test. For criteria 1–3, we provided a score of “1” if the proportion of antibodies in V2 outperformed the proportion of antibodies in V1 by a percentage of  $\geq 5\%$ , and “0” if the percent difference was  $< 5\%$ . For criterion 4, we provided a “1” if V2 was statistically improved in each developability metric over V1 and “0” if there was no significant difference (V1 was not better than V2 in any full comparison). As shown in Figure 5k, V2 outperformed V1 (score  $\geq 2$ ) across 6 of 9 assays, with the remaining 3 assays showing no clear differences between the two versions. Finally, to assess the proportion and total number of therapeutic leads (i.e., those that were developable and bound specifically to the target) that could be selected from each population, we created a score which assigns a “1” if the assay was at or above the GMM thresholds or “0” if it was below the GMM threshold. Individual assay scores along with the summed score are shown in Figure 5l for each population. We observed that 43 of 72 (60%) V1 and 53 of 72 (74%) V2 VHHs exhibited 2 or fewer flags, respectively. This performance trend held even when including non-binding antibodies (Supplementary Figure S15A). Due to differences in the performance of the individual sublibraries (Supplementary Figure S15B) from the V1 and V2 populations, with sublibrary 3 from V2 showing greater polyreactivity than the V1 version, we also included an “optimal combination” library composed of sublibraries 1, 2, 4 from V2 and sublibrary 3 from V1 (Figure 5l, right panel; Supplementary Figure S15A, right panel). This optimal combination of the four sublibraries (using the best performing sublibraries from V1 and V2) showed the highest percentage (77%) of antibodies that exhibited strong specific binding and developability profiles.

## Discussion

Antibody candidates from discovery campaigns for therapeutic purposes should typically have: 1) reasonable affinities (typically stronger is better, except for agonists, where lower affinity may be more effective than higher affinity);<sup>72</sup> 2) broad diversity to engage different epitopes with potentially different biological activities; and 3) good developability properties. We recently described a novel scFv library design<sup>52</sup> which incorporates liability-free natural human CDR diversity directly into therapeutic scaffolds. This new design fulfilled the criteria above, with direct selection of antibodies with affinities as low as 13pM,<sup>73</sup> up to hundreds of different NGS identified clusters,<sup>74,75</sup> and excellent developability properties,<sup>76</sup>

reducing the need to optimize antibodies after selection. This concept was also successfully applied to development of a Fab/scFab library format,<sup>53</sup> and here we describe its application to a new semi-synthetic library design for therapeutic VHH development, which similarly produces a broad diversity of developable VHH leads with significantly higher affinities compared to those from other *in vitro* libraries.<sup>32–38</sup>

Our approach used optimal therapeutic VHH scaffolds and curated, human liability-free HCDR1–2 diversity<sup>52</sup> from the VH3 germline gene family, to boost the clinical potential of selected VHHs. By incorporating a filtration step to eliminate sequences that prevent the binding of protein A, the library provides well-folded, protein A binding molecules that can be directly purified using the same manufacturing techniques used for standard antibodies with an Fc domain, thus simplifying downstream development. We verified this binding was effectively preserved at the individual CDR level (Figures 2d–g) and in the final populations (Figure 3f, red), even following the iterative enhancements of poly-tyrosine removal and heat shock (Figure 3f, blue).

A similar strategy, in which immune camelid CDR3s are combined with humanized VHH scaffolds containing limited developable camelid CDR1 and CDR2 diversity, has been successfully used to generate human-like VHH from immunized animals.<sup>77</sup> The key differences here are that we constructed a naïve library using human HCDR3s from naïve, rather than immune sources, and combined this with extensive human HCDR1 and HCDR2 diversity. This allowed us to build a universal naïve library with VHH sequences as close to human as possible, including within the CDRs. Given the strong resemblance of therapeutic VHHs to the human IGHV3–23 family (Figure 1a), we determined that incorporating CDR diversity from human germline families directly into these effective humanized therapeutics would be feasible. This approach avoids the need for specialized techniques, such as *in silico* humanness predictions on camelid diversity, by utilizing HCDR1 and HCDR2 from human VH3 germline family genes, which have been shown to be well-tolerated in humanized VH3 family VHH therapeutic scaffolds. Additionally, the similar sequence patterns of VH3 family and alpaca CDR1–2 (Figure 2b) bolster the argument that CDRs sourced from humans can be effectively integrated into these scaffolds. By carefully selecting individual CDR populations (Figure 2b,c) displayed in yeast while maintaining protein A binding (Figure 2c), we were able to filter out any CDRs preventing correct folding as detected by SpA.

While a commonly used rationale for using VHHs, particularly camelid-based, relates to their longer HCDR3s, conferring unique properties in binding different surfaces (convex, concave, and flat),<sup>25</sup> we have found this length characteristic is not typical for therapeutic VHHs (Figure 2l). Importantly, because it is known longer HCDR3s elevate the risk of poly-reactivity, the use of longer camelid CDR3s may complicate the development of therapeutic VHHs. Furthermore, longer CDR3s often require disulfide stabilization<sup>57</sup> to other CDRs (e.g., CDR1/2, from which we have removed cysteines), and are more relevant in the context of camelid immune systems, not VHH therapeutics, which must be manufactured, stored, and produced for in-patient delivery. While the use of shorter



human HCDR3s may theoretically reduce the potential for broader epitope coverage, this remains unproven and is balanced by reducing developability risks related to long CDR3s and inter/intra-CDR (e.g., CDR1/3) disulfide bonds. Optimal developability is crucial to the successful development of therapeutic VHHs, where developability, specificity, manufacturing, and storage requirements are essential.

Previous studies have shown stretches of poly-tyrosines may lead to unfavorable polyreactive properties in nanobodies<sup>68</sup> and antibodies,<sup>69</sup> although they may also contribute to specificity.<sup>62,69</sup> In the V2 library poly-tyrosine stretches  $\geq 3$  were significantly reduced compared to the V1 library. This was combined with the application of heat shock and protein A binding at the individual CDR level, to create final populations that displayed well on the yeast surface and maintained protein A binding capabilities at higher temperatures, allowing a comparison between the two designs.

VHHs isolated from both libraries showed protein A binding, high diversity, and broad affinity ranges against IFN $\alpha$ 2 and other targets. It was important to confirm that integrating liability-free *human* HCDRs into germline family-matched therapeutic VHH scaffolds maintained favorable binding and developability properties. For both design strategies strong affinity ( $\leq 10$  nM) binders could be directly selected, with the highest affinity binder from the V2 cohort exhibiting a binding affinity of 97 pM. We validated the libraries against a broad range of antigens (Figure 4f), which showed a broad range of affinities not only for IFN $\alpha$ 2 but against 10 additional target campaigns, with over 50% of campaigns yielding subnanomolar binders and almost all producing at least one single digit nanomolar binder. Finally, testing the VHH for developability revealed that most fell within the parental range and/or above the established GMM threshold, which differentiates between acceptable and unacceptable ranges. Due to the limited available published developability data for isolated VHHs (with VHH-Fc format being more commonly assessed), we developed four criteria to establish baseline cutoffs for favorable or unfavorable developability properties. First, the lowest 10% threshold for the V1 design was based on its theoretical “less optimal” status relative to V2. Second, we used developability values from the parental mean  $\pm 2$  standard deviations, considering parental molecules are presumably well behaved as they’ve achieved clinical-grade status (Phase 2 to Approved). Third, a GMM was used for binary classification, setting thresholds at the boundary between favorable and unfavorable distributions. Finally, the Mann-Whitney U significance test compared developability metrics between V1 and V2 populations, providing non-parametric validation of our observations.

