Mutational approaches to improve the biophysical properties of human single-domain antibodies

Kim, Dae Young; Hussack, Greg; Kandalaft, Hiba; Tanha, Jamshid

Publisher's version / Version de l'éditeur:
https://doi.org/10.1016/j.bbapap.2014.07.008
Biochimica et Biophysica Acta - Proteins and Proteomics, 1844, 11, pp. 1983-2001, 2014-07-24

Mutational approaches to improve the biophysical properties of human single-domain antibodies

Dae Young Kim[1*], Greg Hussack[1*], Hiba Kandalaft[1], and Jamshid Tanha[1,2,3,4]

[1]Human Health Therapeutics Portfolio, National Research Council Canada, Ottawa, ON, Canada K1A 0R6, [2]Department of Biochemistry, Microbiology and Immunology, University of Ottawa, Ottawa, ON, Canada K1H 8M5, [3]School of Environmental Sciences, Ontario Agricultural College, University of Guelph, Guelph, ON, Canada N1G 2W1

[4]Correspondence to: Jamshid.Tanha@nrc-cnrc.gc.ca *These authors contributed equally.

Keywords: Aggregation, antibody engineering, mutation, single domain antibody, stability, V_H, V_L

Abstract

Monoclonal antibodies are a remarkably successful class of therapeutics used to treat a wide range of indications. There has been growing interest in smaller antibody fragments such as Fabs, scFvs and domain antibodies in recent years. In particular, the development of human V_H and V_L single-domain antibody therapeutics, as stand-alone affinity reagents or as "warheads" for larger molecules, are favored over other sources of antibodies due to their apparent lack of immunogenicity in humans. However, unlike camelid heavy-chain variable domains (V_HHs) which almost unanimously resist aggregation and are highly stably, human V_Hs and V_Ls are prone to aggregation and exhibit poor solubility. Approaches to reduce V_H and V_L aggregation and increase solubility are therefore very active areas of research within the antibody engineering community. Here we extensively chronicle the various mutational approaches that have been applied to human V_Hs and V_Ls to improve their biophysical properties such as expression yield, thermal stability, reversible unfolding and aggregation resistance. In addition, we describe stages of the V_H and V_L development process where these mutations could best be implemented.

1. Introduction

Monoclonal antibodies are a proven class of human therapeutics deployed to treat numerous indications, ranging from cancer to autoimmune disorders to infectious diseases. Recombinant antibody engineering and selection has led to smaller monoclonal antibody fragments including antigen-binding fragments (Fabs), single chain variable fragments (scFvs) and domain antibodies (e.g., variable heavy chain (V_H) and light chain (V_L)) with unique biophysical properties and unique applications, such as medical imaging, deep tumor penetration and multi-target specificity formatting, and opened the possibility for non-injected antibody delivery routes such as topical, inhaled or oral administration [1-4].

Of several key developments in the recombinant antibody field in the past few decades one cannot overlook the growing popularity of single-domain antibodies (sdAbs) (Figure 1). The seminal and prophetic paper by Ward *et al.* [5], in the earlier days of antibody engineering, was the inception point for the concept of sdAbs. The authors showed that V_H domains could act as stand-alone antibody affinity reagents, but unfortunately they also revealed a fundamental problem with V_Hs – their general tendency to aggregate.

Shortly after, another fundamental and related discovery provided the vote of confidence for the viability of the idea of sdAbs as affinity reagents for applications such as immunotherapy: the discovery of the existence of a novel type of antibodies termed heavy-chain antibodies (hcAbs) in the immune system of camelid species [6]. Astonishingly, these antibodies were devoid of light chains and had only V_H at their N-terminal antigen binding region, implying - and soon were shown to be an established fact - that autonomous sdAbs can exist as legitimate and *bona fide* antibody binding fragments just like other classical antibody fragments. The discovery of camelid hcAbs, as well as a few subsequent confirmatory studies, also revealed the aggregation-resistance and strictly monomeric character of hcAbs' antigen binding V_H domain, termed "V_HH", in a sharp contrast to conventional V_H domains shown earlier to be aggregation prone [6-10].

As early studies of camelid V_HHs had pinpointed (a handful of) framework region (FR) residues to play a key role in imparting a non-aggregation character to V_HHs [6-8, 11], initial V_H solubilisation strategies were focused on FR engineering of V_Hs, with the complementarity-determining region (CDR) engineering limited to including an inter-CDR1-CDR3 disulfide linkages in V_H constructs [12-16]. Such strategies were successful in several instances, but it was becoming apparent that CDRs were also playing a complementary role to FRs in the aggregation resistance and monomericity of V_Hs. For example, the role of CDR3 was pointed out to be important in the solubility of a V_HH [10], and the same set of engineered V_H FRs could lead to a monomeric or non-monomeric V_H, depending on the CDR sequence they associated with [17]. Even V_HHs, so characterized by high aggregation resistance and with drastic FR solubilizing mutations, become aggregation prone once their CDR sequences are tampered with [18]. It is now firmly established that both FRs and CDRs are determinants of human V_H domain solubility, supported by numerous studies. This principle also holds true for human V_L domains, which are the second type of human sdAbs pursued as immunotherapeutics.

What gives human V_H and V_L sdAbs the edge in immunotherapy applications over non-human sdAbs like camelid-derived V_HHs or shark-derived V_{NAR}s is their perceived lack of immunogenicity (Figure 1). While inherently unstable and aggregation prone, human V_Hs and V_Ls that match V_HHs in terms of solubility and aggregation resistance can now routinely be isolated at adequate frequency against target antigens, when optimized libraries are used in conjunction with

stringent selection conditions. For the most part, human V_H and V_L binders have been isolated from synthetic sdAb display libraries by evolutionary approaches, with library diversity generated through CDR/loop randomization on a single library scaffold. The affinity of V_H and V_L binders isolated from synthetic human sdAb display libraries are typically higher than sdAbs isolated from naïve V_HH display libraries, due to the synthetic libraries' higher CDR/loop sequence complexity, but lower than those obtained from immune V_HH libraries which are enriched with *in vivo* affinity-matured V_HHs.

Other sources of human sdAb binders include naïve human sdAb display libraries and the Crescendo Mouse transgenic mouse platform (http://www.crescendobiologics.com), which uses a mouse knock out for murine antibodies and transgenic for human hcAbs that are similar in structure to camelid hcAbs. Crescendo's transgenic mouse technology promises to provide V_HS with high target affinity, high stability and high solubility, which are common characteristics of V_HHs derived from immunized libraries, but unlike V_HHs in that they do not require further humanization. If required, routine techniques such as affinity maturation experiments combined with evolutionary selection approaches involving both affinity and stability selection pressures, high throughput functional screenings and deep sequencing could be utilized to obtain highly stable and high affinity human V_HS and V_LS, similar to immune V_HHs and $V_{NAR}S$.

To underscore the growing enthusiasm towards human sdAbs, a number of companies and research and technology organizations, such as GSK/Domantis, are focused on the development of V_H- and/or V_L-based therapeutics, some of which have human sdAbs in clinical trials (http://clinicaltrials.gov; http://www.gsk.com).

The focus of this review is to chronicle the mutagenic approaches available for improving the biophysical properties of human sdAbs and to outline at which stages of the development process these mutations could best be implemented. For additional strategies, we direct readers to several other excellent reviews describing approaches to engineer variable domains in the context of therapeutic IgGs [19], camelid and shark variable domains [4, 20], as well as human constant domains, human variable domains and other antibody fragments [1, 20-22].

2. Biophysical properties of human sdAbs and tools for their study

Aggregation and denaturation is a concern in the manufacturability and efficacy of human diagnostic and therapeutic antibodies, as antibody aggregation has been implicated in reduced *in vivo* efficacy and increased immunogenicity [23-26]. Therapeutic antibodies must possess the ability to withstand physical and chemical stresses, and mechanical agitations, such as high protein concentrations, elevated temperature, extreme pH and ionic strength, freeze drying/filtration/centrifugation, and exposure to detergents, organic solvents and proteases [27].

The physical stability of an antibody encompasses both colloidal [28, 29] and conformational (also known as thermodynamic) stabilities, which governs an antibody's resistance to aggregation and unfolding. Inter- and intra-molecular forces determine a protein's propensity to aggregate – a complex process where proteins form non-native, but energetically favorable states [30]. This process can occur through multiple pathways depending on the proteins biophysical properties.

Human sdAbs are classified into seven families (V_H1a, V_H1b, V_H2, V_H3, V_H4, V_H5 and V_H6 families for V_Hs; $V_\kappa1$, $V_\kappa2$, $V_\kappa3$, $V_\kappa4$, $V_\lambda1$, $V_\lambda2$ and $V_\lambda3$ families for V_Ls) [31], which differ in their biophysical properties [32]. Odd numbered V_H families (V_H1, V_H3 and V_H5) showed superior biophysical properties to even numbered V_H families, with the V_H3 family exhibiting the highest stability and expression yield. Fewer differences were observed among the V_L families, with V_κ families exhibiting higher thermodynamic stability and expression yields than V_λ families. Almost all the V_L families in this study behaved as monomers, with the exception of $V_\kappa4$, and $V_\lambda3$ which existed as monomer-dimer equilibriums. Many V_Hs and V_Ls have also been known to bind protein A and protein L, respectively, without interfering with antigen binding [33-36]. The binding of V_Hs to protein A is restricted to the V_H3 family, which represents half of the inherited V_H genes [37-41], while protein L binding by V_L domains is restricted to families $V_\kappa1$, $V_\kappa3$, and $V_\kappa4$ [42-45]. This ability can be exploited for use in technological applications such as tag-free purification of domain antibodies through protein A- or protein L-conjugated sepharose affinity columns [46] and purging synthetic phage-display libraries from aggregating, non-functional domains [47, 48].

Some of the factors that influence thermodynamic and colloidal stability of human sdAbs are germ-line origins of framework regions, CDR3 loop length, the presence of hydrophobic residues at the V_H/V_L interface and CDR, net charge, and the degree of hydrophobic core packing [32]. These factors can be addressed through several protein engineering techniques (Table 1; discussed below), to produce highly non-aggregating, thermodynamically stable domains suitable for therapeutic and diagnostic applications. Moreover, selecting scaffolds from germ-line families with enhanced biophysical properties for library construction greatly increases the likelihood of isolating aggregation-resistant domains [48-51].

Several methods can be used to study the biophysical properties of antibody fragments, and these techniques have been discussed extensively elsewhere [52]. Differential scanning calorimetry (DSC), circular dichroism spectropolarimetry (CD) and intrinsic tryptophan fluorescence spectroscopy have traditionally been used to study protein thermodynamic and pH stability, stability in the presence of denaturants, and to obtain structural information [53, 54]. Fourier transform Raman and infrared spectroscopy [55] can also be employed to study a protein's secondary structure. Differential scanning fluororimetry is an alternative high-throughput

method for determining the thermal stability (T_m) of a protein [56]. The colloidal stability (i.e., the propensity of the folded protein to precipitate [57]) can be studied using dynamic or differential static light scattering, turbidity analysis, size-exclusion chromatography (SEC) and analytical ultracentrifugation. SEC is one of the most widely used techniques to determine the aggregation state of proteins, and when coupled to multi-angle laser light scattering (MALS), protein peaks eluting off a gel filtration columns can be analyzed by MALS to determine molar mass [58],

3. Site-directed mutational approaches for improving biophysical properties of human sdAbs

In this section we summarize site-specific mutations introduced into the CDRs or the FRs of human V_Hs and V_Ls to improve their biophysical properties, and discuss their mutational effects (Table 2). The table contains only the mutations identified by the studies using human V_H or V_L domains and therefore many other mutations studied in the context of Fv, scFv, Fab, and IgG were omitted due to the scope of the review. We have included a supplementary table (Supplementary Table 1) which contains comprehensive details regarding the mutations described in Table 2. For consistency, all mutations are numbered as defined by Kabat *et al.* [59].

3.1. Mutations affecting biophysical properties of human V_Hs

In an effort to obtain antigen-specific V_H domains from antibody producing B cells, Ward *et al.* [5] immunized mice with either lysozyme or keyhole limpet haemocyanin (KLH) and subsequently cloned repertoires of V_H genes from the mouse spleen genomic DNA into *E. coli* for V_H secretion. Screening of several thousand individual colonies by ELISA, performed on bacterial supernatant containing the secreted V_H domains, identified many supernatants with antigen binding activities. V_H domains from the positive colonies were expressed and shown to be target specific, with affinities in the 20 nM range. Moreover, it was shown that V_H domains with antigen binding activity could also be obtained from naïve mice, albeit with lower frequency. However, the V_H domains were aggregation prone, and this was attributed to their exposed hydrophobic surfaces at the V_H/V_L interface.

Engineering efforts to solubilize V_Hs, therefore, initially focused on mutating suspect interface residues concentrated in the FR2 regions, but gradually expanded to other FR residues and CDRs. The same paper that reported the discovery of naturally occurring antibodies devoid of light chains (hcAbs) and suggested that hcAbs bind to their antigens through autonomous non-aggregating V_H domains also presented structural data, albeit limited, that explained how such antigen-binding V_H domains could stay non-aggregating [6]. The paper revealed the sequences of two V_H domains derived from hcAbs, later known as V_HHs, which had several significant mutations

compared to conventional V_Hs. Notable differences were substitutions of highly conserved residues at the former V_H/V_L interface positions: V37F/Y, G44E, L45R and W47G/F. These substitutions suggested a mechanism for how V_HHs could exist as strict monomers and provided a mutational approach for converting aggregating V_Hs to aggregation-resistant ones.

Initial efforts by Davis and Riechmann [12] to analyze a human V_H (Ox13-VH) by Nuclear Magnetic Resonance (NMR) spectroscopy was unsuccessful due to a broad spectral line width; this was attributed to the aggregation prone nature of the V_H. Using the sequences of the two previously published V_HHs [6] as a guide, they substituted the V_H residues with corresponding camel V_HH residues at three V_L interface positions (G44E/L45R/W47G), a process which they referred to as 'camelization'. The mutant, VH-P1, had drastically improved NMR line width, an indication it became aggregation resistant. Interestingly, the stability and expression yield (in *E. coli*) of VH-P1 was lower than the aggregation-prone parent Ox13-VH. A second camelized version, VH-P8 (G44E/L45R/W47I), was similar to Ox13-VH in terms of expression yield and stability but had poorer aggregation resistance compared to VH-P1. In the presence of detergent (CHAPS), the NMR line width of VH-P8 was improved, matching that of VH-P1. These studies highlighted the importance of a few interface residues in the biophysical behaviour of V_Hs and how even one amino acid difference at the V_H/V_L interface can cause differential biophysical behaviour.

Since the VH-P8 variant expressed well, it was chosen for detailed NMR structural studies with the aim of using the structural information towards optimizing V_Hs as affinity reagents. NMR experiments carried out in the presence of CHAPS showed that the secondary structure of VH-P8 was identical to that for V_Hs that are associated with V_Ls in intact antibodies [33]. The NMR 3D structure revealed that the architecture of the two pleated β-sheets in the Ig fold, as well as the H1 and H2 loop conformations, were similar to those of non-modified V_L-associating V_Hs in intact antibodies [15]. However, the H3 loop had a different conformation when compared to that of a V_L-associating V_H (D1.3-VH) of identical CDR3 length. Instead of orienting towards the (absent) V_H/V_L interface as in the case of D1.3-VH, the H3 loop of VH-P8 was oriented towards the H1 and H2 loops. Other notable changes in VH-P8 were the absence of the salt bridge between R94 and D101, typically observed in V_L-associated V_Hs, the solvent-exposed orientation of the now-liberated D101 side chain, and the conformational changes of the interface residues V37, W103 and R38, which lead to the burial of the side chains of V37 and W103 and exposed the R38 side chain in VH-P8. This is in contrast to V_L-associating non-camelized V_Hs, where V37 and W103 side chains are "exposed" to interact with the V_L, and the R38 side chain is buried. These conformational changes and the camelized mutations (G44E/L45R/W47I) increased the hydrophilicity of the interface significantly thereby increasing its aggregation resistance.