It is notable that the overall expression yields are relatively low for VHH, including for the parental therapeutic constructs (Figure 5a,b). This may be attributed to the high-throughput expression system used, which is not specifically optimized for yield or VHH expression. Therefore, these results should be interpreted on the basis of relative comparisons rather than absolute yields.

Given that the developability assessment derived from a single target (IFN $\alpha$ 2), it was crucial that the outputs exhibited

broad diversity across the four different scaffolds in both V1 and V2 designs. The outputs showed highly distinct CDR3 distributions (Figure 4d) with 84 unique sequences in V1 and 88 in V2, with a median Levenshtein distance of 10. This diversity was expanded when considering the high diversity in CDR1 and CDR2 (83 in V1 and V2 for CDR1, and 74 in V1 and 72 in V2 for CDR2).

While we showed that VHH derived from our initial design (V1) population exhibited a strong performance (60% of binders with  $\leq 2$  developability flags) in both affinity and developability, subsequent enhancements, including poly-tyrosine removal and heat shock tolerance, resulted in an even better performance in binding and developability (74% binders with  $\leq 2$  developability flags), validating these library modifications. Surprisingly, we found that sublibrary 3 (based on sonelokimab) exhibited reduced developability in the V2 design. Whether this was target related, or more generally applicable will be investigated in future studies. By combining the best sublibraries (1, 2 and 4 from V2 and 3 from V1), we find that 77% of antibodies show therapeutic lead potential in terms of affinity and developability properties, at least for VHH selected against IFN $\alpha$ 2.

In line with the philosophy guiding the construction of our libraries, we used the four therapeutic VHH scaffolds without modification, none of which contained the four hallmark “camelizing” residues. It was recently shown<sup>78</sup> that incorporation of some of these camelizing residues into isolated human VHs selected from an *in vitro* library, and subsequently affinity matured, was able to improve the developability properties of some of the VHs with a tendency to aggregate. Whether the post-selection introduction of these mutations into VHHs with developability issues, or the creation of new libraries based on the same scaffolds, but containing these mutations, will further improve VHH developability properties remains to be explored.

The nonhuman origin of VHHs might carry an elevated risk of eliciting undesired immune responses when used as therapeutics. However, this risk can be mitigated by the inherent sequence identity of VHHs to the human germline VH3 family,<sup>26</sup> particularly in humanized VHH therapeutics (Figure 1a). In our library design, immunogenicity is potentially further reduced by incorporating human-derived CDRs from VH3 germline genes directly into these humanized frameworks. Many of these VHH therapeutic scaffolds are currently undergoing clinical trials or have already attained approval status in some countries, such as caplacizumab, which provides a high degree of confidence in their robust safety assessment profiles. For instance, studies have demonstrated that caplacizumab (Lib1 scaffold) exhibits favorable immunogenicity profiles, with only 8.3% (3/36) of patients testing positive for anti-drug antibody (ADA) responses.<sup>79</sup> It is noteworthy that vobarilizumab does show an elevated incidence of ADAs (31%) in Phase 2b studies.<sup>79</sup> However, this did not impact pharmacokinetics, efficacy, or safety. These findings highlight the potential of humanized VHH therapeutics to maintain low immunogenicity while retaining therapeutic efficacy.

Over the years, numerous innovative VHH library designs have been proposed.<sup>36,39,48,80–89</sup> Our method focused on using

already humanized well-behaved therapeutic scaffolds, eliminating sequence liabilities during initial construction, and ensuring protein A binding. This approach is expected to optimize and speed up the downstream selection process, resulting in a greater number of lead candidates with favorable biophysical characteristics and a more straightforward manufacturing process. Consequently, the resulting library aims to balance diversity (CDR3) and functionality (CDR1–2 and scaffolds), to enable the development of high-affinity developable binders for therapeutic applications.

## Materials and methods

### Recombinant proteins

Protein A (Thermo Scientific, #21184), von Willebrand Factor (Sino Biological, # 10973-H08C), IL17A/IL17F (Sino Biological, # CT047-H08H), ADAMTS5 (R&D systems, # 2198-AD), IFN- $\alpha$ 2 (Genscript, # Z03002), IL6R (Sino Biological, # 10398-H02H).

### Single-CDR library construction and heat shock filtering

The libraries were built and filtered by MACS, as described before<sup>76</sup> but using magnetic beads functionalized with protein A (Miltenyi Biotec). For the heat shocks, 4 ml of the induced yeast cells were washed, then resuspended in 500  $\mu$ l of phosphate-buffered saline (PBS). They were incubated for 20 minutes at various temperatures (20°C, 60°C, 61.5°C, 64°C, 67.4°C, 71.8°C, 76°C, 78.6°C, and 80°C) in a thermocycler and subsequently chilled on ice for 2 minutes. Following this, cells were purified using protein A beads, and plasmid DNA was extracted using the Zymo yeast mini-prep II kit. Lastly, the region encoding the VHHs was amplified and transformed into EBY100 cells for further rounds.

### HCDR3 recovery and targeted degradation

HCDR3 was retrieved from CD19<sup>+</sup> IgM<sup>+</sup> B cells from 10 leukopaks from human donors, as previously described.<sup>76</sup> For targeted degradations, the Yblock primers (Supplementary Table S2), targeting the bolded portion of the JH6 gene shown in Figure 3a, were annealed to mRNA prior to cDNA synthesis and treated with RNase H to degrade mRNA containing these sequences (or parts of them). Reverse transcription was performed using SuperScript IV First-Strand Synthesis System (Invitrogen) with IgM RT primer (Supplementary Table S2), which hybridizes to the CH1 regions of human IgM, followed by PCR amplification using the F-L-HCDR3 and JH-NheI-universal primers (Supplementary Table S2).<sup>77</sup>

### NGS preparation & analysis of human repertoire sequences

Sequences were prepared<sup>41,75</sup> and processed as described previously.<sup>75</sup> Both MiSeq (single-CDR libraries) and NovaSeq 6000 (HCDR3 diversity) were performed at the genomics core facility at the University of Illinois.

### NGS analysis of alpaca repertoire sequences

Raw sequences were retrieved from Tu et al.<sup>56</sup> and annotated with our internal pipeline.<sup>75</sup>

### Vector preparation

The phagemid vector, pDAN5,<sup>59</sup> was purified from *E. coli* Omnimax 2 cells (Invitrogen, # C854003) by standard alkaline lysis followed by cesium chloride/ethidium bromide gradient centrifugation. The product was linearized using BssHII and NheI restriction enzymes (NEB, #R0199 and #R3131) and gel purified.

### Phage display sublibrary assembly and transformation

CDR1 and CDR2 from the protein A-filtered sublibraries were PCR-amplified along with adjacent framework areas and assembled with HCDR3 retrieved from the B cells.<sup>76</sup> The pieces were assembled in an overlap extension PCR using Q5 polymerase (NEB, #M0493) and inserted into the previously cut pDAN5 phagemid vector. The ligation mixture was then used to transform *E. coli* TG1 cells (Lucigen, catalog #60502–2) through electroporation. The transformed cells were spread on 2 $\times$ YT agar plates, which included 3% glucose, 1.5% sucrose, and 100 mg/ml carbenicillin, and incubated at 30°C overnight. The following day, the colonies were collected, and the bacterial cells from each sublibrary were preserved separately in 2 $\times$ YT containing 16% glycerol and stored at –80°C.

### Phage-display selection

For the phage selections, we utilized 10  $\mu$ L of streptavidin-conjugated magnetic beads (Dynabeads M-280, ThermoFisher, Cat # 11205D), coating the beads with an excess of biotinylated antigen to ensure complete coverage. The automated Kingfisher magnetic bead system (Thermo Fisher Scientific) was used to conduct the actual selection process, with a series of washing steps performed to remove non-binding phage from the beads. Remaining phage were recovered from the beads by acid elution and used to infect F' pilus-carrying bacteria (Omnimax-2, Thermo Fisher Scientific). After propagation of the eluted phage the selection cycle was repeated.

### Yeast display and sorting of VHH

After one or 2 rounds of phage selection, the enriched VHH antibodies were subcloned into the yeast display vector as previously described.<sup>55,76,90</sup> The selected VHH genes were amplified with specific primers that introduced an overlap with the yeast display vector pDNL6. The vector and the fragments were co-transformed into yeast cells to allow cloning by homologous recombination.<sup>55,76,90</sup> The yeast mini-libraries obtained were further enriched for binders using flow cytometry according to previously published protocols.<sup>55,76,90</sup> After induction, 2  $\times$  10<sup>6</sup> yeast cells were stained with 100, 20, and 5 nM of biotinylated antigen. Cells were labeled with streptavidin-AlexaFluor633 to detect

binding of biotinylated target antigens and anti-SV5-PE to assess VHH display levels. Yeast clones showing both antigen binding (AlexaFluor633 positives) and display (PE positives) were sorted. The collected cells were grown at 30°C for 2 days and induced for the next round of sorting at 20°C for 16 h.

### **Soluble expression as VHH and VHH-Fc**

Monomeric VHH and dimeric VHH-Fc (human IgG1 Fc domain, C220S mutation to eliminate unpaired cysteine) fusions were expressed by BioIntron using their high-throughput 4 ml/100 µg scale expression system.

### **Size-exclusion (SEC) HPLC for fraction monomer content and expression yield measurement**

Samples were evaluated for aggregation, degradation and main peak using the retention time of 20 µL of sample of the expressed and purified material (undiluted) into a TSKgel G3000SWxl column (60 cm x 7.5 mm, CAT# 0008541) at flow rate of 1 mL/min with 100 mM phosphate buffer (pH 6.8) using the Shimadzu LC-20AT. The fraction main peak was evaluated for the area under the curve (AUC) at 280 and 214 nm detection at retention time of known monomer peak and calculated as a fraction of total peak mass from aggregate and degradant peaks. Expression yield is calculated on the purified material according to concentration and volume, with concentration determined by the Nanodrop A280.

Affinity evaluation with the Biacore 8K. A Biacore 8K+ (Cytiva) SPR system equipped with CM5 sensor chips (Cytiva) was used to estimate the kinetic parameters of the VHH-Fc captured by an anti-Fc antibody. Briefly, goat Anti-Human IgG Fc (Southern Biotech #2014-01) was amine-coupled in CM5 Series S chips under standard conditions at a flow rate of 10 µL/min. Flow cells 1 and 2 were activated with a freshly prepared 1:1 v/v mixture of aqueous stocks of 75 mg/mL 1-ethyl-3-(3-dimethylaminopropyl) carbodiimide (EDC) + 11.5 mg/mL N-hydroxysuccinimide (NHS) for 7 min. The goat anti-human IgG Fc (same as above) was diluted to 12 µg/mL in 10 mM sodium acetate pH 4.5 and coupled for 7 min. The unreacted esters were blocked with 1 M ethanolamine-HCl at pH 8.5 for 7 min. The coupling level was around 2,000 RU (Response Units). For the affinity estimations, runs were performed with HBSP (10 mM HEPES, pH 7.4, 150 mM NaCl, 0.05% Tween-20) as running buffer at a standard 25°C flow cell temperature. The VHH-Fc antibodies were captured at a flow rate 5 µL/min for 90 seconds over flow cell 2. Capture level was approximately 100 RU. Recombinant purified human IFN $\alpha$  was injected at concentrations of 2 and 20 nM over flow cells 1 and 2 for 2 min (flow rate of 30 µL/min); dissociation time set to 5 min. Two buffer blank injections were run per interaction. The chip surface was regenerated with two injections of 10 mM glycine pH 1.5 (10 s) at a flow rate of 30 µL/min. The data was processed and analyzed with Biacore Insight Evaluation Software Version (Cytiva Life Sciences). Responses from flow cell 1 (reference) were subtracted from the responses from flow cell 2 (sample). Data was fit to a 1:1 Langmuir binding model to determine the apparent association rate constant ( $k_a$ ) and dissociation rate

constants ( $k_d$ ). Their ratio provided the apparent equilibrium dissociation constant or affinity constant ( $K_D = k_d/k_a$ ).

Standup monolayer adsorption chromatography (SMAC) for Hydrophobicity & Aggregation Potential. The SMAC assay was performed on an Agilent 1260 Infinity II instrument equipped with an auto sampler. 0.5 µg (10 µL at 50 µg/mL) of each VHH sample was injected into a Zenix SEC-300 column (Sepax Technologies). A flow rate of 0.50 mL/min with the running buffer containing 1X PBS at pH 7.4 was used. The running time for each sample was 30 minutes after column pre-equilibration. The collected data was analyzed using Agilent HPLC – data analysis software. Retention time for each sample was assigned based on the major elution peak, which is inversely related to their colloidal stability that antibodies prone to precipitation or aggregation retain longer on the column with even broader peaks.

### **Melting temperature & $T_{agg473}$ determination with the UNCLE**

Differential scanning fluorescence assay was performed using Uncle Instrument (Unchained Labs). 6 µL of stock antibody solution (0.5 mg/mL) was mixed with 4 µL of SYPRO Orange dye solution (The original dye in dimethylsulfoxide was 80-fold diluted in the Protein Thermal Shift Buffer, Applied Biosystems by Thermo Fisher Scientific). The final VHH concentration was 0.3 mg/mL. 9 µL of the mixture was loaded into the Unis (Unchained Labs). The Unis was heated from 25°C to 95°C at a ramp rate of 1°C/min. The data were collected by Uncle Client V6.0 software and analyzed by Uncle Analysis V6.0 software. The resulting fluorescence data (Integrated in Counts.nm) and static light scattering data (SLS473 in Counts.nm) plots against temperature were used to calculate melting temperature  $T_m$  and aggregation temperature  $T_{agg}$  at 473 nm at 50% transition based on the two-state model.

### **Polyreactivity Enzyme Linked Immunosorbent Assay (ELISA)**

Polyreactivity probes (50 µL) identified in previous studies<sup>70,71</sup> were immobilized overnight in Maxisorp ELISA plates (31050; ThermoFisher) by incubation at 4°C. Concentration of the probes for immobilization was: Poly-L-Glutamic acid sodium salt (26247-79-0; Alamanda polymers) at 20 µg/mL in PBS; Cardiolipin (50 µg/mL, C0563; Sigma) at 50 µg/mL in Absolute Ethanol, double-strand DNA (dsDNA) (1 µg/mL, D4522; Sigma) at 1 µg/mL in PBS. IFN $\alpha$ 2 (target, Genescript; Z03002) was immobilized in PBS (6 µg/mL) on the same time-temperature conditions as the probes. After adsorption, wells were washed three times with 300 µL PBS and blocked with 3% bovine serum albumin (BSA) at room temperature for 1 h 30 min. The testing VHH-Fc (50 µL) were added to the wells at a concentration of 8 µg/mL (~100 nM) in PBS and incubated for 2 h. Wells were washed three times with PBS and 50 µL of anti-human IgG (1:2000 in PBS 0.5% BSA, 109-035-008; Jackson ImmunoResearch) were added per well. After 1 hour incubation at room temperature, wells were washed three times with PBS and 50 µL of 3,3', 5,5' tetramethylbenzidine (TMB) substrate (T8665-1 L; Sigma) were added per well and incubated for 2–5 min until the reaction started saturating for



positive polyreactivity controls (C-10 VHH-Fc, Gantenerumab IgG, Bococizumab IgG). To stop the reactions, 25  $\mu$ L of 1 M sulfuric acid (H<sub>2</sub>SO<sub>4</sub>) were added per well. Absorbance was read at 450 nm. For analysis, data was blanked against the anti-Fc secondary-only values. All previously listed incubation/washing PBS steps were performed in PBS pH 7.4.