Interestingly, the crystal structure of cVH-E2, a derivative of VH-P8 which differs in the length and sequence of CDR3 and the presence of one additional camelization mutation at position 37 (V37F), revealed that the overall folds are essentially the same, but the conformation of the V_H/V_L interface involving the camelized residues were significantly different [60], suggesting V37F camelization as an underlying factor. The different V_H/V_L interface conformation, however, still serves to suppress aggregation, and the fourth camelization mutation at position 37 (Val to Phe or Tyr) was suggested to be necessary for maintaining the "true structural integrity" of camelized V_Hs.

To develop antigen-binding V_H domains, camelized VH-P1 (G44E/L45R/W47G) was used as a scaffold to construct a V_H phage display library, with library diversity generated through CDR3 randomization [13]. Many anti-protein and anti-hapten V_Hs were identified and demonstrated affinities (for haptens) in the range of 100 - 400 nM, with no evidence of non-specific binding. Expression yields of camelized V_Hs increased significantly when a G47I mutation was introduced, with only small changes in affinity. NMR analysis of one of the G47I mutant binders indicated the V_H was highly soluble and stable. An affinity-improved version of the hapten binder VH-Ox21 was identified from a phage display library. This V_H (VH-OS1), which had an affinity of 25 nM compared to 146 nM for VH-Ox21, had two mutations in CDR1 with respect to VH-Ox21: A33Y and S35W [61]. Interestingly, when introducing the G47I mutation into the camelized version (44E/45R/47G) of VH-OS1, it became sticky and bound to its target as well as many other ligands, which is an indication of aggregation. However, when comparing the I47 version of VH-OS1 to the non-sticky I47 version of VH-Ox21, it became clear that CDR1 A33Y and S35W mutations are implicated in the sticky nature of VH-OS1, as these are the only two differences. The presence of more hydrophobic amino acids at positions 33 and 35 of VH-OS1 is consistent with this view. Position 35, in particular, has been implicated in the solubility of V_Hs in a number of studies [62-66]. In a second affinity maturation approach, a phage display library based on the camelized V_H VH-Ox21 was constructed by randomizing CDR2 residues 52, 52a and 53, and panned for affinity improved V_Hs. Phage Enzyme-linked Immunosorbent Assay (ELISA) revealed that the binders were a mix of aggregating and non-aggregating antibodies, underlining the role of CDR sequence in the aggregation behaviour of V_Hs. One V_H, VH-Ox21.2.4 (S52G/G52aL/S53P and expressed with I47), exhibited a K_D of 47 nM compared to 185 nM for the parent VH-Ox21. Construction of a similar CDR2-randomized library based on a nitrophenacetyl (NIP)-specific camelized V_H, VH-N3c1 [13], also yielded many positive clones. One clone in particular, VH-N3c1.2.1 (S52A/G52aP/S53G and expressed with I47), was non-sticky and bound to its target with a K_D of 31 nM, which was 10-fold stronger than the wild-type VH-N3c1 in the I47 format. Finally, an affinity improved version of the anti-lysozyme binder, VH-LS2 (VH-LS2.5.1), isolated

from a CDR2 phage display of VH-LS2 was expressible as a soluble V_H in both G47 and I47 camelized formats. The mutant, which differed from the wild-type V_H at two CDR2 positions (N52 and W47), bound specifically to lysozyme.

In a series of experiments, termed "camelization II", Davies and Riechmann [14] set out to determine the effect of incorporating structural features that were seen as characteristics of camel V_HHs [11] on the stability of a primary-camelized (44E/45R/47G) V_H, VH-Ox21-AVRL. These included a V37F mutation, an R94A mutation and incorporation of an inter-CDR1-CDR3 disulfide linkage connecting C33 and C100b. A V37F mutation led to a T_m increase of 1.6°C. Incorporation of Cys residues at positions 33 and 100b of VH-Ox21-AVRL and VH-Ox21-AFRL resulted in the formation of intra-domain disulfide linkages in both, with no trace of inter-domain disulfide linkages, as was shown by SDS-PAGE. This led to T_m increases of 1.3°C and 4.5°C for VH-Ox21-CVRC and VH-Ox21-CFRC, without affecting their proper folding as determined by CD experiments. However, the additional disulfide linkages led to decreased expression yield in *E. coli*, which is consistent with more recent findings of disulfide linkage-engineered sdAbs [35, 36, 67]. Finally, an R94A mutation increased the T_m of VH-Ox21-AFRL to 63.9°C, but decreased the T_m of the more camelized VH-Ox21-CFRC to 58.5°C. Although the beneficial effect of the secondary camelization on the stability of VH-Ox21 was small, V_H binders that had the same camelization mutations as VH-Ox21-CFRC (A33C/V37F/G44E/L45R/W47G/L100bC) were isolated from a V_H phage display library that were much more stable than VH-Ox21-CFRC. Specifically, an phenyloxazolone (Ox)-BSA specific V_H (Oct6) and an anti-lysozyme V_H (VH-Lc1) were found to have the intended intra-domain CDR1-CDR3 disulfide linkage, were properly folded, were non-sticky as they only bound to their target antigen, and importantly, had T_ms values far higher (average ΔT_m = 12.9°C) than VH-Ox21-CFRC, which differs only in CDR3, underlining the importance of CDR3 in folding stability of V_Hs. Unfortunately, increases in stability did not translate to higher expression yields as all three V_Hs (VH-Ox21-CFRC, Oct6 and VH-Lc1) had similar, low expression yields below 1 mg/L of bacterial culture. However, as observed previously for other V_Hs with primary camelized mutations, the expression yields of the Oct6 and VH-Lc1 V_Hs were increased with the G47I mutation to 7 – 8 mg/L of bacterial culture without compromising affinity and specificity. The G47I mutation also increased the T_ms of VH-Ox21-AVRL and VH-Oct 6 by ~11°C and 1°C, respectively. Interestingly, it was found that I47-containing V_Hs with a lower T_m (VH-Ox21-AVRL) had higher expression yield than a G47-containing V_H with higher T_m (VH-Oct6). This was attributed to the stabilizing effect of the G47I mutation during the *in vivo* protein folding pathway, leading to a higher ratio of folded to misfolded proteins that eventually would be lost to aggregation and/or proteolysis.

A similar camelization strategy was applied to an aggregating human V_H (BT32/A6 V_H) of a different family (V_H3) and it was shown that camelization improved solution properties of V_Hs [16]. For example, while the non-camelized BT32/A6 V_H had size exclusion chromatogram and NMR spectral features typical of aggregating V_Hs, camelized BT32/A6.ERG (44E/45R/47G) and BT32/A6.ERI (44E/45R/47I) were strictly monomeric V_Hs. Interestingly, the collection of high quality NMR data for camelized V_Hs was achieved without CHAPS, in contrast to a previously described camelized V_H [15] which required CHAPS to reduce aggregation and achieve adequate solubility, indicating the good quality of BT32/A6 V_H as a library scaffold. Camelized BT32/A6 V_H with several mutations (33C/44E/45R/47G/93A/94A/100eC) was used as the scaffold to make a phage display library with diversity restricted to CDR3 and a pair of Cys residues at positions 33 and 100e to facilitate the formation of an inter-CDR1-CDR3 disulfide linkage, a feature seen frequently in V_HHs. By mass spectrometry, it was shown that the inter-CDR1-CDR3 disulfide linkage was formed, but it did not necessarily impart favourable biophysical properties. For example, while the camelized R3A10.G47I V_H with the inter-CDR disulfide linkage had partially sharp NMR spectra, R3A10.G47I/Cys$^-$ version lacking the disulfide linkage had significantly enhanced solubility and dramatically improved NMR spectral properties, while maintaining its intact tertiary structure. Several binders identified against one "antigen" (3B1 scFv) did not show signs of stickiness by phage ELISA. Many more binders against a second "antigen" (M2 IgG) were also obtained and several of these, when expressed in *E. coli* as soluble V_Hs, showed good expression yields (>10 mg/mL) and were strictly monomeric as determined by SEC.

To develop human V_H domains that exhibit the non-aggregating properties of camelid V_HHs but without relying on the CDR3 loops, Barthelemy *et al.* [66] performed a comprehensive study using combinatorial phage-displayed libraries coupled with protein A selection and biophysical characterization. A V_H phage display library (Library A) using human VH-4D5 (V_H3 family) as the scaffold was generated by randomizing CDR3 (93-102; length variation: 7 – 17) and positions 35, 37, 45 and 47, which are known to have solubilization activity in V_Hs and V_HHs [49, 62-66]. Protein A binding has been known to correlate with intrinsic V_H stability, so the library was panned against protein A to select for correctly folded V_Hs [18, 47, 68, 69]. Analysis of twenty five selected sequences revealed no preference for solubilization residues typically found in V_HHs at interface positions 37, 45 and 47 or for hydrophobic residues in CDR3, which contrasts a previous study with a V_HH domain where hydrophobic residues were strongly biased in the CDR3 [18, 68], suggesting that the majority of selected V_H domains are not stabilized by the CDR3-framework interactions seen with camelid V_HHs. Biophysical analysis of six V_Hs revealed that compared to the parent VH-4D5, which had an expression yield of 1 mg/L, a T_m of 58°C and irreversibly unfolded, four of six V_Hs had improved soluble expression yields (7 – 14 mg/L), T_ms

(65°C – 73°C), monomeric SEC behavior and reversible unfolding. In two V_Hs, position 35 was occupied by Gly, suggesting that the V_H domains were stabilized by a similar mechanism observed in an aggregation-resistant V_H domain (HEL4; discussed below) which also contained G35 [62]. To explore if the V_H biophysical properties could further be improved by targeting additional residues, another library (Library B) was generated by using the lead VH-A1 (from Library A) as the scaffold and randomizing at ten positions, (37, 39, 44, 45, 47, 50, 91, 103 and 105) at the V_H/V_L interface and position 35, with no CDR3 length variation (i.e., same CDR3 length as VH-A1 that is eight residues long). Library B was subjected to protein A selection and 384 unique sequences were analysed. The parent VH-A1 residues were conserved at most positions, Gly was favored at position 35, consistent with other observations [62, 63, 65, 66] and position 103 was favorably occupied by Arg instead of the parental Trp, indicating that W103R might improve V_H stability as suggested previously [70]. Of the Library B clones, only one V_H (VH-B1) showed a better expression yield (19 mg/L) than the parent VH-A1 (14 mg/L). VH-B1 was essentially monomeric unfolded reversibly and had a high T_m nearly identical to the parent. To determine the structural contribution of CDR3 residues in V_H domains, libraries were generated by using eleven V_Hs (ten VH-B variants and VH-A1) as scaffolds, randomizing each CDR3 residue for wild-type (WT) and Ala residues followed by selection for protein A binding. WT/Ala ratios were close to one at all positions for most of the V_H domains including, VH-A1 and VH-B1, indicating that stability in the V_Hs was CDR3-independent, which contrasted the results from a V_HH domain (anti-HCG) that showed a strong reliance of stability on hydrophobic CDR residues at particular positions. To determine the structural contribution of the residues present in the V_L interface towards stability, a V_H library was generated by using VH-B1 as a scaffold and randomizing nine positions near the V_H/V_L interface (37, 39, 44, 45, 47, 50, 91, 103 and 105), and position 35, and the library was subjected to selection for protein A binding. Parental residues were conserved at most positions but no position was biased toward a particular amino acid and most positions could be replaced with small and/or hydrophilic residues. This suggested improving the solution properties of VH-B1 through additional point mutations at the V_L interface, without compromising stability, may be possible. This contrasted results from a V_HH domain (anti-HCG) that showed a strong bias to parental residues at all positions, indicating the stability of camelid V_HH domains is dependent on the hydrophobic residues present in the V_H/V_L interface and in CDR3. To further test the capacity of VH-B1, the V_H was engineered to have the same sequences as VH-4D5 except for five mutations (H35G, Q39R, L45E, R50S, and S93A) and was named VH-B1a. VH-B1a expressed well, was monomeric, unfolded reversibly with a higher T_m (79°C) than VH-B1, but an increased SEC elution time indicatedVH-B1a might stick to the column. The crystal structure of VH-B1a revealed that the main chain of VH-B1a and VH-4D5, when in association with V_L, superimposed with a

minor deviation but there were significant differences in the CDR3 conformations, indicating the CDR3 loop is flexible and may be able to adopt multiple conformations in solution. Finally, to optimize VH-B1a, exposed hydrophobic residues were mutated and the variants characterized. W47 was replaced with five possible residues (W47A, W47V, W47L, W47E and W47T) to reduce hydrophobicity and all mutations showed improved SEC characteristics with minimal impact on thermal stability. Additional hydrophilic mutations were introduced in the W47L mutant at positions 37 and 103 (W47L/V37S and W47L/V37T; W47L/W103R, W47L/W103S, and W47L/W103T). VH-B1a variants containing the W47L/V37S or W47L/V37T exhibited excellent SEC behavior and reversible unfolding but reduced the T_m (72°C compared to 79°C for VH-B1a). VH-B1a variants with Ser or Thr substitutions at position 37 in the W47L mutant showed good SEC behavior. However, VH-B1a variant with W47L/W103R was aggregating (dimer by SEC) although it was highly thermostable (T_m = 83°C) and unfolded reversibly. Two additional hydrophilic substitutions were made in the W47L mutant at residues R39 and E45 (W47L/R39S, W47L/R39T, W47L/R39H, W47L/R39K, W47L/R39D, and W47L/R39E; W47L/E45H, W47L/E45S, and W47L/E45T), revealing that while at position 39 all six mutations were tolerated well, at position 45 the E45H substitution was tolerated but the E45S substitution significantly increased SEC elution time and the E45T substitution caused severe aggregation. The contributions of individual mutations to VH-B1a behaviour was then examined by back mutating 4 positions (35, 39, 45, and 50). The replacement of G35 by His significantly reduced the fraction of the monomer, while thermal stability or reversible unfolding behaviour was unchanged. The R39Q back mutation slightly increased T_m and reduced the monomeric fraction. The E45L substitution increased T_m and also increased the SEC elution time dramatically, indicating that the hydrophobicity of Leu promoted stickiness to the column. The S50R substitution reduced T_m but improved the monomer fraction and reduced elution time. In the background of W47L or W47T, the S50R substitution improved the monomeric fraction in all cases. Taken together, the hydrophilic substitutions at positions 39, 45 and 47 improve the solubility of human V_H domains. G35 is beneficial when a Trp is present at position 47 and not required when W47 is substituted by less hydrophobic, or hydrophilic residues. Based on the back mutation analysis, the R50S substitution is not appropriate for the autonomous behaviour of human V_H domains. The mutations identified by Barthelemy *et al.* [66] are rarely found from natural V_H and V_HH repertoires, highlighting the importance of the comprehensive and unbiased approach in recovering mutations improving stability and solubility of human sdAbs that helps to expand the boundary of antibody engineering beyond the natural sequence diversity. Subsequent studies showed the suitability of VH-B1a as a library scaffold for V_H synthetic libraries [50]. A derivative of VH-B1a with three mutations (W47L/S50R/W103S) was used to construct libraries by CDR randomization. Panning against

VEGF identified many binders with 2 with IC_{50} of ~60 nM. Further affinity maturation yielded one binder with affinity as low as K_D = 10 nM. Interestingly, the X-ray crystal structure data show that the V_H paratope is formed by combining CDR3 loop and former V_H/V_L interface.

From a naïve llama sdAb phage display library, Tanha *et al.* [71] isolated many anti-idiotypic sdAbs that were structurally V_Hs, but characteristically behaved like V_HHs. For example, while many sdAbs contained V37, G44, L45 and W47 that are typical of conventional V_Hs, they demonstrated reversible thermal unfolding, strict monomericity, high solubility and aggregation resistance as demonstrated by size-exclusion chromatography, high quality NMR data [72], and high expression yields in *E. coli* (70 – 80 mg/L). The V_Hs showed a remarkable homology to human V_H3 family, and a FR sequence comparison between these V_Hs and the human V_H3 family consensus sequence nominated five amino acid substitutions to be responsible for the aforementioned biophysical properties of llama V_Hs: E6A, S74A, R83K, A84P and L108Q [73]. Thus, in subsequent studies it was shown that the solubility and thermal refolding efficiency of a typical V_H (BT32/A6) was improved by introducing E6A and L108Q mutations in FR1 and FR4, respectively, to generate BT32/A6.L1 with an improved thermal refolding efficiency from 13% (parent) to 44% (BT32/A6.L1). Incorporation of three more mutations (S74A/R83K/A84P) into FR3 of BT32/A6.L1 generated BT32/A6.L2 and resulted in a further increase in thermal refolding efficiency from 44% to 74% without compromising solubility. For both mutants the aforementioned biophysical gains were not at the expense of expression, as both had yields similar to BT32/A6. A marginal but incremental 1°C increase in T_ms was observed for each mutation set. Interestingly, the thermal refolding efficiency of both mutants was significantly higher than BT32/A6.ERI (27%), a camelized version of BT32/A6 with G44E/L45R/Y47I mutations that possessed excellent solution properties [16] and similar T_m (63.7°C) to BT32/A6. It was suggested that the mutations improved the thermal refolding efficiency of BT32/A6.L1 and BT32/A6.L2 by stabilizing the folding intermediates that lead to the native structure *in vivo* [73]. The modest improvement in thermal refolding efficiency of BT32/A6.ERI over BT32/A6 suggests that the camelization mutations at position 44, 45 and 47 may not have a significant role in the reversible thermal unfolding of V_HHs, although this may be a simplistic view of a more complex sequence context dependent situation. A phage display library constructed by CDR3 randomization on the BT32/A6.L1 scaffold yielded V_Hs with good antigen binding activity and aggregation resistance, demonstrating that with as little as two presumably non-human mutations (E6A and L108Q) one could obtain immunotherapeutics with almost 100% human sequences in nature.

Mutations introduced into the CDRs are another strategy shown to be effective in improving solubility and aggregation resistance of human V_Hs [47, 62-65]. HEL4 (K_D = 130 nM),

an anti-lysozyme V_H isolated from a human V_H phage display library, was well expressed in *E. coli* (9.5 mg/L), was monomeric by SEC and analytical centrifugation at concentrations up to 200 μM and highly non-aggregating at high concentrations post thermal denaturation, despite the presence of presumable aggregation-promoting residues at the former V_H/V_L interface (G44/L45/W47) [62]. In contrast, DP47d, which differs from HEL4 only in CDR sequences, expressed in lower yield (2.9 mg/L), was sticky, irreversibly unfolded and formed aggregates when heated above 55°C, despite having a significantly higher thermodynamic stability than and a similar T_m to HEL4. When comparing HEL4 to V_L-associating V_Hs, crystallographic data revealed an unusual flipping of the HEL4 W47 side chain into a cavity formed by G35, V37 and A93, leading to an increase in the hydrophilicity at the former V_L interface. Introduction of a S35G mutation, which allowed a cavity to form, or a W47R mutation into DP47d reduced its stickiness, but compromised its thermodynamic stability. Neither mutation, however, improved reversible thermal unfolding. In addition to the W47 flip, other residues including those in CDR3 were also suggested to play a role in the biophysical behavior of HEL4 [62]. Furthermore, it was thought that while the W47 flip contributed to the ability of the native state of HEL4 to resist aggregation, it did not help in terms of HEL4's ability to resist aggregation following thermal unfolding. In subsequent experiments, Jespers *et al.* [47] developed an elegant method that efficiently selected aggregation-resistant V_Hs from phage display libraries that were overwhelmingly populated by aggregation prone V_Hs. The method involved heating and cooling phage display libraries that displayed multiple copies of V_Hs on the surface of phage at 80°C for 10 min to denature V_Hs and promote aggregation (hence inactivation). Panning experiments are performed against a target ligand to select V_Hs that unfolded reversibly (aggregation resistant), over those that did not (aggregation prone). The method was applied to a DP47d V_H-based multivalent human V_H phage display library (library I), whose members universally bound protein A. Several cycles of heat denaturation and panning against protein A revealed that the majority of the two hundred V_H phage tested retained their protein A binding after heating. A unique subset of the binders (twenty V_Hs) only differed in their CDR sequence and length, thus demonstrating that mutations in the CDRs alone could transform aggregation prone V_Hs into aggregation resistant ones. Interestingly, eighteen of the twenty V_Hs had an acidic theoretical pI. Six of the twenty V_Hs were analysed for biophysical properties and, similar to V_HHs, expressed at high yields in *E. coli* (22 - 39 mg/L compared to 2.9 mg/L for DP47d), were non-sticky and monomeric by gel filtration and unfolded reversibly following two cycles of thermal denaturation. A second, larger, DP47d V_H-based phage display library (library II) with different randomization positions in the three CDRs was panned against human serum albumin (HSA), with or without the heat-denaturation step. While heat-denaturation selection yielded one anti-HSA V_H (H26), selection without heat-

denaturation yielded six unique anti-HSA V_Hs; however, only H26 showed complete reversible unfolding both on the surface of phage and was soluble at 50 μM, underlining again how mutations in CDRs alone were sufficient to render V_Hs aggregation resistant. Consistent with findings from library I, the aggregation resistant H26 had an acidic pI (5.7), whereas the six aggregation prone V_Hs did not (pI = 7.4 ± 1.2, mean ± SD). It was hypothesized that the CDR mutations conferred aggregation resistance by reducing the aggregation tendency of the unfolded (or partially unfolded) states of V_Hs, rather than by stabilizing their native state. This hypothesis was based on the fact that the aggregation resistant V_Hs had significantly lower thermodynamic stability than aggregation prone V_Hs with the same human V germ-line segment, including DP47d [47, 74]. While the high expression yield of the selected V_Hs was explainable in terms of their resistance to aggregation [47], their lack of stickiness to the SEC column matrix, despite the absence of any solubilizing camelid residues at the former V_L interface, may be due to conformational changes that make the interface more hydrophilic as seen in the case of HEL4 [62], or to the folding of CDR3 loops over the V_L interface residues as seen for a V_H and V_HHs [10, 72]. A generalized infectivity-based selection strategy, based on the aforementioned approach [75], when applied to the DP47d phage Library I [47], yielded phage V_Hs that retained their binding after heating. The aggregation tendencies of the soluble forms of five of the selected V_Hs were determined by monitoring their solution turbidity at 80°C in phosphate-buffered saline (PBS). While DP47d formed aggregates within 1 min, none of the five selected V_Hs formed any aggregates for 10 min, three had longer turbidity lag times than HEL4 and one did not aggregate following a 15 h incubation at 80°C. More recently, it was shown that the heat denaturation approach [47] enables the rapid prediction and ranking of large numbers of soluble V_Hs, with respect to expression and refolding efficiencies [76].

Christ *et al.* [77, 78] employed the heat denaturation approach [47] to generate a library highly enriched in aggregation resistant human V_H domains. Following heat-induced aggregation of a previously described V_H phage display library (Repertoire U;[62]), a sub-repertoire of phage containing aggregation resistant V_Hs (Repertoire S) was selected through binding to protein A. By combinatorial assembly of the "aggregation-resistant CDRs" from repertoire S, a large aggregation-resistant repertoire (Repertoire C) was generated. While repertoire C was similar in size to the original unselected library (2×10^9 clones), monoclonal phage ELISA showed that it had a significantly higher number of aggregation resistant clones (71% versus 8%, respectively), demonstrating the effectiveness of the combinatorial CDR assembly approach in generating V_H repertoires highly enriched in aggregation resistant V_Hs. The resulting V_H phage display library yielded affinity reagents that were both aggregation resistant and of good antigen binding affinities (K_D = μM – nM). By SPR one antigen-specific V_H maintained half of its activity following

heating (80°C, 1 μM) while a second V_H ("Clone 2") was fully active after heating (80°C, 5 μM). The thermodynamic stability of Clone 2 was lower than those of aggregating V_Hs with the same (DP47/V3-23) framework.

To identify CDR sequence determinants of aggregation in human V_Hs, Dudgeon *et al.* [65] applied the heat denaturation and protein A capture method to a human V_H phage display library (Library I, [47]). The resulting heat-selected repertoire was more aggregation resistant compared to an unselected control repertoire, as determined by phage ELISA performed on 40 clones from each pool. A sequence comparison of heat-selected and unselected clones revealed significant bias for V29, G35 and E32 in CDR1, R52 in CDR2 and S102and V100lin CDR3, suggesting these amino acids contribute to the resistance of heat-induced V_H aggregation. With respect to physico-chemical properties, the heat selected V_Hs had a significant increase in net negative charge and a reduction in hydrophobicity and β-sheet propensity compared to unselected V_Hs and CDRs.

By using a phenotypic phage plaque selection method on a naïve human V_H phage display library, To *et al.* [48] identified more than a dozen human V_Hs with diverse V_H3 V/J germ-line sequences that were monomeric and additionally demonstrated desirable properties such as protein A binding, excellent soluble expression yields in *E. coli* (up to 60 mg/mL), good refolding efficiency following thermal denaturation, resistance to aggregation following a long incubation at 37°C (17 days in PBS up to 3 mg/mL), and resistance to trypsin digestion. The identification of determinants of the observed biophysical properties was not possible, although for a few V_Hs, a V37F mutation, implicated in the solubility of V_HHs, may have influenced solubility and stability. One of the non-aggregating V_Hs, HVHP430, was used as a scaffold to construct a relatively large phagemid vector-based V_H phage display library with a disulfide bond-constrained H3 loop and diversity residing in CDR1/H1, CDR3 and at position 94 [63]. Selecting the library in a monovalent format against several targets (α-amylase, lysozyme and carbonic anhydrase) resulted in binders that were aggregating by SEC and precipitated at 4°C shortly after purification, further highlighting again the role of CDRs in the aggregation resistance of V_Hs. When the library was selected in a multivalent display format (by incorporating hyperphage M13KO7ΔpIII for super-infection) with a transient heat denaturation approach [47], binders with a significantly higher mean and median monomericity index were isolated, with four being strictly monomeric, non-aggregating and showed heat refolding efficiencies of ~90 – 100%. Furthermore, this selection approach enriched for V_Hs with acidic theoretical isoelectric points (pIs) and inter-CDR1-CDR3 disulfide linkages – features associated with non-aggregating V_HHs as well – and suggested the importance of an inter-CDR1-CDR3 disulfide linkage and acidic residues in imparting stability and aggregation resistance to V_Hs. Consistent with the above suggestion, three out of the four non-aggregating V_Hs had acidic pIs, whereas 16 of the 17 aggregating V_Hs had non-acidic pIs, with the

one aggregating V_H being only slightly acidic (theoretical pI = 6.4). A pI distribution analysis of camelid V_HH and human V_H pools revealed a higher proportion of acidic domains in the non-aggregating camelid V_HH pools versus a higher proportion of basic domains in the aggregating camelid or human V_H pools [63]. A tendency for aggregation-resistant V_HHs to have acidic pI was also reported previously [47, 62]. Although these studies did not pinpoint whether it is the global pI effect or local pI effect in CDRs that was at work, it does propose incorporation of acidic residue into human V_Hs as a strategy of imparting aggregation resistance, as was demonstrated in subsequent studies (see below).

The heat-denaturation method described above was efficient for isolating aggregation resistant antibody domains, but the domains had relatively low thermodynamic stability. Using an analogous "acid-denaturation" method applied to a V_H phage display library, Famm *et al.* [79] showed that sdAbs with both high thermodynamic stability and aggregation resistance can be obtained simultaneously. The method involved incubating a V_H phage display library [47] in acidic conditions that would promote V_H aggregation on the surface of phage (2 h, pH 3.2, 37°C), followed by neutralization and subsequent binding against protein A to select for refolded V_Hs. Following a few cycles of panning under acidic conditions, the library was highly enriched for V_Hs which resisted acid-induced aggregation. By turbidity assays, four of the selected domains, expressed and purified as soluble proteins, demonstrated superior resistance to acid-induced aggregation compared to the aggregation-prone DP47d V_H and the aggregation-resistant HEL4 V_H. Importantly, compared to the four heat-induced aggregation-resistant V_Hs isolated from the same library by the heat-denaturation method, the four acid-induced aggregation-resistant V_Hs had higher apparent T_ms and thermodynamic stability at pH 3.2 and 7.0. Hence, it was suggested that thermostability was a factor in imparting aggregation resistance to the acid-selected V_Hs and not to the heat-selected V_Hs [79]. Moreover, comparing the sequences of acid-selected V_Hs to V_Hs from the unselected library identified residue R28 in CDR1 to be a key contributor of resistance to acid aggregation. As the thermodynamic stability of the V_Hs remained unchanged with the mutations, it was concluded that R28 improved acid-induced aggregation resistance through stabilization of the denatured state.

A number of studies have gone on to show that the aggregation resistance of human domains can be effectively improved by incorporating charged amino acids – in particular negatively-charged residues – into hypervariable loops/CDRs [49, 64, 80-82]. By CDR loop swap and mutational analysis of aggregation-resistant and aggregation-prone human V_Hs, Perchiacca *et al.* [64] identified aggregation hotspots in, and in the vicinity of CDR1 and showed that negatively charged mutations at these spots transformed an aggregation-prone V_H into aggregation resistant ones. Human V_Hs DP47d and HEL4 differ only in their CDR/loop sequences (26-35, 50-56, 94-

102), yet DP47d is aggregation-prone whereas HEL4 is aggregation-resistant. Since the aggregation prone DP47d has a lower net charge than the aggregation-resistant HEL4 at pH 7, the authors asked if simply charging DP47d would transform it into an aggregation-resistant antibody. To this end, two different negatively- and positively-charged DP47d mutants were created by mutating solvent-exposed non-CDR loop residues L11D, P41E, G66E, S75D, and S112E (pI: 4.8) and E1R, L11K, P41D, D62K, and S112R (pI: 9.8), respectively. Although the charged mutations improved solubility to some extent with the negatively-charged mutations imparting a better effect, thermal unfolding and non-native protein solubility analysis (heat-induced aggregation analysis) revealed that similar to the parent DP47d, both acidic and basic mutants failed to unfold reversibly and also aggregated near and above their T_ms. Moreover, with respect to conformational stability, all aggregating and non-aggregating V_Hs had similar free energy of folding (ΔG_{N-U}). These results indicated that the different aggregation behavior between DP47d and HEL4 V_Hs could not be explained by a simple difference in net charge, but might be due to the difference in CDR sequence. To address whether the aggregation resistance of HEL4 was a result of a global effect of all 3 CDRs or is imparted by a single CDR, DP47d mutants were constructed where CDR1, CDR2 and CDR3 were replaced with those of HEL4. The three mutants were shown to be well folded and to have similar secondary structure. Interestingly, thermal unfolding and non-native protein solubility analysis revealed that HEL4 CDR2- or CDR3-swapped DP47d failed to reversibly unfold and did not show improved solubility, while the HEL4 CDR1-swapped mutant matched HEL4 in terms of reversibility of thermal unfolding and aggregation resistance. Moreover, DP47d and the CDR-swapped V_Hs were found to have similar free energy of folding (ΔG_{N-U}), indicating that the aggregation resistance of HEL4 CDR1-swapped DP47d was not due to increased conformational stability. The contribution of CDR1 to aggregation resistance of V_Hs has been demonstrated elsewhere [49, 65]. Through a systematic mutagenesis approach, Perchiacca et al [64] further showed that transferring individual residues from HEL4 CDR1/H1 to corresponding positions in DP47d CDR1/H1 failed to confer aggregation resistance to DP47d where six DP47d mutants constructed for this purpose (T28R, F29I, S31D, Y32E, A33D and S35G) had characteristics of aggregation prone V_Hs (e.g., irreversible unfolding). In contrast to the single mutants, a mutant of DP47d incorporating the HEL4 CDR1 negatively-charged Asp-Glu-Asp triad (S31D/Y32E/A33D) was aggregation-resistant and unfolded reversibly as was shown by thermal unfolding and non-native protein solubility experiments. Thus, as few as three negatively-charged mutations in CDR1 (residues 31-33) transformed the highly aggregating DP47d to a highly aggregation resistant HEL4-like V_H. Remarkably, it was shown that as little as one negative charge mutation in the structural region of H1 loop (F29D), resulted in DP47d being completely aggregation resistant, indistinguishable from HEL4. While human germ-line V_H3 domains typically

have a Phe at position 29, camelid V_HHs typically have a Tyr or Ser at the same position [74, 83, 84], suggesting the importance of position 29 in aggregation resistance and reversible unfolding of V_H domains. However, a study by Mandrup *et al.* [85] showed that the introduction of Asp at position 29 in HEL4 (I29D) caused a reduction in the thermodynamic stability. These findings suggest that the incorporation of acidic amino acids at positions 29, 31-33 into V_H library design may increase the aggregation resistance of the repertoire; however, incorporation of just one at position 29 may be preferable to avoid compromising repertoire diversity and hence the depth and breadth of the repertoires affinity/specificity.