### Baculovirus particle (BVP) binding assay

Immobilization of BVP particles (LakePharma; #25690) was performed by overnight incubation (4°C) at 30 ng/ $\mu$ L (in 50  $\mu$ L sodium carbonate buffer 50 mM pH 9.6) in Maxisorp ELISA plates (31050; ThermoFisher). Next day, unbound particles were removed by six washing steps with 300  $\mu$ L PBS.<sup>70</sup> PBS BSA 0.5% (200  $\mu$ L) was added for blocking and incubated for 1 h 30 min. Wells were washed three times with 300  $\mu$ L PBS and the testing VHH-Fcs added at 8  $\mu$ g/mL (~100 nM) and incubated for 2 h. Wells were washed six times with 300  $\mu$ L PBS and 50  $\mu$ L of anti-Fc (1:2000 in PBS 0.5% BSA, 109–035–008; Jackson ImmunoResearch) secondary antibody added per well. After 1 h incubation, wells were washed three times with PBS and 50  $\mu$ L of TMB added. Reaction was developed for 2–5 min and stopped when polyreactivity controls (C-10 VHH-Fc, Gantenerumab IgG, Bococizumab IgG) saturated by using 25  $\mu$ L of 1 M H<sub>2</sub>SO<sub>4</sub> per well. Absorbance was read at 450 nm. For analysis, data was blanked against the anti-Fc secondary-only values. All listed incubation/washing PBS steps were performed in PBS pH 7.4.

### Bioinformatics processing of developability data

We established four separate thresholds to compare the V1 and V2 populations including 1) use of non-parametric Mann-Whitney U statistical test (also known as the Wilcoxon rank-sum test) using the `wilcox.test` R function, 2) parental (therapeutic) mean  $\pm$  2 standard deviations, 3) Gaussian Mixture Model based on combined V1 and V2 from the `normalmixEM`, part of the `mixtools` R package and 4) worst 10% threshold based on V1 population rank ordered in direction of worst to most favorable.

### Disclosure statement


MFE, EM, LARC, DK, HT, JL, RDN, AF, AD, LS, SD, FF, ARMB are or were employees of Specifica, an IQVIA business at the time of writing. CLL is currently an employee of Sanofi, Cambridge, USA. All other authors declare no conflicts of interest. AART and ARMB are listed as co-inventors on patent describing this technology.

### Funding

The author(s) reported there is no funding associated with the work featured in this article.

### ORCID

M. Frank Erasmus  <http://orcid.org/0000-0002-8046-5048>  
 Andre A. R. Teixeira  <http://orcid.org/0000-0002-8348-0235>  
 Esteban Molina  <http://orcid.org/0000-0003-2813-5997>  
 Luis Antonio Rodriguez Carnero  <http://orcid.org/0000-0003-0684-4039>

Jianquan Li  <http://orcid.org/0009-0004-3205-9514>  
 Camila Leal-Lopes  <http://orcid.org/0000-0001-7058-3350>  
 Adeline Fanni  <http://orcid.org/0000-0002-7240-2792>  
 Ashley DeAguiro  <http://orcid.org/0000-0001-7110-9167>  
 Laura Spector  <http://orcid.org/0000-0002-3965-9115>  
 Sara D'Angelo  <http://orcid.org/0000-0002-8329-9938>  
 Fortunato Ferrara  <http://orcid.org/0000-0002-4615-035X>  
 Andrew R. M. Bradbury  <http://orcid.org/0000-0002-5567-8172>

### References

- Morrison C. Nanobody approval gives domain antibodies a boost. *Nat Rev Drug Discov.* 2019;18(7):485–487. doi: [10.1038/d41573-019-00104-w](https://doi.org/10.1038/d41573-019-00104-w).
- Revetts H, De Baetselier P, Muyldermans S. Nanobodies as novel agents for cancer therapy. *Expert Opin Biol Ther.* 2005;5(1):111–124. doi: [10.1517/14712598.5.1.111](https://doi.org/10.1517/14712598.5.1.111).
- Bannas P, Hambach J, Koch-Nolte F. Nanobodies and nanobody-based human heavy chain antibodies as antitumor therapeutics. *Front Immunol.* 2017;8:1603. doi: [10.3389/fimmu.2017.01603](https://doi.org/10.3389/fimmu.2017.01603).
- Hu Y, Liu C, Muyldermans S. Nanobody-based delivery systems for diagnosis and targeted tumor therapy. *Front Immunol.* 2017;8:1442. doi: [10.3389/fimmu.2017.01442](https://doi.org/10.3389/fimmu.2017.01442).
- Hamers-Casterman C, Atarhouch T, Muyldermans S, Robinson G, Hammers C, Songa EB, Bendahman N, Hammers R. Naturally occurring antibodies devoid of light chains. *Nature.* 1993;363(6428):446–448. doi: [10.1038/363446a0](https://doi.org/10.1038/363446a0).
- Stanfield RL, Dooley H, Verdino P, Flajnik MF, Wilson IA. Maturation of shark single-domain (IgNAR) antibodies: evidence for induced-fit binding. *J Mol Biol.* 2007;367(2):358–372. doi: [10.1016/j.jmb.2006.12.045](https://doi.org/10.1016/j.jmb.2006.12.045).
- Desmyter A, Decanniere K, Muyldermans S, Wyns L. Antigen specificity and high affinity binding provided by one single loop of a camel single-domain antibody. *J Biol Chem.* 2001;276(28):26285–26290. doi: [10.1074/jbc.M102107200](https://doi.org/10.1074/jbc.M102107200).
- Spinelli S, Tegoni M, Frenken L, van Vliet C, Cambillau C. Lateral recognition of a dye hapten by a llama VHH domain. *J Mol Biol.* 2001;311(1):123–129. doi: [10.1006/jmbi.2001.4856](https://doi.org/10.1006/jmbi.2001.4856).
- Conway JO, Sherwood LJ, Collazo MT, Garza JA, Hayhurst A, Hofmann A. Llama single domain antibodies specific for the 7 botulinum neurotoxin serotypes as heptaplex immunoreagents. *PLOS ONE.* 2010;5(1):e8818. doi: [10.1371/journal.pone.0008818](https://doi.org/10.1371/journal.pone.0008818).
- Ishiwatari-Ogata C, Kyuuma M, Ogata H, Yamakawa M, Iwata K, Ochi M, Hori M, Miyata N, Fujii Y. Ozoralizumab, a humanized anti-TNF $\alpha$  NANOBODY<sup>®</sup> compound, exhibits efficacy not only at the onset of arthritis in a human TNF transgenic mouse but also during secondary failure of administration of an anti-TNF $\alpha$  IgG. *Front Immunol.* 2022;13:853008. doi: [10.3389/fimmu.2022.853008](https://doi.org/10.3389/fimmu.2022.853008).
- Rajabzadeh A, Hamidieh AA, Rahbarizadeh F. Spinoculation and retronectin highly enhance the gene transduction efficiency of mucin-1-specific chimeric antigen receptor (CAR) in human primary T cells. *Cell J.* 2021;22(1):502–513. doi: [10.1186/s12860-021-00397-z](https://doi.org/10.1186/s12860-021-00397-z).
- Heukers R, Mashayekhi V, Ramirez-Escudero M, de Haard H, Verrips TC, van Bergen En Henegouwen PMP, Oliveira S. VHH-Photosensitizer conjugates for targeted photodynamic therapy of met-overexpressing tumor cells. *Antibodies (basel).* 2019;8(2):26. doi: [10.3390/antib8020026](https://doi.org/10.3390/antib8020026).
- Flicker S, Zettl I, Tillib SV. Nanobodies—useful tools for allergy treatment? *Front Immunol.* 2020;11:576255. doi: [10.3389/fimmu.2020.576255](https://doi.org/10.3389/fimmu.2020.576255).
- Nosenko MA, Atretkhany KSN, Mokhonov VV, Efimov GA, Kruglov AA, Tillib SV, Drutskaya MS, Nedospasov SA. VHH-Based bispecific antibodies targeting cytokine production. *Front Immunol.* 2017;8:1073. doi: [10.3389/fimmu.2017.01073](https://doi.org/10.3389/fimmu.2017.01073).
- Dörner T, Beneden M, Beneden KV, Dombrecht EJ, Beuf KD, Schoen RK, Zeldin RK. FRI0239 Results of a phase 2b study of vobarilizumab, an anti-interleukin-6 receptor nanobody, as