Solvent-exposed hydrophobic residues in the CDRs of antibodies contribute to antigen binding, but they may also mediate their aggregation. In such cases it would be ideal to eliminate aggregation while leaving the hydrophobic residues intact to maintain the binding properties, e.g., affinity and specificity, uncompromised. In a series of experiments involving human V_Hs called "gammabodies" [81] with aggregation promoting hydrophobic residues in CDR3, Perchiacca *et al.* [86] showed that the positional-incorporation of negatively charged mutations in CDR3 can eliminate hydrophobic CDR3-mediated aggregation, yet maintaining the V_Hs affinity and specificity. Two aggregating human V_H domains, Aβ15-24 and Aβ33-42, were constructed with CDR3-inserted hydrophobic Aβ peptides determined to be the cause of aggregation. To address if the aggregation behavior of V_Hs could be improved by a simple net charge difference, Aβ15-24 and Aβ33-42 were engineered with negatively charged mutations at solvent-exposed positions in non-CDR loops (L11D, P41D, K43E, S74D, S112D), resulting in a decrease in pI. This approach did not improve expression yields or drastically alter secondary structures or conformational stability; however, relative improvements in solution properties of the charged mutants were observed. Thermal denaturation and non-native protein solubility analysis showed increased aggregation resistance at $T > 70°C$ and revealed a partial improvement in reversible thermal unfolding as a result of the charged mutations. Thus, it appears that introducing negatively-charged mutations outside CDR loops does not completely suppress aggregation of sdAbs with hydrophobic CDR loops. In sharp contrast, insertion of negatively-charged mutations near the extremities of the hydrophobic CDR3 did suppress the aggregation of Aβ15-24 and Aβ33-42. Six (single, double and triple) negatively-charged mutants of both Aβ15-24 and Aβ33-42 (D-Aβ15-24-D, DD-Aβ15-24-DD, and DED-Aβ15-24-DED; D-Aβ33-42-D, DD-Aβ33-42-DD, and DED-Aβ33-42-DED) with charged residues at both ends of the Aβ peptide segments in CDR3 were all shown to have a three-fold higher expression yield (16 – 20 mg/L) than the parent V_Hs Aβ15-24 and Aβ33-42. Solution properties of the mutants improved as a function of the number of negatively-charged amino acid insertions, with two or more insertions providing the maximal benefit. All four double and triple mutants were shown to be completely aggregation-resistant by non-native

protein solubility analysis at temperatures as high as 95°C, and fully refolded following thermal denaturation. Although the insertion of an Asp near the ends of CDR3 did not completely prevent aggregation, it had a better effect than the five non-CDR loop mutations in reducing the heat-induced aggregation of both Aβ15-24 and Aβ33-42. Single mutations improved native solubility (solubility at 25°C), but with less efficacy than the five non-CDR mutations. Two charge mutations resulted in a significant leap in solubility but it was only after the introduction of three negatively-charged mutations that both triple mutants remained fully soluble up to the seven days tested. Consistent with the solubility results, SEC profiles showed improved aggregation resistance as a function of an increase in the number of negatively-charged mutations. However, improved biophysical properties cannot be attributed to an increase in thermodynamic stability as all charged mutants and parental sdAbs had essentially the same conformational stability. Importantly, by comparing the binding affinities and specificities of Aβ15-24 and Aβ33-42 to those of corresponding single, double and triple negatively-charged mutants, it was demonstrated that the charged mutational approach endows aggregation resistance without compromising binding properties. This was also demonstrated for other aggregating V_Hs displaying hydrophobic Aβ peptides in their CDR3 [81].

More detailed studies subsequently showed a direct correlation between the solubilization effect of charged mutations and their closeness to shorter stretches of hydrophobic clusters within CDR3, which presumably mediated V_H aggregation. Four triple mutants of Aβ15-24 and Aβ33-42 with the Asp-Glu-Asp triad (DED) inserted at either end of their CDR3's were constructed (Aβ15-24-DED, DED-Aβ15-24; Aβ33-42-DED, and DED-Aβ33-42). For Aβ15-24, the inserted Asp-Glu-Asp triad would be in similar proximity from either point of insertion to its most hydrophobic segment which is near the center of the CDR3, whereas, for Aβ33-42, the CDR3 C-terminal insertion would make the Asp-Glu-Asp triad closer to the most hydrophobic segment which is near the C-terminus of the CDR3. Both Aβ15-24-DED and DED-Aβ15-24 had a similar expression yields and similar to that of the aggregation-resistant DED-Aβ15-24-DED (14 – 16 mg/L). For Aβ33-42, whereas Aβ33-42-DED had a similar expression yield to DED-Aβ33-42-DED (15 mg/L), DED-Aβ33-42 had a similar expression yield to the aggregation prone parental Aβ33-42. Consistent with these findings, native and non-native protein solubility, thermal unfolding and SEC analyses showed that both Aβ15-24-DED and DED-Aβ15-24 were aggregation resistant and only Aβ33-42-DED, with the charged triad closer to the most hydrophobic segment of CDR3, was non-aggregating. Although aggregating, DED-Aβ33-42 still outperformed the parental Aβ33-42, once again underlining the importance of negatively-charged mutation in improving aggregation resistance. Improved biophysical properties cannot be explained by conformational stability as all charged mutants and parental V_Hs had essentially the same secondary structure and

conformational stability. These results were further supported by more defined experiments performed with two more aggregating V_Hs, Aβ12-21 and Aβ18-27, which contain different ten-amino acid Aβ insertions into their CDR3, with the (same) most hydrophobic portion (VFFA) near the CDR3 C-terminus for Aβ12-21 and N-terminus for Aβ18-27, and their mutants with DED triad mutations inserted at either end (single DED triad) and at both ends (double DED triad) of CDR3 (Aβ12-21-DED, DED-Aβ12-21, DED-Aβ12-21-DED; Aβ18-27-DED, DED-Aβ18-27, and DED-Aβ18-27-DED. Native and non-native protein solubility and thermal unfolding experiments revealed that aggregation resistance was observed for Aβ12-21-DED and DED-Aβ12-21-DED, and for DED-Aβ18-27 and DED-Aβ18-27-DED, which was in agreement with the above conclusions that aggregation suppression is a function of the proximity of the charged insertions to the most hydrophobic portion of the CDRs. Nonetheless, for the aggregating DED-Aβ12-21 and Aβ18-27-DED V_Hs, significant improvement in aggregation resistance was still observed with the insertion of charged mutations. The expression yields of the non-aggregating V_Hs were three-fold higher (17 – 18 mg/L) than those of aggregating ones (6 – 7 mg/L) and the parent V_Hs (5 – 6 mg/L). Once again, secondary structures and thermodynamic stabilities of all the V_Hs were essentially the same and similar to the previous V_Hs above. The solubilization method appears to be generally applicable to aggregating V_Hs where antigen binding activity and aggregation-mediating hydrophobic patches all reside within one CDR. Applying the method where antigen binding and hydrophobic patches involve several CDRs is more challenging.

In related experiments, Perchiacca *et al.* [82] revealed that CDR3 N-terminal Asp containing Aβ18-27 V_Hs (D-Aβ18-27, DD-Ab18-27 and DDD-Aβ18-27) had improved biophysical properties proportional to the number of Asp insertions. Expression yields increased stepwise for each Asp insertion, from 21 mg/L for the aggregating Aβ18-27 to 45 mg/L for the triple mutant DDD-Aβ18-27. Reversible thermal unfolding and resistance to aggregation was partially increased with one Asp insertion but reached 100% with only two Asp insertions. Secondary structures and conformational stabilities remained the same; however, the insertion of three Asp residues at the C-terminus of CDR3 further away from the aggregation hotspot (Aβ18-27-DDD) failed to confer aggregation resistance to Aβ18-27. Contrary to the report showing that Asp had a higher solubilization activity than Glu [49], aggregation analyses of similar constructs with Glu instead of Asp insertions (E-Aβ18-27, EE-Aβ18-27 and EEE-Aβ18-27) showed that Asp and Glu had similar solubilization activity. Moreover, analysis of Ala insertion mutants (A-Aβ18-27, AA-Aβ18-27 and AAA-Aβ18-27) showed that the acquired aggregation resistance in charged mutants was not due to the CDR3 length increase. The authors extended their studies to mutants with positively-charged amino acid (Arg and Lys) instead of Asp (R-Aβ18-27, RR-Aβ18-27, RRR-Aβ18-27, K-Aβ18-27, KK-Aβ18-27, KKK-Aβ18-27). The secondary structure and conformational stability

remained the same, but interestingly, expression yields (9 – 19 mg/L) decreased as the number of insertions increased. This was not thought to be due to increased aggregation tendency or reduced conformational stability, but rather due to the poor capability of the positively-charged mutants to cross the cytoplasmic membrane that is required to fold in the periplasm, which was observed for other positively-charged proteins by Wilson and Finlay [87]. Aggregation resistance of the positively-charged mutants increased with an increase in the number of mutations. Partial solubilization was also observed with KKK or RRR at the C-terminus of Aβ18-27 CDR3; however, results with respect to the effect of insertion location of amino acids (KKK or RRR) on aggregation propensity of Aβ18-27 were inconsistent. A similar set of experiments with aggregating Aβ33-42, which differs from Aβ18-27 with respect to the inserted Aβ peptide and the location of its aggregation hotspot (CDR3 C-terminus) and its mutants with single, double and triple negatively (Asp or Glu)- or positively (Arg or Lys)-charged mutations inserted near CDR3 C-terminus revealed similar results to Aβ18-27. That is: i) secondary structure and conformational stability were similar for all, ii) in the case of Asp mutants, expression yields and aggregation resistance increased with increasing mutation numbers with complete aggregation resistance plateauing at two insertions, iii) in the case of Arg mutants, with increasing number of mutations, expression yields decreased whereas aggregation resistance increased but was partial, iv) Glu and Lys mutants showed similar solution behaviours to Asp and Arg mutants, respectively, v) the acquired solubilization was not due to a CDR3 length increase, and vi) charge insertions away from the CDR3 aggregation hotspot (CDR3 N-terminus) did not confer aggregation resistance.

Thus, inter-molecular charge-charge repulsion involving inserted flanking acidic residues at the edge of the hydrophobic CDR3 cannot explain their aggregation suppression activities, since positively charged insertions at the same insertion locations would have been expected to be equally effective in suppressing aggregation of Aβ18-27 or Aβ33-42 V_Hs. The next set of experiments revealed a more complex nature of the mechanism underlying charged mutation-mediated aggregation-suppression. Two sets of mutations, (set 1: L11R/G26R/S74R/E85R/A94R/Q105K/S112R; set 2: L11D/G26E/K43D/K64D/Q105E/S112D) were introduced at non-CDR solvent-exposed positions of the same V_H to obtain a positive and a negative scaffold, respectively. The hydrophobic Aβ33-42 CDR3 was inserted into both scaffolds each having either triple Asp, Arg or Ala (control) insertions at the CDR3 C-terminus, giving six V_Hs with similar secondary structure and conformational stability, although conformational stability was significantly less than Aβ33-42 with wild-type scaffold. Consistent with earlier findings, Asp insertion mutants (similar to Ala insertion mutants) gave higher expression yields than those of Arg insertion mutants (15 - 26 mg/L versus 10 mg/L) with better expression yields of Asp insertion mutants in a negative scaffold (26 mg/L versus 15 mg/L for a negative and a positive

scaffold, respectively). Subsequent biophysical analyses involving native and non-native protein solubility analysis, and SEC concluded that positively-charged CDR mutations had better solubility behavior in the context of a positive scaffold compared to the context of a negative scaffold, while negatively-charged ones in the context of positive scaffold were more aggregation prone than in the context of negative or almost neutral (wild-type) scaffolds. Importantly, binding measurements against Aβ42 amyloid fibrils of various triple charge mutants of Aβ18-27 and Aβ33-42 revealed that insertion of solubilizing negatively (Asp or Glu) or positively (Arg or Lys) charged mutations near the edges of hydrophobic CDR3 did not alter binding. Thus, charged mutations can be introduced into hydrophobic antibody loops to suppress aggregation without compromising the binding properties such as affinity and specificity. These studies also highlight the highly positional character of the charged mutational approach and the importance of identifying aggregation hotspots in CDRs for a more effective and efficient solubility engineering.

Single Asp mutants of a human V_H domain of V_H3 family (V3-23/DP47, J_H4b) were constructed by amino acid substitutions at several surface exposed positions which included H1, H2, and several non-CDR positions as reported by Dudgeon *et al.* [49]. Individual phage-displayed domains were assessed for aggregation resistance in phage ELISA by measuring binding to protein A (V_Hs) following transient heat denaturation. Measurements showed that Asp substitutions which improved aggregation resistance were all clustered in H1 (28, 30-33 and 35) with no contribution from other positions including H2. Double Asp mutant combinations in H1 (residues 31-33, 35) provided higher aggregation resistance. Further analysis of phage-V_Hs with single and multiple substitutions in the H1 loop showed that negatively-charged residues, more so Asp than Glu, were better than positively-charged Arg and Lys in improving the aggregation resistance of V_Hs, indicating that it is more than just charge that affects aggregation resistance and contrasting more recent findings [82]. Moreover, incorporating double Asp substitutions (H1: 32D/33D) into a synthetic V_H phage display library with extensive randomization at fourteen CDR positions significantly improved the aggregation resistance of the V_H repertoire, indicating a more general applicability of the approach. Consistent with the above findings, it was shown that the introduction of single, double and triple negatively-charged amino acids into the H1 of V_Hs (31-33) improved many of their biophysical properties, including i) increased soluble expression yields by ~seven-fold, ii) resistance to aggregation at $T > T_m$ (80°C) as determined by turbidometry, iii) resistance to freeze/drying or membrane filtration, iv) reduced stickiness to column matrix, and v) increased refolding efficiencies measured by SEC following exposure to denaturing heat (10 min at 80°C). Biophysical property improvements were shown to be a function of the number of mutations, where two or more mutations were needed to achieve maximal improvements. X-ray crystal structures of triple mutants of a DP47d 31-33DED V_H did not reveal any structural basis for

the observed biophysical behavior of the mutants as no significant conformational change (e.g., at CDRs or at the V_H/V_L interface) were observed as a result of the mutations. Additionally, the thermodynamic stability of several single and triple Asp/Glu V_H mutants tested either remained unchanged or decreased relative to wild-type DP47d domains, raising the possibility that the mutations improved aggregation resistance by stabilizing aggregation prone folding intermediates. It has to be noted that the approach by Dudgeon *et al.* [49] was different from similar approaches such as the "supercharged" approach by Lawrence *et al.* [88] that engineer proteins with a relatively a large number (30 – 80) of charged residues at non-specific surface positions and relied on the global net charge, or the "negatively charged peptide" approach by Schaefer and Plückthun [89] which was not positional but also relied on increased net charge. Altogether, the charge mutational strategy involving introducing negatively-charged amino acids into CDRs appear to be an effective and general strategy for imparting aggregation resistance to human V_Hs. The amino acid substitution is positional and a threshold number of negatively-charged amino acid insertions are required to achieve complete aggregation resistance. A lack of a direct correlation between V_H aggregation resistance and thermodynamic stabilities indicates the charged mutations impart aggregation resistance by reducing the tendency of the unfolded/partially unfolded state to aggregate. Strangely, positively-charged amino acid insertions provide partial aggregation resistance at best. The differential solubilization activity of Asp/Glu and Arg/Lys has been attributed to favorable or non-favorable interactions with water and increased [82].

Chen *et al.* [90] generated a human V_H phage display library using a V_H scaffold (named m0) which was isolated from a hcAb derived from a Fab library. The m0 V_H was randomly mutated at four positions in CDR1 (26, 28, 30 and 31) with four amino acid residues (Ala, Asp, Ser or Tyr), while CDR2 and CDR3 were grafted from five human Fab libraries through a PCR-mediated amplification of CDR2 and CDR3. From the sequence diversity analysis, it was found that the mutations in CDR1 were not biased to a particular residue in all positions although relatively higher frequencies of Asp and Tyr were observed for all positions compared to those of Ala and Ser. After four rounds of selection for binding to protein A, the comparative frequency analysis of clones before and after the selection demonstrated that Asp and Ser were significantly enriched at CDR1 positions 26, 28 and 31, and CDR1 position 30, respectively, after selection (two fold increase on average), while the occurrence of Tyr was more than two-fold decreased in all positions compared to the frequency before selection. This indicates that Asp and Ser residues in those CDR1 positions could stabilize V_Hs, which is in agreement with previous findings [47, 49, 62-64]. No sequence or length bias was found in CDR3 of the clones analyzed.

Inspired by the work on V_HHs [67, 91-93], Kim *et al.* [35, 94] employed a disulfide linkage engineering approach to stabilize human V_Hs. To this end, model human V_Hs (V_H3 family) with

cysteine substitutions at positions 49 and 69 in FR2 and FR3 were generated and shown to be expressed well in soluble form in *E. coli* with expression yields comparable to wild-type V_Hs. Importantly, all formed the intended 49C-69C extra disulfide linkage as shown by mass spectrometry fingerprinting analysis. As observed for llama V_HHs [67, 91-93, 95], the addition of the 49C-69C disulfide linkage dramatically increased the T_m of all V_Hs, by 14°C - 18°C, with two mutant V_Hs exhibiting exceptionally high T_ms (89.2°C and 87.5°C), likely due to the presence of another extra, inter-CDR1-CDR3 disulfide linkage, which was previously demonstrated to stabilize human V_Hs [14, 63]. The high thermal stability of the two mutant V_Hs in turn imparted unusually high resistance to GI proteases, which were previously shown to be correlated [67, 96]. Importantly, the mutant V_Hs showed that the extra disulfide linkage also improved aggregation resistance. For example, the mutant V_Hs showed better SEC behavior such as lower high-molecular weight aggregates, higher elution recoveries and symmetric, monomeric peaks with no tailing, or a mutant V_H, in particular, showed a two-fold increase in thermal reversibility compared to its wild-type counterpart. However, SPR binding analysis of mutant V_Hs revealed affinity decreases of two-to-ten-fold for protein A, indicating structural alteration as a result of the 49C-69C disulfide linkage introduction, further suggesting that this may lead to affinity and specificity compromises as has been shown to be the case for a number of V_HHs with the same 49C-69C disulfide linkage [67, 91, 93]. Thus, selection for affinity should be post 49C-69C disulfide linkage introduction.

While the focus of this review is on human domain antibodies, there are a number of studies describing the isolation and engineering of soluble, non-aggregating murine V_H and V_L sdAbs that could be applied to human sdAbs. For example, a consensus sequence-based mutagenesis approach, founded on the hypothesis that the replacement of native amino acid residues with canonical or consensus residues lead to protein stabilization [97], was used as an alternative approach to improve stability and aggregation resistance. Wirtz and Steipe [98] successfully predicted such stabilizing mutations, and when engineered onto a murine V_H (IcaH-01), stability and aggregation resistance were increased. Specifically, IcaH-01 variants with a single mutation (A16G, R43Q, I58T or P74S) showed small but significant increases in midpoint denaturation concentrations in Gdn-HCl unfolding experiments. When the beneficial mutations were combined, the stabilizing effects were additive. For example, IcaH-501 (P7S/A16G/R43Q/I58T/P74S) was twenty-fold more stable in heat denaturation experiments than the parent IcaH-01. Elsewhere, Reiter *et al.* [99] searched a murine IgG database for rare V_H residues located at the V_H/V_L interface and produced a V_H phage display library based on a scaffold that contained a Lys at position 44 (normally G44). The group randomized CDR3 at nine positions, but did not mutate CDR1 or CDR2. After four rounds of panning against two test antigens, several

clones were positive by phage ELISA. One clone was expressed and found to be non-aggregating by SEC and analytical ultracentrifugation. A soluble ELISA clearly showed the isolated domain was not sticky to other non-target proteins and SPR showed the clone possessed an affinity of K_D = 20 nM. In a third example, camelization of an aggregating murine V_H (1F7) failed to solubilize the V_H [100]. By comparing molecular dynamic simulations of 1F7 versus a llama V_HH, a mobile region in the V_H was unidentified. This region corresponded to residues 9 – 16 and was thought to be the cause of aggregation of the V_H. Interestingly, residues in this region are involved in hydrophobic-hydrophobic interactions with C_H1, which is thought to be mobile in the stand alone V_H due to the loss of the interactions. Subsequently, the V_H residues in the mobile region were replaced with corresponding ones from the V_HH without making the region more hydrophilic. Subsequent molecular dynamic simulations suggested increased stability of the region as a result of the "mutation". It was not shown by biochemical studies if the mutations could indeed stabilize the V_H. The molecular dynamic simulation approach may be useful in identifying aggregation causing regions that are the result of dynamic effects.

3.2. Mutations affecting biophysical properties of human V_Ls

V_L domains are similar to V_H domains in terms of overall tertiary structure in that two β-sheets, formed by multiple anti-parallel β-strands, are packed face to face to form a β-sandwich structure [36]. Multiple studies have pointed out that human V_Ls have lower aggregation propensity than human V_Hs, indicating that human V_Ls may be better candidates for human therapeutics [1, 32, 36, 51, 101-108]. However, compared to human V_Hs, there are fewer studies dedicated to characterizing and improving the stability and aggregation resistance of human V_L domains through mutagenesis. Historically, most human V_L domain efforts were focused on V_L domain-mediated amyloidosis (light-chain amyloidosis or AL), which is characterized by extracellular deposition of fibrillar aggregates leading to malfunction of multiple organs in human [109], and this led to additional studies that characterized mutants of these V_Ls [110-112]. Recently, human V_Ls have been engineered and stabilized by Dudgeon *et al.* [49] and Kim *et al.* [36].

Using the sequence analysis of the AL-related V_Ls, Hurle *et al.* [109] identified mutations at particular positions that might be important in maintaining stability and structural intactness of the V_L domain (Table 2 and Supplementary Table 1). They first expressed and purified the wild-type domain of the Bence-Jones protein REI V_L in *E. coli* and determined minimal aggregates were formed. Then they engineered individual V_Ls with one of six point mutations (Y49F, G57E, R61N, G68D, L78T or A84T) and successfully expressed all in *E. coli*, although some formed inclusion bodies. In five of the six mutants, the amino acid replacements destabilized the V_Ls resulting in

increased aggregate formation (weakened aggregation resistance) as determined by increased Congo red binding to aggregates. Only the L78T mutation showed low Congo red binding, identical to the wild-type V_L. Unfolding experiments using Gdn-HCl demonstrated five of the six mutants were less stable than wild-type V_Ls, while the Y49F mutant was essentially unchanged relative to wild-type. When the data were combined by plotting *in vitro* aggregate formation versus Gdn-HCl denaturation midpoints, all mutants were considered less stable than the wild-type V_L. The R61N mutation was the most destabilizing because of the loss of a conserved salt bridge/H-bond with D82. Overall, this supported the hypothesis that specific mutations on the AL-related V_L domains are crucial to facilitate the destabilization of the V_L domains allowing the V_L domains to adopt aggregation prone non-native folding state.

To address the thermodynamic contribution of the canonical disulfide linkage of the V_L domain, Frisch *et al.* [110] engineered the REI V_L domain with multiple stabilizing mutations. The mutations were first selected through sequence alignments against the kappa light chain sequence database and then characterized in two different states (oxidized (SS) or reduced ($(SH)_2$), revealing that either single mutations or combinations of two or three mutations increased thermodynamic stability and that the mutational effect was additive. The thermodynamic contribution of the disulfide linkage was 19 kJ/mol on average. In particular, the single mutation T39K increased the thermodynamic stability of wild-type REI V_L from 25 kJ/mol to 30 kJ/mol, and other single mutations Y32H, F73L and Y71F increased wild-type V_L stability by similar amounts. When combinations of single mutations were introduced, even greater thermodynamic stability was achieved, with the T39K/Y71F double mutant reaching 36 kJ/mol. When the F73L mutation was added, giving T39K/Y71F/F73L, the triple mutant reached 40.6 kJ/mol. Importantly, when the Y32H mutation was introduced into the non-expressing C23V mutant, the resulting V_L mutant (C23V/Y32H) expressed in soluble form in *E. coli*, which otherwise would have been insoluble inclusion bodies. It was demonstrated that the V_L mutants had similar tertiary structures with and without the canonical disulfide linkage. Thus, it was concluded that the canonical disulfide linkage was not necessarily required for proper folding of the V_L domain as long as the loss of the disulfide linkage could be compensated for by specific amino acid replacements stabilizing the tertiary fold (Table 2 and Supplementary Table 1).

To gain structural insight of light-chain fibril formation (AL), Raffen *et al.* [111] used a site-directed mutagenesis approach to study the mutational effects on the stability and the fibril formation of the three amyloidogenic V_L domains (V_L domain of human κ 4 protein Len (VL-Len), VL-Rec and VL-Sma). In total, twenty V_L mutants were produced, which led to both stabilized and destabilized V_Ls compared to wild-type. Specifically, using Gdn-HCl midpoint of unfolding concentrations (C_m) as a comparison of V_L stability, VL-Len and VL-Rec possessed a C_m of 1.76 M

and 0.76 M, respectively. VL-Rec differed from VL-Len at fourteen positions, of which eight were located in CDR1. When the CDR1 of VL-Len was replaced with the CDR1 of VL-Rec, giving $LenCDR1_{REC}$, the C_m was between that of wild-type VL-Len and VL-Rec. Introduction of other VL-Rec containing mutants into VL-Len included a single Y27dD mutation, or doublet or triplet mutants containing Y27dD, that lead to a considerable increase in stability. Other single point mutations in VL-Len generally reduced V_L domain stability, including two CDR1 mutations (N28F and V27bL). Overall, it was found that there was a good correlation between the stability and the fibril formation, where only destabilizing mutations induced the formation of fibril aggregates although there were some destabilizing mutations that did not induce the fibril formation, further confirming the study by Hurle *et al.* [109]. Furthermore, the authors noted the V_L domains were susceptible to global stability changes imparted by CDR residue mutations.

Using the V_L domain of human κ 4 V_L Len (VL-Len), Pokkuluri *et al.* [112] demonstrated that the Gln to Asp mutations at positions 38 (VL-Len Q38D) and 89 (VL-Len Q89D), which are located at the end and in the middle of a β-strand, decreased the thermodynamic stability of the wild-type VL-Len domain. The C_m values of VL-Len Q38D and VL-Len Q89D, 1.30 M and 0.55 M, revealed that the mutational effect of Q89D was more dramatic than Q38D in destabilizing the wild-type V_L-Len (C_m value of 1.76 M). Interestingly, the destabilizing effects of the Gln to Asn substitution at the same positions (Q38N and Q89N) were milder compared to the mutational effects observed from both Q38D and Q89D with C_m values of 1.51 M and 1.39 M for VL-Len Q38N and VL-Len Q89N, respectively. A crystallographic study did not show a structural basis for the instability given by the mutations since there was no major structural perturbation in the crystal structures of the mutants. Based on the observation that the negatively charged Asp residues were more destabilizing than the neutral Asn residues, which provide similar steric constraints as the Asp residues but do not carry a negative charge, therefore, an unfavorable charge-charge interaction seemed to explain the instability. Although several other factors such as changes in dihedral angles near positions 38 and 89, rotomer conformations of the Asp residues and contact distances between atoms cannot be excluded. Glu residues also carry a negative charge but it has a longer side chain than Asp residues, so an unfavorable local repulsion with the negative charges from the carbonyl oxygen groups of the backbone can be avoided. Indeed, the destabilizing effect of the Gln to Glu substitution at the position 89 (Q89E) was minor compared to the Q89D mutation (ΔG_{unf} for VL-Len Q89E = 6.3 kcal mol^{-1}) [112].

In contrast to the results from the mutagenesis approach with human V_Hs by Dudgeon *et al.* [49] (see section 3.1), the introduction of mutations at equivalent positions in CDR1 of a V_k1 family human V_L (O2/O12/DPK9, J_k1) did not impart the biophysical properties to the extent shown in the human V_Hs. Instead, negatively charged mutations at five positions in CDR2 (50, 51,

52, 53 and 56) and two positions in non-CDR (24 and 49) clustered in the CDR2 of the human V_L (CDR-L2) were identified to play a major role in aggregation resistance. The mutational effect was found to be a function of the mutation numbers, in that the combination of the charged mutations indeed improved biophysical properties (e.g., increased soluble expression, decreased SEC elution volume, increased thermal refolding and decreased turbidity upon freeze dry). The thermodynamic stability of several single and triple Asp/Glu V_L mutants either remained unchanged or decreased compared to wild-type DPK9 domain, indicating that the mutations improved aggregation resistance by stabilizing aggregation prone folding intermediates, rather than stabilizing the folded state. In agreement with the human V_H mutagenesis study by Dudgeon *et al.* [49] (see section 3.1), the incorporation of double Asp substitutions (L2: 52D/53D) into a synthetic V_L phage display library that was extensively randomized at nine CDR positions significantly improved the aggregation resistance of the human V_L phage display repertoire, indicating a more general applicability of the approach for human V_L domains. The aggregation resistance observed from the phage display repertoires was largely independent to CDR sequences at other positions and was highly positional, such that the effect differed depending on the positions in CDR-L2 (e.g., Q55D showed much less binding to protein L (< 20%) compared to the other positions in CDR-L2). A crystal structure of human V_L triple mutant (DPK9 50,52-53DDD) revealed that there was no major structural perturbation in either the V_H/V_L interface or the CDR loops, while structural changes were restricted to mutant side chains, suggesting that the mutations may be compatible with V_H/V_L pairing and antigen binding when combined into existing mAbs and antibody fragments (e.g., scFv). It was demonstrated that variants of Trastuzumab (Herceptin®) carrying single and double Asp substitutions were shown to be fully functional in cellular binding, inhibitory profiles, serum clearance in an IgG format and showed enhanced aggregation resistance and unchanged binding affinity (3.8 nM and 4.1 nM for WT and 30D/52D, respectively) as an scFv.