- monotherapy in patients with moderate to severe rheumatoid arthritis. *Ann Rheumatic Dis.* 2017;76:575–575. doi: [10.1136/annrheumdis-2017-eular.3746](https://doi.org/10.1136/annrheumdis-2017-eular.3746).
16. Dong J, Huang B, Wang B, Titong A, Gallolu Kankanamalage S, Jia Z, Wright M, Parthasarathy P, Liu Y. Development of humanized tri-specific nanobodies with potent neutralization for SARS-CoV-2. *Sci Rep.* 2020;10(1):17806. doi: [10.1038/s41598-020-74761-y](https://doi.org/10.1038/s41598-020-74761-y).
  17. Guttler T, Aksu M, Dickmanns A, Stegmann KM, Gregor K, Rees R, Taxer W, Rymarenko O, Schünemann J, Dienemann C, et al. Neutralization of SARS-CoV-2 by highly potent, hyperthermostable, and mutation-tolerant nanobodies. *EMBO J.* 2021;40(19):e107985. doi: [10.15252/embj.2021107985](https://doi.org/10.15252/embj.2021107985).
  18. Xu J, Xu K, Jung S, Conte A, Lieberman J, Muecksch F, Lorenzi JCC, Park S, Schmidt F, Wang Z, et al. Nanobodies from camelid mice and llamas neutralize SARS-CoV-2 variants. *Nature.* 2021;595(7866):278–282. doi: [10.1038/s41586-021-03676-z](https://doi.org/10.1038/s41586-021-03676-z).
  19. Biswas M, Yamazaki T, Chiba J, Akashi-Takamura S. Broadly neutralizing antibodies for influenza: passive immunotherapy and intranasal vaccination. *Vaccines (Basel).* 2020;8(3):424. doi: [10.3390/vaccines8030424](https://doi.org/10.3390/vaccines8030424).
  20. Hussack G, Ryan S, van Faassen H, Rossotti M, MacKenzie CR, Tanha J. Neutralization of clostridium difficile toxin B with VHH-Fc fusions targeting the delivery and CROPs domains. *PLOS ONE.* 2018;13(12):e0208978. doi: [10.1371/journal.pone.0208978](https://doi.org/10.1371/journal.pone.0208978).
  21. Lam KH, Perry K, Shoemaker CB, Jin R. Two VHH antibodies neutralize botulinum neurotoxin E1 by blocking its membrane translocation in host cells. *Toxins (basel).* 2020;12(10):616. doi: [10.3390/toxins12100616](https://doi.org/10.3390/toxins12100616).
  22. Maffey L, Vega CG, Mino S, Garaicoechea L, Parreno V, Ho PL. Anti-VP6 VHH: an experimental treatment for rotavirus A-Associated disease. *PLOS ONE.* 2016;11(9):e0162351. doi: [10.1371/journal.pone.0162351](https://doi.org/10.1371/journal.pone.0162351).
  23. Strokappe NM, Hock M, Rutten L, McCoy LE, Back JW, Caillat C, Haffke M, Weiss RA, Weissenhorn W, Verrips T, et al. Super potent bispecific llama VHH antibodies neutralize HIV via a combination of gp41 and gp120 epitopes. *Antibodies (basel).* 2019;8(2):38. doi: [10.3390/antib8020038](https://doi.org/10.3390/antib8020038).
  24. Scully M, Cataland SR, Peyvandi F, Coppo P, Knöbl P, Kremer Hovinga JA, Metjian A, de la Rubia J, Pavenski K, Callewaert F. Caplacizumab treatment for acquired thrombotic thrombocytopenic purpura. *N Engl J Med.* 2019;380(4):335–346. doi: [10.1056/NEJMoa1806311](https://doi.org/10.1056/NEJMoa1806311).
  25. Muyldermans S. Nanobodies: natural single-domain antibodies. *Annu Rev Biochem.* 2013;82(1):775–797. doi: [10.1146/annurev-biochem-063011-092449](https://doi.org/10.1146/annurev-biochem-063011-092449).
  26. Steeland S, Vandenbroucke RE, Libert C. Nanobodies as therapeutics: big opportunities for small antibodies. *Drug Discov Today.* 2016;21(7):1076–1113. doi: [10.1016/j.drudis.2016.04.003](https://doi.org/10.1016/j.drudis.2016.04.003).
  27. Jovcevska I, Muyldermans S. The therapeutic potential of nanobodies. *BioDrugs.* 2020;34(1):11–26. doi: [10.1007/s40259-019-00392-z](https://doi.org/10.1007/s40259-019-00392-z).
  28. De Meyer T, Muyldermans S, Depicker A. Nanobody-based products as research and diagnostic tools. *Trends Biotechnol.* 2014;32(5):263–270. doi: [10.1016/j.tibtech.2014.03.001](https://doi.org/10.1016/j.tibtech.2014.03.001).
  29. Muyldermans S. A guide to: generation and design of nanobodies. *FEBS J.* 2021;288(7):2084–2102. doi: [10.1111/febs.15515](https://doi.org/10.1111/febs.15515).
  30. Kinoshita S, Nakakido M, Mori C, Kuroda D, Caaveiro JMM, Tsumoto K. Molecular basis for thermal stability and affinity in a VHH: contribution of the framework region and its influence in the conformation of the CDR3. *Protein Sci: Publ Protein Soc.* 2022;31(11):e4450. doi: [10.1002/pro.4450](https://doi.org/10.1002/pro.4450).
  31. Fernandez-Quintero ML, Guarnera E, Musil D, Pekar L, Sellmann C, Freire F, Sousa RL, Santos SP, Freitas MC, Bandejas TM, et al. On the humanization of VHHs: prospective case studies, experimental and computational characterization of structural determinants for functionality. *Protein Sci: Publ Protein Soc.* 2024;33(11):e5176. doi: [10.1002/pro.5176](https://doi.org/10.1002/pro.5176).