Recently, an approach for stabilizing human V_Ls with an extra disulfide linkage was described [36], similar to that applied to human V_Hs [35, 94]. Kim *et al.* [36] engineered eight human V_Ls from different germ-line origins (2 V_κ1s, 5 V_κ3s, 1 V_λ1) with two extra cysteine residues at positions FR2 48 and FR3 64 and compared the mutants with their corresponding wild-type V_Ls in terms of biophysical properties. The mutants formed the intended disulfide linkage and were expressed in soluble form in *E. coli* with yields ranging from 1.1 – 11 mg/L, which was significantly lower than the corresponding wild-type V_Ls at 6.2 – 77 mg/L. With the exception of one, all the mutants were observed to be monomeric in SEC-MALS, indicating that similar to their wild-type counterparts, the mutant V_Ls were aggregation resistant. As observed for human V_Hs and llama V_HHs with the extra disulfide linkage [35, 67], the mutant V_Ls showed significantly

increased thermal stability by 12.4°C (median ΔT_m) with an increase in the range of 5.4°C - 17.3°C. Although the study demonstrated that the thermal stability improvement occurred regardless of the germ-line origins, the improvement was more prominent for the $V_\kappa 3$ and $V_\lambda 1$ subgroups. Experiments revealed that the extra disulfide linkage significantly improved resistance to pepsin, but did not significantly alter trypsin or chymotrypsin resistance. The improvement in pepsin resistance strongly correlated with increase in thermal stability. The higher resistance to pepsin was thought to be due to the fact that the engineered disulfide linkage changed the conformation of the mutant V_Ls to a more compact and thermodynamically stable structure that resist acid-induced unfolding under the pepsin digestion conditions. Differential SPR binding against generic ligand protein L and protease resistance profiles between of wild-type and mutant V_Ls indicated conformational changes as a result of the disulfide linkage introduction, which may undesirably lead to affinity and specificity compromises. Overall, the disulfide linkage engineering approach appeared to be a simple and universally applicable strategy to increase V_L stability without compromising aggregation resistance.

4. Strategies for obtaining functional human sdAb reagents

While the specific mutations described above lead to the improvement of human sdAb biophysical properties, they are just one of several strategies that can be utilized to obtain robust binders (Table 1). In Figure 2 we propose a practical, chronological workflow for selecting human sdAb affinity reagents.

The starting point for most is selecting a non-engineered scaffold in which to build a library of human V_Hs or V_Ls. A major consideration is whether the library will be human V_H or V_L, and from what human germ-line family. As highlighted above, germ-line families have different inherent biophysical properties and selecting a family with a greater tendency for aggregation will inherently require more modification and optimization. When selecting a lead scaffold for library construction, many engineering and/or rational design approaches can be applied. These include, selecting V_Hs/V_Ls with optimal isoelectric points, introducing an additional stabilizing disulfide bond, selecting V_Hs/V_Ls with protein A or protein L binding, camelization (applying V_HH residues at equivalent positions of V_Hs/V_Ls (positions 37, 44, 45 and 47)), CDR modifications including adding charged tags on either end of the CDR, hydrophobicity or CDR loops, and *in silico* algorithms for determining aggregation hotspots. Another approach is to introduce the CDRs of various V_Hs/V_Ls into a lead scaffold to ensure the scaffold's biophysical properties are maintained in the presence of diverse CDR sequences.

Upon selecting a lead scaffold, CDRs can be randomized and the library constructed. In randomizing the CDRs of the optimized scaffold, CDR length – especially CDR3 – can be altered.

In addition, certain CDR residues can be maintained, or randomized in a restricted fashion. The sequence diversity of human sdAb libraries should be as large as possible (approaching 1×10^{10}) to ensure the isolation of high affinity binders and the choice of display format should also be considered (i.e., yeast-display, phage-display, ribosome-display) [113]. Depending on the display format, selecting a host cell line is also important. After production of the library, either synthetically or through mutagenesis approaches, affinity selection of binders to a desired target can be commenced. A number of affinity selection strategies can be employed to reduce the number of aggregation prone binders and select for soluble, stable human sdAbs. These include the use of deep sequencing of eluted pools of binders, tailored elution approaches (competition, pH, denaturants, proteases), the pre-treatment of libraries with heat, pH, proteases or denaturants, utilizing solid-phase or solution-phase immobilized targets, or intracellular antibody capture (IAC) technology [114].

Finally, the biophysical properties of human sdAb binders can be improved or modified after selection using many of the same methods applied at the scaffold selection stage, including disulfide bond engineering, addition of charged fusion tags and mutating protease-sensitive residues. Furthermore, human sdAbs require an engineering for their specific application, including half-life extension (i.e., fusion to Fc or an anti-serum albumin sdAb), glycosylation to increase their stability, and labelling strategy (i.e., for antibody-drug conjugates or for attaching fluorophores for imaging/diagnostics).

5. CONCLUSIONS

The demands of therapeutic antibodies for treating human malignancies including cancers are growing exponentially. Although a great deal of understanding has been achieved in many aspects of human sdAb engineering, particularly in our understanding of thermodynamic folding stability due to powerful approaches performed on *in vitro* phage display method with structure-based mutagenesis and high-throughput selection, there are still areas that deserve attention. This includes understanding aggregation or colloidal stability, and in doing so will allow developing more robust human therapeutics. Combining *in silico* engineering and high-throughput approaches, systematic mutational approaches will be crucial to gain a deeper understanding of human sdAbs in terms of stability and aggregation, which will surely be important for further advancement in generating a wide range of human antibodies that facilitate broader antibody applications.

There have been several strategies for generating non-aggregating human sdAbs. Some have relied on mutations in CDRs and some in FRs. Conformational changes in response to such mutations have been diverse, highlighting structural plasticity of human sdAbs in accommodating

aggregation resistance. None of the solutions devised so far have been universal, although some maybe more general than others. When the solubilization approach is applied to an individual binder, if successful, there is always the risk of compromising affinity and specificity, akin to antibody humanization. When applied en masse to human sdAbs phage display libraries, there is the risk of populating the library with aggregating species, which obscures the selection of non-aggregating species. This is where combining stability selection with affinity selection becomes crucial for efficient isolation of aggregation-resistant binders. The heat-denaturation method developed by Jespers *et al.* [47] appears to be an effective one for this purpose. Engineering for aggregation resistance has often come at the cost of reduced thermodynamic stability. However, the acid-denaturation selection method by Famm *et al.* [79] seems to be an avenue for generating aggregation-resistant human sdAbs with high thermodynamic stability, as is the disulfide linkage engineering approach by Kim *et al.* [35, 94].

The whole purpose of resorting to human sdAbs as opposed to other non-human sdAbs is the low immunogenicity in immunotherapy applications. Thus, to be true to the cause, one should try to minimize the deviations from human sequences and ideally utilize fully human FRs in order to minimize immunogenicity. Therefore, the CDR mutational approach should be considered the preferred strategy for obtaining aggregation-resistant human sdAbs. As such, one would construct phage display libraries by surmounting CDR repertoires on a fully human scaffold, followed by subjecting the library to affinity-stability selection in order to isolate non-aggregating binders. However, it remains to be seen if mutations in the CDRs alone can lead to binders that are both non-aggregating and of high affinity, similar to V_HHs.

REFERENCES

1. **Holt, L.J., et al., *Domain antibodies: proteins for therapy.* Trends in biotechnology, 2003. 21(11): p. 484-90.**
2. **Holliger, P. and P.J. Hudson, *Engineered antibody fragments and the rise of single domains.* Nature biotechnology, 2005. 23(9): p. 1126-36.**
3. **Weisser, N.E. and J.C. Hall, *Applications of single-chain variable fragment antibodies in therapeutics and diagnostics.* Biotechnology advances, 2009. 27(4): p. 502-20.**
4. **Muyldermans, S., *Nanobodies: natural single-domain antibodies.* Annual review of biochemistry, 2013. 82: p. 775-97.**
5. **Ward, E.S., et al., *Binding activities of a repertoire of single immunoglobulin variable domains secreted from Escherichia coli.* Nature, 1989. 341(6242): p. 544-6.**
6. **Hamers-Casterman, C., et al., *Naturally occurring antibodies devoid of light chains.* Nature, 1993. 363(6428): p. 446-8.**
7. **Arbabi Ghahroudi, M., et al., *Selection and identification of single domain antibody fragments from camel heavy-chain antibodies.* FEBS letters, 1997. 414(3): p. 521-6.**
8. **Spinelli, S., et al., *The crystal structure of a llama heavy chain variable domain.* Nature structural biology, 1996. 3(9): p. 752-7.**

9. Sheriff, S. and K.L. Constantine, *Redefining the minimal antigen-binding fragment.* Nature structural biology, 1996. 3(9): p. 733-6.
10. Desmyter, A., et al., *Crystal structure of a camel single-domain VH antibody fragment in complex with lysozyme.* Nature structural biology, 1996. 3(9): p. 803-11.
11. Muyldermans, S., et al., *Sequence and structure of VH domain from naturally occurring camel heavy chain immunoglobulins lacking light chains.* Protein engineering, 1994. 7(9): p. 1129-35.
12. Davies, J. and L. Riechmann, *'Camelising' human antibody fragments: NMR studies on VH domains.* FEBS letters, 1994. 339(3): p. 285-90.
13. Davies, J. and L. Riechmann, *Antibody VH domains as small recognition units.* Bio/technology, 1995. 13(5): p. 475-9.
14. Davies, J. and L. Riechmann, *Single antibody domains as small recognition units: design and in vitro antigen selection of camelized, human VH domains with improved protein stability.* Protein engineering, 1996. 9(6): p. 531-7.
15. Riechmann, L., *Rearrangement of the former VL interface in the solution structure of a camelised, single antibody VH domain.* Journal of molecular biology, 1996. 259(5): p. 957-69.
16. Tanha, J., et al., *Optimal design features of camelized human single-domain antibody libraries.* J Biol Chem, 2001. 276(27): p. 24774-80.
17. Martin, F., et al., *Affinity selection of a camelized V(H) domain antibody inhibitor of hepatitis C virus NS3 protease.* Protein engineering, 1997. 10(5): p. 607-14.
18. Bond, C.J., J.C. Marsters, and S.S. Sidhu, *Contributions of CDR3 to V H H domain stability and the design of monobody scaffolds for naive antibody libraries.* Journal of molecular biology, 2003. 332(3): p. 643-55.
19. Igawa, T., et al., *Engineering the variable region of therapeutic IgG antibodies.* MAbs, 2011. 3(3): p. 243-52.
20. Perchiacca, J.M. and P.M. Tessier, *Engineering aggregation-resistant antibodies.* Annual review of chemical and biomolecular engineering, 2012. 3: p. 263-86.
21. Rouet, R., D. Lowe, and D. Christ, *Stability engineering of the human antibody repertoire.* FEBS letters, 2014. 588(2): p. 269-77.
22. Lee, C.C., J.M. Perchiacca, and P.M. Tessier, *Toward aggregation-resistant antibodies by design.* Trends in biotechnology, 2013. 31(11): p. 612-20.
23. Hermeling, S., et al., *Structure-immunogenicity relationships of therapeutic proteins.* Pharmaceutical research, 2004. 21(6): p. 897-903.
24. Schellekens, H., *Factors influencing the immunogenicity of therapeutic proteins.* Nephrology, dialysis, transplantation : official publication of the European Dialysis and Transplant Association - European Renal Association, 2005. 20 Suppl 6: p. vi3-9.
25. Singh, S.K., *Impact of product-related factors on immunogenicity of biotherapeutics.* Journal of pharmaceutical sciences, 2011. 100(2): p. 354-87.
26. Jefferis, R., *Aggregation, immune complexes and immunogenicity.* MAbs, 2011. 3(6): p. 503-4.
27. Wang, W., et al., *Antibody structure, instability, and formulation.* Journal of pharmaceutical sciences, 2007. 96(1): p. 1-26.
28. Neal, B.L., D. Asthagiri, and A.M. Lenhoff, *Molecular origins of osmotic second virial coefficients of proteins.* Biophysical journal, 1998. 75(5): p. 2469-77.
29. Chang, R.C., D. Asthagiri, and A.M. Lenhoff, *Measured and calculated effects of mutations in bacteriophage T4 lysozyme on interactions in solution.* Proteins, 2000. 41(1): p. 123-32.
30. Papandreou, N., et al., *Universal positions in globular proteins.* European journal of biochemistry / FEBS, 2004. 271(23-24): p. 4762-8.

31. Knappik, A., et al., *Fully synthetic human combinatorial antibody libraries (HuCAL) based on modular consensus frameworks and CDRs randomized with trinucleotides.* Journal of molecular biology, 2000. 296(1): p. 57-86.
32. Ewert, S., et al., *Biophysical properties of human antibody variable domains.* Journal of molecular biology, 2003. 325(3): p. 531-53.
33. Riechmann, L. and J. Davies, *Backbone assignment, secondary structure and protein A binding of an isolated, human antibody VH domain.* Journal of biomolecular NMR, 1995. 6(2): p. 141-52.
34. Akerstrom, B., et al., *On the interaction between single chain Fv antibodies and bacterial immunoglobulin-binding proteins.* Journal of immunological methods, 1994. 177(1-2): p. 151-63.
35. Kim, D.Y., et al., *Disulfide linkage engineering for improving biophysical properties of human VH domains.* Protein Eng Des Sel, 2012. 25(10): p. 581-9.
36. Kim, D.Y., et al., *Antibody light chain variable domains and their biophysically improved versions for human immunotherapy.* MAbs, 2014. 6(1): p. 219-35.
37. Sasso, E.H., G.J. Silverman, and M. Mannik, *Human IgM molecules that bind staphylococcal protein A contain VHIII H chains.* Journal of immunology, 1989. 142(8): p. 2778-83.
38. Sasso, E.H., G.J. Silverman, and M. Mannik, *Human IgA and IgG F(ab')2 that bind to staphylococcal protein A belong to the VHIII subgroup.* Journal of immunology, 1991. 147(6): p. 1877-83.
39. Hillson, J.L., et al., *The structural basis of germline-encoded VH3 immunoglobulin binding to staphylococcal protein A.* The Journal of experimental medicine, 1993. 178(1): p. 331-6.
40. Sasano, M., D.R. Burton, and G.J. Silverman, *Molecular selection of human antibodies with an unconventional bacterial B cell antigen.* Journal of immunology, 1993. 151(10): p. 5822-39.
41. Silverman, G.J., M. Sasano, and S.B. Wormsley, *Age-associated changes in binding of human B lymphocytes to a VH3-restricted unconventional bacterial antigen.* Journal of immunology, 1993. 151(10): p. 5840-55.
42. Bjorck, L., *Protein L. A novel bacterial cell wall protein with affinity for Ig L chains.* Journal of immunology, 1988. 140(4): p. 1194-7.
43. Nilson, B.H., et al., *Protein L from Peptostreptococcus magnus binds to the kappa light chain variable domain.* The Journal of biological chemistry, 1992. 267(4): p. 2234-9.
44. Beckingham, J.A., et al., *Interactions between a single immunoglobulin-binding domain of protein L from Peptostreptococcus magnus and a human kappa light chain.* The Biochemical journal, 1999. 340 (Pt 1): p. 193-9.
45. Axcrona, K., L. Bjorck, and T. Leanderson, *Multiple ligand interactions for bacterial immunoglobulin-binding proteins on human and murine cells of the hematopoetic lineage.* Scandinavian journal of immunology, 1995. 42(3): p. 359-67.
46. Cossins, A.J., et al., *Recombinant production of a VL single domain antibody in Escherichia coli and analysis of its interaction with peptostreptococcal protein L.* Protein expression and purification, 2007. 51(2): p. 253-9.
47. Jespers, L., et al., *Aggregation-resistant domain antibodies selected on phage by heat denaturation.* Nature biotechnology, 2004. 22(9): p. 1161-5.
48. To, R., et al., *Isolation of monomeric human V(H)s by a phage selection.* J Biol Chem, 2005. 280(50): p. 41395-403.
49. Dudgeon, K., et al., *General strategy for the generation of human antibody variable domains with increased aggregation resistance.* Proceedings of the National Academy of Sciences of the United States of America, 2012. 109(27): p. 10879-84.
50. Ma, X., et al., *Design of synthetic autonomous VH domain libraries and structural analysis of a VH domain bound to vascular endothelial growth factor.* Journal of molecular biology, 2013. 425(12): p. 2247-59.

51. Hussack, G., et al., *A V(L) single-domain antibody library shows a high-propensity to yield non-aggregating binders.* Protein engineering, design & selection : PEDS, 2012. 25(6): p. 313-8.
52. Lowe, D., et al., *Aggregation, stability, and formulation of human antibody therapeutics.* Advances in protein chemistry and structural biology, 2011. 84: p. 41-61.
53. Jelesarov, I. and H.R. Bosshard, *Isothermal titration calorimetry and differential scanning calorimetry as complementary tools to investigate the energetics of biomolecular recognition.* Journal of molecular recognition : JMR, 1999. 12(1): p. 3-18.
54. Kelly, S.M. and N.C. Price, *The use of circular dichroism in the investigation of protein structure and function.* Current protein & peptide science, 2000. 1(4): p. 349-84.
55. Griffiths, P. and J.A. De Haseth, *Fourier transform infrared spectrometry*. Vol. 171. 2007: John Wiley & Sons.
56. Senisterra, G.A. and P.J. Finerty, Jr., *High throughput methods of assessing protein stability and aggregation.* Molecular bioSystems, 2009. 5(3): p. 217-23.
57. Wang, Y., et al., *Quantitative Evaluation of Colloidal Stability of Antibody Solutions using PEG-Induced Liquid-Liquid Phase Separation.* Molecular pharmaceutics, 2014.
58. Li, Y., W.F.t. Weiss, and C.J. Roberts, *Characterization of high-molecular-weight nonnative aggregates and aggregation kinetics by size exclusion chromatography with inline multi-angle laser light scattering.* Journal of pharmaceutical sciences, 2009. 98(11): p. 3997-4016.
59. Kabat, E., et al., *Sequences of Proteins of Immunological Interest . NIH Publication No. 91-3242. US Department of Health and Human Services.* Public Health Service, National Institutes of Health, Bethesda, MD, 1991.
60. Dottorini, T., et al., *Crystal structure of a human VH: requirements for maintaining a monomeric fragment.* Biochemistry, 2004. 43(3): p. 622-8.
61. Davies, J. and L. Riechmann, *Affinity improvement of single antibody VH domains: residues in all three hypervariable regions affect antigen binding.* Immunotechnology : an international journal of immunological engineering, 1996. 2(3): p. 169-79.
62. Jespers, L., et al., *Crystal structure of HEL4, a soluble, refoldable human V(H) single domain with a germ-line scaffold.* Journal of molecular biology, 2004. 337(4): p. 893-903.
63. Arbabi-Ghahroudi, M., et al., *Aggregation-resistant VHs selected by in vitro evolution tend to have disulfide-bonded loops and acidic isoelectric points.* Protein Eng Des Sel, 2009. 22(2): p. 59-66.
64. Perchiacca, J.M., M. Bhattacharya, and P.M. Tessier, *Mutational analysis of domain antibodies reveals aggregation hotspots within and near the complementarity determining regions.* Proteins, 2011. 79(9): p. 2637-47.
65. Dudgeon, K., K. Famm, and D. Christ, *Sequence determinants of protein aggregation in human VH domains.* Protein engineering, design & selection : PEDS, 2009. 22(3): p. 217-20.
66. Barthelemy, P.A., et al., *Comprehensive analysis of the factors contributing to the stability and solubility of autonomous human VH domains.* The Journal of biological chemistry, 2008. 283(6): p. 3639-54.
67. Hussack, G., et al., *Engineered single-domain antibodies with high protease resistance and thermal stability.* PLoS One, 2011. 6(11): p. e28218.
68. Bond, C.J., et al., *A structure-based database of antibody variable domain diversity.* Journal of molecular biology, 2005. 348(3): p. 699-709.
69. de Wildt, R.M., et al., *Antibody arrays for high-throughput screening of antibody-antigen interactions.* Nature biotechnology, 2000. 18(9): p. 989-94.
70. Desmyter, A., et al., *Antigen specificity and high affinity binding provided by one single loop of a camel single-domain antibody.* The Journal of biological chemistry, 2001. 276(28): p. 26285-90.
71. Tanha, J., et al., *Selection by phage display of llama conventional V(H) fragments with heavy chain antibody V(H)H properties.* J Immunol Methods, 2002. 263(1-2): p. 97-109.

72. Vranken, W., et al., *Solution structure of a llama single-domain antibody with hydrophobic residues typical of the VH/VL interface.* Biochemistry, 2002. 41(27): p. 8570-9.
73. Tanha, J., et al., *Improving solubility and refolding efficiency of human V(H)s by a novel mutational approach.* Protein Eng Des Sel, 2006. 19(11): p. 503-9.
74. Ewert, S., et al., *Biophysical properties of camelid V(HH) domains compared to those of human V(H)3 domains.* Biochemistry, 2002. 41(11): p. 3628-36.
75. Famm, K. and G. Winter, *Engineering aggregation-resistant proteins by directed evolution.* Protein engineering, design & selection : PEDS, 2006. 19(10): p. 479-81.
76. Dudgeon, K., R. Rouet, and D. Christ, *Rapid prediction of expression and refolding yields using phage display.* Protein engineering, design & selection : PEDS, 2013. 26(10): p. 671-4.
77. Christ, D., K. Famm, and G. Winter, *Repertoires of aggregation-resistant human antibody domains.* Protein engineering, design & selection : PEDS, 2007. 20(8): p. 413-6.
78. Christ, D., K. Famm, and G. Winter, *Tapping diversity lost in transformations--in vitro amplification of ligation reactions.* Nucleic acids research, 2006. 34(16): p. e108.
79. Famm, K., et al., *Thermodynamically stable aggregation-resistant antibody domains through directed evolution.* Journal of molecular biology, 2008. 376(4): p. 926-31.
80. Perchiacca, J.M., et al., *Structure-based design of conformation- and sequence-specific antibodies against amyloid beta.* Proceedings of the National Academy of Sciences of the United States of America, 2012. 109(1): p. 84-9.
81. Ladiwala, A.R., et al., *Rational design of potent domain antibody inhibitors of amyloid fibril assembly.* Proceedings of the National Academy of Sciences of the United States of America, 2012. 109(49): p. 19965-70.
82. Perchiacca, J.M., C.C. Lee, and P.M. Tessier, *Optimal charged mutations in the complementarity-determining regions that prevent domain antibody aggregation are dependent on the antibody scaffold.* Protein engineering, design & selection : PEDS, 2014. 27(2): p. 29-39.
83. Wang, N., et al., *Conserved amino acid networks involved in antibody variable domain interactions.* Proteins, 2009. 76(1): p. 99-114.
84. Nguyen, V.K., et al., *Camel heavy-chain antibodies: diverse germline V(H)H and specific mechanisms enlarge the antigen-binding repertoire.* The EMBO journal, 2000. 19(5): p. 921-30.
85. Mandrup, O.A., et al., *A novel heavy domain antibody library with functionally optimized complementarity determining regions.* PLoS One, 2013. 8(10): p. e76834.
86. Perchiacca, J.M., et al., *Aggregation-resistant domain antibodies engineered with charged mutations near the edges of the complementarity-determining regions.* Protein engineering, design & selection : PEDS, 2012. 25(10): p. 591-601.
87. Wilson, D.R. and B.B. Finlay, *Phage display: applications, innovations, and issues in phage and host biology.* Canadian journal of microbiology, 1998. 44(4): p. 313-29.
88. Lawrence, M.S., K.J. Phillips, and D.R. Liu, *Supercharging proteins can impart unusual resilience.* Journal of the American Chemical Society, 2007. 129(33): p. 10110-2.
89. Schaefer, J.V. and A. Pluckthun, *Engineering aggregation resistance in IgG by two independent mechanisms: lessons from comparison of Pichia pastoris and mammalian cell expression.* Journal of molecular biology, 2012. 417(4): p. 309-35.
90. Chen, W., et al., *Construction of a large phage-displayed human antibody domain library with a scaffold based on a newly identified highly soluble, stable heavy chain variable domain.* Journal of molecular biology, 2008. 382(3): p. 779-89.
91. Saerens, D., et al., *Disulfide bond introduction for general stabilization of immunoglobulin heavy-chain variable domains.* Journal of molecular biology, 2008. 377(2): p. 478-88.

92. Hagihara, Y., S. Mine, and K. Uegaki, *Stabilization of an immunoglobulin fold domain by an engineered disulfide bond at the buried hydrophobic region.* The Journal of biological chemistry, 2007. 282(50): p. 36489-95.
93. Chan, P.H., et al., *Engineering a camelid antibody fragment that binds to the active site of human lysozyme and inhibits its conversion into amyloid fibrils.* Biochemistry, 2008. 47(42): p. 11041-54.
94. Kim, D.Y., W. Ding, and J. Tanha, *Solubility and stability engineering of human VH domains.* Methods Mol Biol, 2012. 911: p. 355-72.
95. Hussack, G., et al., *Protease-resistant single-domain antibodies inhibit Campylobacter jejuni motility.* Protein engineering, design & selection : PEDS, 2014.
96. Kim, D.Y. and J. Tanha, *Engineering of immunoglobulin domains WO/2012/100343*, 2012.
97. Steipe, B., et al., *Sequence statistics reliably predict stabilizing mutations in a protein domain.* Journal of molecular biology, 1994. 240(3): p. 188-92.
98. Wirtz, P. and B. Steipe, *Intrabody construction and expression III: engineering hyperstable V(H) domains.* Protein science : a publication of the Protein Society, 1999. 8(11): p. 2245-50.
99. Reiter, Y., et al., *An antibody single-domain phage display library of a native heavy chain variable region: isolation of functional single-domain VH molecules with a unique interface.* Journal of molecular biology, 1999. 290(3): p. 685-98.
100. Voordijk, S., et al., *Molecular dynamics simulations highlight mobile regions in proteins: A novel suggestion for converting a murine V(H) domain into a more tractable species.* Journal of molecular biology, 2000. 300(4): p. 963-73.
101. Ohage, E. and B. Steipe, *Intrabody construction and expression. I. The critical role of VL domain stability.* Journal of molecular biology, 1999. 291(5): p. 1119-28.
102. Paz, K., et al., *Human single-domain neutralizing intrabodies directed against Etk kinase: a novel approach to impair cellular transformation.* Molecular cancer therapeutics, 2005. 4(11): p. 1801-9.
103. van den Beucken, T., et al., *Building novel binding ligands to B7.1 and B7.2 based on human antibody single variable light chain domains.* Journal of molecular biology, 2001. 310(3): p. 591-601.
104. Soderlind, E., M. Vergeles, and C.A. Borrebaeck, *Domain libraries: synthetic diversity for de novo design of antibody V-regions.* Gene, 1995. 160(2): p. 269-72.
105. Colby, D.W., et al., *Potent inhibition of huntingtin aggregation and cytotoxicity by a disulfide bond-free single-domain intracellular antibody.* Proceedings of the National Academy of Sciences of the United States of America, 2004. 101(51): p. 17616-21.
106. Colby, D.W., et al., *Development of a human light chain variable domain (V(L)) intracellular antibody specific for the amino terminus of huntingtin via yeast surface display.* Journal of molecular biology, 2004. 342(3): p. 901-12.
107. Lee, W.R., et al., *Gene silencing by cell-penetrating, sequence-selective and nucleic-acid hydrolyzing antibodies.* Nucleic acids research, 2010. 38(5): p. 1596-609.
108. Schiefner, A., et al., *A disulfide-free single-domain V(L) intrabody with blocking activity towards huntingtin reveals a novel mode of epitope recognition.* Journal of molecular biology, 2011. 414(3): p. 337-55.
109. Hurle, M.R., et al., *A role for destabilizing amino acid replacements in light-chain amyloidosis.* Proceedings of the National Academy of Sciences of the United States of America, 1994. 91(12): p. 5446-50.
110. Frisch, C., et al., *Contribution of the intramolecular disulfide bridge to the folding stability of REIv, the variable domain of a human immunoglobulin kappa light chain.* Folding & design, 1996. 1(6): p. 431-40.
111. Raffen, R., et al., *Physicochemical consequences of amino acid variations that contribute to fibril formation by immunoglobulin light chains.* Protein science : a publication of the Protein Society, 1999. 8(3): p. 509-17.

112. Pokkuluri, P.R., et al., *Factors contributing to decreased protein stability when aspartic acid residues are in beta-sheet regions.* Protein science : a publication of the Protein Society, 2002. 11(7): p. 1687-94.
113. Lee, C.V., et al., *High-affinity human antibodies from phage-displayed synthetic Fab libraries with a single framework scaffold.* Journal of molecular biology, 2004. 340(5): p. 1073-93.
114. Tanaka, T. and T.H. Rabbitts, *Protocol for the selection of single-domain antibody fragments by third generation intracellular antibody capture.* Nature protocols, 2010. 5(1): p. 67-92.
115. Arbabi-Ghahroudi, M., J. Tanha, and R. MacKenzie, *Isolation of monoclonal antibody fragments from phage display libraries.* Methods Mol Biol, 2009. 502: p. 341-64.
116. Arbabi-Ghahroudi, M., R. Mackenzie, and J. Tanha, *Site-directed mutagenesis for improving biophysical properties of VH domains.* Methods Mol Biol, 2010. 634: p. 309-30.
117. Wu, S.J., et al., *Structure-based engineering of a monoclonal antibody for improved solubility.* Protein engineering, design & selection : PEDS, 2010. 23(8): p. 643-51.
118. Chen, W., et al., *A large human domain antibody library combining heavy and light chain CDR3 diversity.* Molecular immunology, 2010. 47(4): p. 912-21.
119. Ignatovich, O., et al., *Creation of the large and highly functional synthetic repertoire of human VH and Vkappa domain antibodies.* Methods in molecular biology, 2012. 911: p. 39-63.
120. Spinelli, S., et al., *Lateral recognition of a dye hapten by a llama VHH domain.* Journal of molecular biology, 2001. 311(1): p. 123-9.
121. Streltsov, V.A., J.A. Carmichael, and S.D. Nuttall, *Structure of a shark IgNAR antibody variable domain and modeling of an early-developmental isotype.* Protein science : a publication of the Protein Society, 2005. 14(11): p. 2901-9.
122. Chennamsetty, N., et al., *Design of therapeutic proteins with enhanced stability.* Proceedings of the National Academy of Sciences of the United States of America, 2009. 106(29): p. 11937-42.

Table 1: Approaches for improving the biophysical properties of human single domain antibodies.

Approach	Description	Reference
V_H/V_L subclass selection	Selecting a V_H or V_L germ-line subclass with superior biophysical properties as library scaffolds could provide a greater frequency of isolated binders with desired properties	[21, 32]
In silico engineering	Identification of aggregation "hot spots" for modification	[52, 64]
Library selection	Applying heat or extreme pH during sdAb selection (i.e., via phage display) can lead to antibodies with desirable biophysical properties	[47, 48, 63, 79]
Site-directed mutagenesis	Influence of overall charge, camelization, V_H/V_L interface mutations, charged FR and CDR mutations and charged fusion tags	[16, 20, 22, 47, 49, 64, 82, 86, 115-117]
Random mutagenesis	Random mutagenesis of libraries	[106]
Disulfide linkage engineering	Introduction of an additional disulfide bridge within V_H or V_L domains can lead to increased thermodynamic and proteolytic stability	[35, 36, 94, 96]
CDR grafting/rearrangement	CDR amplification and random rearrangement	[64, 90, 118, 119]

Table 2: Site-specific mutations affecting biophysical properties of human single-domain antibodies.