  32. Li Q, Zhang F, Lu Y, Hu H, Wang J, Guo C, Deng Q, Liao C, Wu Q, Hu T, et al. Highly potent multivalent VHH antibodies against chikungunya isolated from an alpaca naïve phage display library. *J Nanobiotechnol.* 2022;20(1):231. doi: [10.1186/s12951-022-01417-6](https://doi.org/10.1186/s12951-022-01417-6).
  33. Monegal A, Ami D, Martinelli C, Huang H, Aliprandi M, Capasso P, Francavilla C, Ossolengo G, de Marco A. Immunological applications of single-domain llama recombinant antibodies isolated from a naïve library. *Protein Eng Des Sel.* 2009;22(4):273–280. doi: [10.1093/protein/gzp002](https://doi.org/10.1093/protein/gzp002).
  34. Olichon A, de Marco A. Preparation of a naïve library of camelid single domain antibodies. *Methods Mol Biol.* 2012;911:65–78. doi: [10.1007/978-1-61779-968-6\\_5](https://doi.org/10.1007/978-1-61779-968-6_5).
  35. Wang M, Wei L, Xiang H, Ren B, Liu X, Jiang L, Yang N, Shi J. A megadiverse naïve library derived from numerous camelids for efficient and rapid development of VHH antibodies. *Analytical Biochem.* 2022;657:114871. doi: [10.1016/j.ab.2022.114871](https://doi.org/10.1016/j.ab.2022.114871).
  36. McMahon C, Baier AS, Pascolutti R, Wegrecki M, Zheng S, Ong JX, Erlandson SC, Hilger D, Rasmussen SGF, Ring AM, et al. Yeast surface display platform for rapid discovery of conformationally selective nanobodies. *Nat Struct Mol Biol.* 2018;25(3):289–296. doi: [10.1038/s41594-018-0028-6](https://doi.org/10.1038/s41594-018-0028-6).
  37. Shin JE, Riesselman AJ, Kollasch AW, McMahon C, Simon E, Sander C, Manglik A, Kruse AC, Marks DS. Protein design and variant prediction using autoregressive generative models. *Nat Commun.* 2021;12(1):2403. doi: [10.1038/s41467-021-22732-w](https://doi.org/10.1038/s41467-021-22732-w).
  38. Contreras MA, Serrano-Rivero Y, González-Pose A, Salazar-Urbe J, Rubio-Carrasquilla M, Soares-Alves M, Parra NC, Camacho-Casanova F, Sánchez-Ramos O, Moreno E, et al. Design and construction of a synthetic nanobody library: testing its potential with a single selection round strategy. *Oxycedrus Needles Berries Mol.* 2023;28(9):3708. doi: [10.3390/molecules28093708](https://doi.org/10.3390/molecules28093708).
  39. Sevy AM, Chen M-T, Castor M, Sylvia T, Krishnamurthy H, Ishchenko A, Hsieh C-M. Structure- and sequence-based design of synthetic single-domain antibody libraries. *Protein Eng Des Sel.* 2020;33. doi: [10.1093/protein/gzaa028](https://doi.org/10.1093/protein/gzaa028).
  40. Yuan TZ, Garg P, Wang L, Willis JR, Kwan E, Hernandez AGL, Tuscano E, Sever EN, Keane E, Soto C, et al. Rapid discovery of diverse neutralizing SARS-CoV-2 antibodies from large-scale synthetic phage libraries. *MAbs-Austin.* 2022;14(1):2002236. doi: [10.1080/19420862.2021.2002236](https://doi.org/10.1080/19420862.2021.2002236).
  41. Nakakido M, Kinoshita S, Tsumoto K. Development of novel humanized VHH synthetic libraries based on physicochemical analyses. *Sci Rep.* 2024;14(1):19533. doi: [10.1038/s41598-024-70513-4](https://doi.org/10.1038/s41598-024-70513-4).
  42. Davies J, Riechmann L. ‘Camelising’ human antibody fragments: NMR studies on VH domains. *FEBS Lett.* 1994;339(3):285–290. doi: [10.1016/0014-5793\(94\)80432-X](https://doi.org/10.1016/0014-5793(94)80432-X).
  43. Barthelemy PA, Raab H, Appleton BA, Bond CJ, Wu P, Wiesmann C, Sidhu SS. Comprehensive analysis of the factors contributing to the stability and solubility of autonomous human VH domains. *J Biol Chem.* 2008;283(6):3639–3654. doi: [10.1074/jbc.M708536200](https://doi.org/10.1074/jbc.M708536200).
  44. Jespers L, Schon O, Famm K, Winter G. Aggregation-resistant domain antibodies selected on phage by heat denaturation. *Nat Biotechnol.* 2004;22(9):1161–1165. doi: [10.1038/nbt1000](https://doi.org/10.1038/nbt1000).
  45. Jespers L, Schon O, James LC, Vepritsiev D, Winter G. Crystal structure of HEL4, a soluble, refoldable human VH single domain with a germ-line scaffold. *J Mol Biol.* 2004;337(4):893–903. doi: [10.1016/j.jmb.2004.02.013](https://doi.org/10.1016/j.jmb.2004.02.013).
  46. Rossotti MA, Belanger K, Henry KA, Tanha J. Immunogenicity and humanization of single-domain antibodies. *FEBS J.* 2022;289(14):4304–4327. doi: [10.1111/febs.15809](https://doi.org/10.1111/febs.15809).
  47. Vincke C, Loris R, Saerens D, Martinez-Rodriguez S, Muyldermans S, Conrath K. General strategy to humanize a camelid single-domain antibody and identification of a universal humanized nanobody scaffold. *J Biol Chem.* 2009;284(5):3273–3284. doi: [10.1074/jbc.M806889200](https://doi.org/10.1074/jbc.M806889200).
  48. Tanha J, Xu P, Chen Z, Ni F, Kaplan H, Narang SA, MacKenzie CR. Optimal design features of camelized human

- single-domain antibody libraries. *J Biol Chem.* 2001;276(27):24774–24780. doi: [10.1074/jbc.M100770200](https://doi.org/10.1074/jbc.M100770200).
49. Ewert S, Cambillau C, Conrath K, Pluckthun A. Biophysical properties of camelid V HH domains compared to those of human V H 3 domains. *Biochemistry.* 2002;41(11):3628–3636. doi: [10.1021/bi011239a](https://doi.org/10.1021/bi011239a).
  50. Graille M, Stura EA, Corper AL, Sutton BJ, Taussig MJ, Charbonnier J-B, Silverman GJ. Crystal structure of a staphylococcus aureus protein a domain complexed with the Fab fragment of a human IgM antibody: structural basis for recognition of B-cell receptors and superantigen activity. *Proc Natl Acad Sci USA.* 2000;97(10):5399–5404. doi: [10.1073/pnas.97.10.5399](https://doi.org/10.1073/pnas.97.10.5399).
  51. Henry KA, Sulea T, van Faassen H, Hussack G, Purisima EO, MacKenzie CR, Arbabi-Ghahroudi M. A rational engineering strategy for designing protein A-Binding camelid single-domain antibodies. *PLOS ONE.* 2016;11(9):e0163113. doi: [10.1371/journal.pone.0163113](https://doi.org/10.1371/journal.pone.0163113).
  52. Teixeira AAR, Erasmus MF, D'Angelo S, Naranjo L, Ferrara F, Leal-Lopes C, Durrant O, Galmiche C, Morelli A, Scott-Tucker A, et al. Drug-like antibodies with high affinity, diversity and developability directly from next-generation antibody libraries. *Mabs-austin.* 2021;13(1):1980942. doi: [10.1080/19420862.2021.1980942](https://doi.org/10.1080/19420862.2021.1980942).
  53. Ferrara F, Fanni A, Teixeira AAR, Molina E, Leal-Lopes C, DeAgüero A, D'Angelo S, Erasmus MF, Spector L, Rodriguez Carnero LA, et al. A next-generation Fab library platform directly yielding drug-like antibodies with high affinity, diversity, and developability. *Mabs-Austin.* 2024;16(1):2394230. doi: [10.1080/19420862.2024.2394230](https://doi.org/10.1080/19420862.2024.2394230).
  54. Marks JD, Hoogenboom HR, Bonnert TP, McCafferty J, Griffiths AD, Winter G. By-passing immunization. Human antibodies from V-gene libraries displayed on phage. *J Mol Biol.* 1991;222(3):581–597. doi: [10.1016/0022-2836\(91\)90498-U](https://doi.org/10.1016/0022-2836(91)90498-U).
  55. Ferrara F, Naranjo LA, Kumar S, Gaiotto T, Mukundan H, Swanson B, Bradbury ARM. Using phage and yeast display to select hundreds of monoclonal antibodies: application to antigen 85, a tuberculosis biomarker. *PLOS ONE.* 2012;7(11):e49535. doi: [10.1371/journal.pone.0049535](https://doi.org/10.1371/journal.pone.0049535).
  56. Tu Z, Huang X, Fu J, Hu N, Zheng W, Li Y, Zhang Y. Landscape of variable domain of heavy-chain-only antibody repertoire from alpaca. *Immunology.* 2020;161(1):53–65. doi: [10.1111/imm.13224](https://doi.org/10.1111/imm.13224).
  57. Govaert J, Pellis M, Deschacht N, Vincke C, Conrath K, Muyldermans S, Saerens D. Dual beneficial effect of interloop disulfide bond for single domain antibody fragments. *J Biol Chem.* 2012;287(3):1970–1979. doi: [10.1074/jbc.M111.242818](https://doi.org/10.1074/jbc.M111.242818).
  58. Mendoza MN, Jian M, King MT, Brooks CL. Role of a noncanonical disulfide bond in the stability, affinity, and flexibility of a VHH specific for the listeria virulence factor InlB. *Protein Sci: Publ Protein Soc.* 2020;29(4):990–1003. doi: [10.1002/pro.3831](https://doi.org/10.1002/pro.3831).
  59. Sblattero D, Bradbury A. Exploiting recombination in single bacteria to make large phage antibody libraries. *Nat Biotechnol.* 2000;18(1):75–80. doi: [10.1038/71958](https://doi.org/10.1038/71958).
  60. Olsen TH, Boyles F, Deane CM. Observed antibody space: a diverse database of cleaned, annotated, and translated unpaired and paired antibody sequences. *Protein Sci: Publ Protein Soc.* 2022;31(1):141–146. doi: [10.1002/pro.4205](https://doi.org/10.1002/pro.4205).
  61. Tiller KE, Li L, Kumar S, Julian MC, Garde S, Tessier PM. Arginine mutations in antibody complementarity-determining regions display context-dependent affinity/specificity trade-offs. *J Biol Chem.* 2017;292(40):16638–16652. doi: [10.1074/jbc.M117.783837](https://doi.org/10.1074/jbc.M117.783837).
  62. Birtalan S, Zhang Y, Fellouse FA, Shao L, Schaefer G, Sidhu SS. The intrinsic contributions of tyrosine, serine, glycine and arginine to the affinity and specificity of antibodies. *J Mol Biol.* 2008;377(5):1518–1528. doi: [10.1016/j.jmb.2008.01.093](https://doi.org/10.1016/j.jmb.2008.01.093).
  63. Parker JM, Guo D, Hodges RS. New hydrophilicity scale derived from high-performance liquid chromatography peptide retention data: correlation of predicted surface residues with antigenicity and X-ray-derived accessible sites. *Biochemistry.* 1986;25(19):5425–5432. doi: [10.1021/bi00367a013](https://doi.org/10.1021/bi00367a013).
  64. Jain T, Boland T, Vasquez M. Identifying developability risks for clinical progression of antibodies using high-throughput in vitro and in silico approaches. *Mabs-Austin.* 2023;15(1):2200540. doi: [10.1080/19420862.2023.2200540](https://doi.org/10.1080/19420862.2023.2200540).
  65. Estep P, Caffry I, Yu Y, Sun T, Cao Y, Lynaugh H, Jain T, Vásquez M, Tessier PM, Xu Y, et al. An alternative assay to hydrophobic interaction chromatography for high-throughput characterization of monoclonal antibodies. *Mabs-Austin.* 2015;7(3):553–561. doi: [10.1080/19420862.2015.1016694](https://doi.org/10.1080/19420862.2015.1016694).
  66. Akbar R, Robert PA, Pavlović M, Jeliazkov JR, Snapkov I, Slabodkin A, Weber CR, Scheffer L, Miho E, Haff IH, et al. A compact vocabulary of paratope-epitope interactions enables predictability of antibody-antigen binding. *Cell Rep.* 2021;34(11):108856. doi: [10.1016/j.celrep.2021.108856](https://doi.org/10.1016/j.celrep.2021.108856).
  67. D'Angelo S, Ferrara F, Naranjo L, Erasmus MF, Hrabec P, Bradbury ARM. Many routes to an antibody heavy-chain CDR3: necessary, yet insufficient, for specific binding. *Front Immunol.* 2018;9:395. doi: [10.3389/fimmu.2018.00395](https://doi.org/10.3389/fimmu.2018.00395).
  68. Harvey EP, Shin J-E, Skiba MA, Nemeth GR, Hurley JD, Wellner A, Shaw AY, Miranda VG, Min JK, Liu CC, et al. An in silico method to assess antibody fragment polyreactivity. *Nat Commun.* 2022;13(1):7554. doi: [10.1038/s41467-022-35276-4](https://doi.org/10.1038/s41467-022-35276-4).
  69. Chen HT, Zhang Y, Huang J, Sawant M, Smith MD, Rajagopal N, Desai AA, Makowski E, Licari G, Xie Y, et al. Human antibody polyreactivity is governed primarily by the heavy-chain complementarity-determining regions. *Cell Rep.* 2024;43(10):114801. doi: [10.1016/j.celrep.2024.114801](https://doi.org/10.1016/j.celrep.2024.114801).
  70. Jain T, Sun T, Durand S, Hall A, Houston NR, Nett JH, Sharkey B, Bobrowicz B, Caffry I, Yu Y. Biophysical properties of the clinical-stage antibody landscape. *Proc Natl Acad Sci USA.* 2017;114(5):944–949. doi: [10.1073/pnas.1616408114](https://doi.org/10.1073/pnas.1616408114).
  71. Carnero LAR, Bedinger D, Cocklin S, Li J, Erasmus MF, D'Angelo S, Leal-Lopes C, Teixeira AAR, Ferrara F, Bradbury ARM, et al. Identification of polyreactive antibodies by high throughput enzyme-linked immunosorbent assay and surface plasmon resonance. *J Immunol Methods.* 2025;539:113855. doi: [10.1016/j.jim.2025.113855](https://doi.org/10.1016/j.jim.2025.113855).
  72. Yu X, Orr CM, Chan HTC, James S, Penfold CA, Kim J, Inzhelevskaya T, Mockridge CI, Cox KL, Essex JW, et al. Reducing affinity as a strategy to boost immunomodulatory antibody agonism. *Nature.* 2023;614(7948):539–547. doi: [10.1038/s41586-022-05673-2](https://doi.org/10.1038/s41586-022-05673-2).
  73. Erasmus MF, Dovner M, Ferrara F, D'Angelo S, Teixeira AA, Leal-Lopes C, Spector L, Hopkins E, Bradbury ARM. Determining the affinities of high-affinity antibodies using KinExA and surface plasmon resonance. *Mabs-Austin.* 2023;15(1):2291209. doi: [10.1080/19420862.2023.2291209](https://doi.org/10.1080/19420862.2023.2291209).
  74. Ferrara F, Erasmus MF, D'Angelo S, Leal-Lopes C, Teixeira AA, Choudhary A, Honnen W, Calianese D, Huang D, Peng L, et al. A pandemic-enabled comparison of discovery platforms demonstrates a naive antibody library can match the best immune-sourced antibodies. *Nat Commun.* 2022;13(1):462. doi: [10.1038/s41467-021-27799-z](https://doi.org/10.1038/s41467-021-27799-z).
  75. Erasmus MF, Ferrara F, D'Angelo S, Spector L, Leal-Lopes C, Teixeira AA, Sørensen J, Nagpal S, Perea-Schmittle K, Choudhary A, et al. Insights into next generation sequencing guided antibody selection strategies. *Sci Rep.* 2023;13(1):18370. doi: [10.1038/s41598-023-45538-w](https://doi.org/10.1038/s41598-023-45538-w).
  76. Azevedo Reis Teixeira A, Erasmus MF, D'Angelo S, Naranjo L, Ferrara F, Leal-Lopes C, Durrant O, Galmiche C, Morelli A, Scott-Tucker A, et al. Drug-like antibodies with high affinity, diversity and developability directly from next-generation antibody libraries. *Mabs-Austin.* 2021;13(1):1980942. doi: [10.1080/19420862.2021.1980942](https://doi.org/10.1080/19420862.2021.1980942).
  77. Arras P, Yoo HB, Pekar L, Schröter C, Clarke T, Krah S, Klewinghaus D, Siegmund V, Evers A, Zielonka S, et al. A library approach for the de novo high-throughput isolation of humanized VHH domains with favorable developability properties following camelid immunization. *Mabs-Austin.* 2023;15(1):2261149. doi: [10.1080/19420862.2023.2261149](https://doi.org/10.1080/19420862.2023.2261149).

78. Belanger K, Wu C, Sulea T, van Faassen H, Callaghan D, Aubry A, Sasseville M, Hussack G, Tanha J. Optimization of synthetic human V H affinity and solubility through in vitro affinity maturation and minimal camelization. *Protein Sci: Publ Protein Soc.* 2025;34(5):e70114. doi: [10.1002/pro.70114](https://doi.org/10.1002/pro.70114).
79. Ackaert C, Smiejewska N, Xavier C, Sterckx YGJ, Denies S, Stijlemans B, Elkrim Y, Devoogdt N, Caveliers V, Lahoutte T, et al. Immunogenicity risk profile of nanobodies. *Front Immunol.* 2021;12:632687. doi: [10.3389/fimmu.2021.632687](https://doi.org/10.3389/fimmu.2021.632687).
80. Ju M-S, Lee SM, Kwon HS, Lee JC, Park JC, Jung ST, Jung ST. A synthetic library for rapid isolation of humanized single-domain antibodies. *Biotechnol Bioproc E.* 2017;22(3):239–247. doi: [10.1007/s12257-017-0082-7](https://doi.org/10.1007/s12257-017-0082-7).
81. Misson Mindrebo L, Liu H, Ozorowski G, Tran Q, Woehl J, Khalek I, Smith J, Barman S, Zhao F, Keating C, et al. Fully synthetic platform to rapidly generate tetravalent bispecific nanobody-based immunoglobulins. *Proc Natl Acad Sci USA.* 2023;120(24):e2216612120. doi: [10.1073/pnas.2216612120](https://doi.org/10.1073/pnas.2216612120).
82. Murakami T, Kumachi S, Matsunaga Y, Sato M, Wakabayashi-Nakao K, Masaki H, Yonehara R, Motohashi M, Nemoto N, Tsuchiya M, et al. Construction of a humanized artificial VHH library reproducing structural features of camelid VHHs for therapeutics. *Antibodies (basel).* 2022;11(1):10. doi: [10.3390/antib11010010](https://doi.org/10.3390/antib11010010).
83. Rouet R, Dudgeon K, Christie M, Langley D, Christ D. Fully human VH single domains that rival the stability and cleft recognition of camelid antibodies. *J Biol Chem.* 2015;290(19):11905–11917. doi: [10.1074/jbc.M114.614842](https://doi.org/10.1074/jbc.M114.614842).
84. Sabir JS, Atef A, El-Domyati FM, Edris S, Hajrah N, Alzohairy AM, Bahieldin A. Construction of naïve camelids VHH repertoire in phage display-based library. *C R Biol.* 2014;337(4):244–249. doi: [10.1016/j.crvi.2014.02.004](https://doi.org/10.1016/j.crvi.2014.02.004).
85. Sun Z, Li W, Mellors JW, Orentas R, Dimitrov DS. Construction of a large size human immunoglobulin heavy chain variable (VH) domain library, isolation and characterization of novel human antibody VH domains targeting PD-L1 and CD22. *Front Immunol.* 2022;13:869825. doi: [10.3389/fimmu.2022.869825](https://doi.org/10.3389/fimmu.2022.869825).
86. Tanha J, Dubuc G, Hiram T, Narang SA, MacKenzie CR. Selection by phage display of llama conventional VH fragments with heavy chain antibody VHH properties. *J Immunological Methods.* 2002;263(1–2):97–109. doi: [10.1016/S0022-1759\(02\)00027-3](https://doi.org/10.1016/S0022-1759(02)00027-3).
87. Yan J, Li G, Hu Y, Ou W, Wan Y. Construction of a synthetic phage-displayed nanobody library with CDR3 regions randomized by trinucleotide cassettes for diagnostic applications. *J Transl Med.* 2014;12(1):343. doi: [10.1186/s12967-014-0343-6](https://doi.org/10.1186/s12967-014-0343-6).
88. Moutel S, Bery N, Bernard V, Keller L, Lemesre E, de Marco A, Ligat L, Rain J-C, Favre G, Olichon A, et al. NaLi-H1: a universal synthetic library of humanized nanobodies providing highly functional antibodies and intrabodies. *Elife.* 2016;5. doi: [10.7554/eLife.16228](https://doi.org/10.7554/eLife.16228).
89. Hairul Bahara NH, Chin ST, Choong YS, Lim TS. Construction of a semisynthetic human VH single-domain antibody library and selection of domain antibodies against  $\alpha$ -crystalline of mycobacterium tuberculosis. *J Biomol Screen.* 2016;21(1):35–43. doi: [10.1177/1087057115609144](https://doi.org/10.1177/1087057115609144).
90. Ferrara F, D'Angelo S, Gaiotto T, Naranjo L, Tian H, Gräslund S, Dobrovetsky E, Hrabec P, Lund-Johansen F, Saragozza S, et al. Recombinant renewable polyclonal antibodies. *MAbs-Austin.* 2015;7(1):32–41. doi: [10.4161/19420862.2015.989047](https://doi.org/10.4161/19420862.2015.989047).