Wild-type sdAbs	Mutations[§]	Effects					References
		Expression yield[a]	Stability[b]	TRE[c]	Aggregation Resistance[d]	sdAb binding[e]	
V_H							
VH-Ox13	•VH-P1 (G44E/L45R/W47G/***I51V***)	↓	↓		↑		[12-14]
	•VH-P8 (G44E/ L45R/W47I/***I51V***)	↓	↓		↑		
	•VH-Ox21-AVRL (G44E/ L45R/W47G/***I51V/L95K***)	↓	↓				
VH-Ox21-AVRL	•AFRL (V37F)	↑	↑				
	•AFAL (V37F/R94A)	↑	↑				
	•CVRC (***A33C/L100bC***)	↓	↑				
	•CFRC (***A33C***/V37F/***L100bC***)	↑	↑				
	•CFAC (***A33C***/V37F/R94A/***L100bC***)	↑	↑				
	•G47I	↑	↑				
IcaH-01 VH	•A16G		↑		↑		[98]
	•A16G/***I58T***		↑		↑		
	•A16G/***I58T***/R43Q		↑		↑		
	•A16G/***I58T***/R43Q/P75S		↑		↑		
	•A16G/***I58T***/R43Q/P75S/P7S	↑	↑		↑		
DP47d VH	•W47R	↑	↓		↑		[47]
	•***S35G***	↑	↓		↑		
BT32/A6 VH	•E6A/S23A/S82aN/V93A/ T108Q	≈	↑	↑	↑		[73]
	•E6A/S23A/S82aN/V93A/ T108Q/S74A/R83K/A84P	≈	↑	↑	↑		
VH-4D5	•VH-A1 (***H35G***/L45Y/***S93R/R94T/W95F/ G96T/G97T/Y100aN/A100bS/M 100cK/D101K/Y102A***)	↑	↑	↑	↑		[66]
	•VH-A2 (***H35G***/V37F/L45M/W47M/***S93T /R94S/G96Y/G97K/Y100aN/A10 0bS/M100cT/D101V/Y102I***)	↑	↑	↑	↑		
	•VH-A3 (***H35A***/L45V/W47M/***S93T/R94S/ W95K/G96K/G97K/A100bS/M10 0cS/D101P/Y102I***)	↑	↑	↑	↑		
	•VH-A4 (***H35S***/V37A/L45M/W47S/***S93V/ R94T/W95G/G96N/G97R/ A100bT/M100cL/D101K/Y102K***)	↑	↑	↑	↑		
	•VH-A5 (***H35S***/L45R/W47E/***S93I/W95K/ G96L/Y100aN/A100bR/M100cS/ D101N/Y102A***)	↑	↑	↑	↑		
	•VH-A6 (***H35G***/V37L/L45V/W47V/***S93G /R94V/W95S/G96I/G97N/ A100bK/ D101F/Y102H***)	↑	↑	↑	↑		
VH-A1	•VH-B1 (Q39R/Y45E/R50S/***T94S/F95L/ N98D/K101T***/W103R)	↑	↑	↑	↑		
VH-4D5	•VH-B1a (***H35G***/Q39R/L45E/***R50S/S93A***)	↑	↑	↑	↑		
VH-B1a	•W47A		↓	↑	↑		
	•W47V		↑	≈	↑		
	•W47L		↑	≈	↑		
	•W47E		↓	≈	↑		
	•W47T		↓	↑	↑		
	•W47L/V37S		↓	↑	↑		
	•W47L/V37T				≈		
	•W47L/W103R		↑	↑	↓		
	•W47L/W103S				≈		

	•W47L/W103T				≈		
	•W47L/R39S				≈		
	•W47L/R39T				↓		
	•W47L/R39H				↓		
	•W47L/R39K				≈		
	•W47L/R39D				≈		
	•W47L/R39E				≈		
	•W47L/E45H				↓		
	•W47L/E45S				↓		
	•W47L/E45T				↓		
	•***G35H***		↑	↑	↓		
	•R39Q		↑	≈	↓		
	•E45L		↑	≈	↓		
	•***S50R***		↓	≈	↑		
	•W47L/***G35H***		↓	≈	↓		
	•W47L/R39Q		↓	≈	↓		
	•W47L/E45L		↑	≈	↓		
	•W47L/***S50R***		↓	≈	↑		
	•W47T/***G35H***		↓	≈	↓		
	•W47T/R39Q		↓	≈	≈		
	•W47T/E45L		↑	≈	↓		
	•W47T/***S50R***		≈	≈	↑		
DP47d VH	•T28R*		≈		↑		[79]
HEL4 VH	•R28T*		≈		↓		
PH3-10 VH	•R28T*		≈		↓		
DP47d VH	•E1R/L11K/P41D/***D62K***/S112R		↓		↑		[64]
	•L11D/P41E/G66E/S75D/ S112E		↓		↑		
	•W47R		↓		≈		
	•W47L		↓		↑		
	•T28R		↓		≈		
	•F29I		↓		↑		
	•***S31D***		↓		↑		
	•***Y32E***		↓		↑		
	•***S35G***		↓		↑		
	•31***SYA***33 → 31***DED***33		↓		↑		
	•F29D		↓		↑		
	•$CDR1_{Hel4}$		↓		↑		
	•$CDR2_{Hel4}$		↓		↑		
	•$CDR3_{Hel4}$		↓		≈		
huVHAm302	•S49C/I69C	≈	↑	↓	↑	↓	[35]
huVHAm427	•S49C/I69C	≈	↑	↓	≈	↓	
huVHAm431	•S49C/I69C	≈	↑	↑	↑	↓	
huVHPC235	•S49C/I69C	≈	↑	↑	↑	↓	
DP47d VH	•***S31D***		≈				[49]
	•***Y32E***		↓				
	•***A33D***		↓				
	•***31-33DED***		↓				
	•Single mutations : ***S31D*** ***Y32E*** ***A33D***	↑		↑	↑		
	•Double mutations : ***31/32DE*** ***31/33DD*** ***32/33DD***	↑		↑	↑		
	•Triple mutations : ***31-33DED***	↑		↑	↑		
Aβ15-24 VH	•Charged Aβ15-24 (Aβ15-24 +L11D/P41D/K43E/ S74D/S112D)	≈	≈		↑		[81, 86]
	•***D/DD/DED***-Charged Aβ15-24-***D/DD/DED***	↑	≈		↑	≈	
	•***DED***- Charged Aβ15-24	↑	≈		↑		
	•Charged Aβ15-24-***DED***	↑	≈		↑		
Aβ33-42 VH	•Charged Aβ33-42 (Aβ33-42 + L11D/ P41D/K43E /S74D/S112D)	≈	≈		↑		
	•***D/DD/DED***- Charged Aβ33-42-	↑	≈		↑	≈	

	D/DD/DED						
	•***DED***-ChargedAβ33-42	≈	≈		↑		
	•Charged Aβ33-42-***DED***	↑	≈		↑		
Aβ12-21 VH	•***DED***-Aβ12-21-***DED***	↑	≈		↑		
	•***DED***-Aβ12-21	≈	≈		↑		
	•Aβ12-21-***DED***	↑	≈		↑		
Aβ18-27 VH	•***DED***-Aβ18-27-***DED***	↑	≈		↑		
	•***DED***-Aβ18-27	↑	≈		↑		
	•Aβ18-27-***DED***	≈	≈		↑		
Aβ18-27 VH	•***D/DD/DDD***-Aβ18-27	↑	≈	↑	↑	≈	[82]
	•***E/EE/EEE***-Aβ18-27	↑	≈	↑	↑	≈	
	•***A/AA/AAA***-Aβ18-27	≈	≈		≈	≈	
	•***R/RR/RRR***-Aβ18-27	≈	≈		↑	≈	
	•***K/KK/KKK***-Aβ18-27	↓	≈		↑	≈	
Aβ33-42 VH	•Aβ33-42-***D/DD/DDD***	↑	≈		↑	≈	
	•Aβ33-42-***R/RR/RRR***	≈	≈		↑	≈	
	•Aβ33-42-***A/AA/AAA***	≈	≈		≈	≈	
Aβ33-42-DDD	•L11D/G26E/K43D/K64D/Q105E/S112D	↑	↓		↑	≈	
Aβ33-42-RRR	•L11D/G26E/K43D/K64D/Q105E/S112D	↓	↓		↑	≈	
Aβ33-42-AAA	•L11D/G26E/K43D/K64D/Q105E/S112D	≈	↓		≈	≈	
Aβ33-42-DDD	•L11R/G26R/S74R/E85R/A94R/Q105K/S112R	↓	↓		↑	≈	
Aβ33-42-RRR	•L11R/G26R/S74R/E85R/A94R/Q105K/S112R	↓	↓		↑		
Aβ33-42-AAA	•L11R/G26R/S74R/E85R/A94R/Q105K/S112R	↓	↓		≈		
V_L							
REI-VL	•Y49F		≈		↓		[109]
	•G57E		↓		↓		
	•R61N		↓		↓		
	•G68D		↓		↓		
	•L78T		↓		≈		
	•A84T		↓		↓		
VL-LEN	•***Y27dD/N29D***		↑		≈		[111]
	•***Y27dD***		↑		≈		
	•***Y27dD/N29D/K30T***		↑		≈		
	•***V27bL***		↑		≈		
	•***S97T***		↑		≈		
	•***Y96P***		↓		↓		
	•***K30T***		↓		↓		
	•L15P		↓		↓		
	•***N28F***		↓		↓		
	•Len***CDR1***$_{REC}$		↓		↓		
	•Len***CDR1***$_{REC}$***Y96P***		↓		≈		
	•Len***CDR1***$_{REC}$***T30K***		↓		↓		
	•***S29N***		↑		≈		
	•***T94H***		↑		≈		
	•***S97T***		↑		≈		
	•I106L		↑		≈		
	•***K30R***		↓		≈		
	•P40L		↓		↓		
	•***Q89H***		↓		≈		
	•***Y96Q***		↓		↓		
VL-REC	•Rec***T30K***		↓		≈		
REI-VL	•T39K^{SS}		↑				[110]
	•T39K$^{(SH)2}$		↓				
	•***Y32H***^{SS}		↑				
	•***Y32H***$^{(SH)2}$		↓				
	•F73L^{SS}		↑				
	•Y71F^{SS}		↑				
	•Y71F$^{(SH)2}$		↓				
	•***Y32H***/T39K^{SS}		↑				
	•***Y32H***/T39K$^{(SH)2}$		↓				
	•T39K/Y71F^{SS}		↑				

	•T39K/Y71F$^{(SH)2}$		↓				
	•T39K/F73L^{SS}		↑				
	•T39K/F73L$^{(SH)2}$		↓				
	•Y71F/F73L^{SS}		↑				
	•Y71F/F73L$^{(SH)2}$		↓				
	•T39K/Y71F/F73L^{SS}		↑				
	•T39K/Y71F/F73L$^{(SH)2}$		↓				
	•C23V/***Y32H***	↑	↓				
	•C23V/***Y32H***/T39K		↓				
VL-Len	•***Y27dD***		↑				[112]
	•M4L/***Y27dD/T94H***		↑				
	•Q38D		↓				
	•Q38N		↓				
	•Q38D/***Y27dD***		↑				
	•Q38N/***Y27dD***		↑				
	•***Q89D/Y27dD***	↑	↓				
	•***Q89N/Y27dD***		↑				
	•***Q89D***/M4L/***Y27dD/T94H***	↑	↓				
	•***Q89N***/M4L/***Y27dD/T94H***		↑				
	•***Q89D***		↓				
	•***Q89N***		↓				
DPK9 VL	•***A50D***		↑				[49]
	•***S52D***		↓				
	•***S53D***		↓				
	•***50,52-53DDD***		↓				
	Single mutations:	↑		↑	≈		
	A50D						
	S52D						
	S53D						
	A51D						
	Double mutations:	↑		↑	≈		
	50/52DD						
	50/53DD						
	51/53DD						
	52/53DD						
	Triple mutations:	↑		↑	≈		
	50,52-53DDD						
HVLP324	•I48C/G64C	↓	↑		≈	≈	[36]
HVLP325	•I48C/G64C	↓	↑		≈	≈	
HVLP335	•I48C/G64C	↓	↑		≈	≈	
HVLP342	•I48C/G64C	↓	↑		≈	≈	
HVLP351	•I48C/G64C	↓	↑		≈	≈	
HVLP364	•I48C/G64C	↓	↑		↓	≈	
HVLP3103	•I48C/G64C	↓	↑		≈	≈	
HVLP389	•I48C/G64C	↓	↑		≈	≈	

[a] Yield of soluble protein purified from 1 liter of bacterial culture (mg/L).

[b] Assessed by T_m (°C) or ΔG_{UNF} /ΔG_F (kJ/mol or kcal/mol) or $[Urea]_{50}$. The mutations assessed in acidic conditions (pH 3.2) are marked with asterisks (*).

[c] Assessed by the reversibility of thermal unfolding using CD or surface plasmon resonance (SPR).

[d] Assessed by SEC (monomer fraction or % recovery or elution time (min) or volume (ml)), % soluble after melting above 70°C (non-native solubility) or % soluble at 25°C (native solubility) or fibril formation using Thioflavin T or t_{50} (min)/T_{50} (°C) or Congo red binding (mol/mol).

[e] Assessed by K_D (nM) on SPR or IC_{50} (nM) using ELISA.

The symbols, ↑ (increased), ↓ (decreased), ≈ (unchanged), indicate the changes of the biophysical properties resulting from the mutations compared to corresponding wild-type values. [§]Residue numbering is as defined by Kabat *et al.* [59]. TRE, thermal refolding efficiency. CDR mutations are bolded and italicized, while the rest of the mutations are framework region (FR) sequence-related.

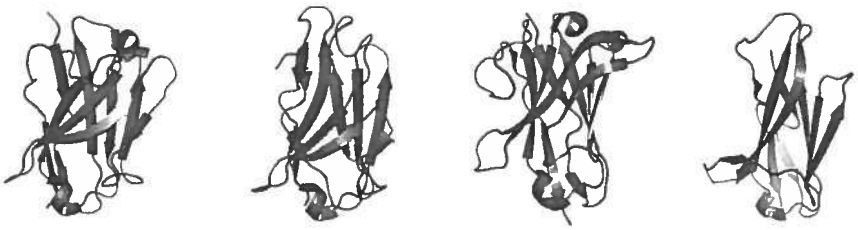
FIGURE LEGENDS

Figure 1: Single domain antibodies, sources and general biophysical characteristics. Single domain antibodies can be produced from conventional human IgGs (V_Hs and V_Ls), from Camelidae heavy-chain IgGs (V_HHs) and from cartilaginous fish IgNARs (V_{NAR}s). Libraries of sdAbs can be constructed from naïve, synthetic, semi-synthetic, transgenic animal or immunized sources. All sdAbs have a similar tertiary structure with two β-sheets formed through multiple anti-parallel β-strands and 3 complementarity determining regions (CDRs), some or all of which are involved in antigen/target binding. The inherent biophysical properties of sdAbs differ drastically depending on the type of antibody with human V_Hs and V_Ls generally exhibiting poor solubility and an increased tendency for aggregation, while V_HHs and V_{NAR}s tend to be highly soluble, resist aggregation and are readily expressed at high yields recombinantly. One advantage of V_Hs and V_Ls is the lower perceived risk of immunogenicity relative to V_HHs and V_{NAR}s. Representative protein structures of V_H, V_L, V_HH and V_{NAR} sdAbs are taken from Barthelemy *et al.* [66], Schiefner *et al.* [108], Spinelli *et al.* [120] and Streltsov *et al.* [121], respectively, with the N-termini at the top of the images and the C-termini at the bottom.

Figure 2: Approaches for isolating and engineering human single domain antibody affinity reagents for therapeutic applications. Work-flow for isolating human sdAb affinity reagents for therapeutic applications can be broken down into several key steps, ranging from selection of human sdAb scaffolds to affinity/stability selection following library construction to post-selection engineering of human sdAbs. In particular, unlike numerous physico-chemical methods employed, *in silico* engineering relies on computational simulation methods allowing predict hydrophobic patches on the surface of an antibody molecule through assessment of each amino acid's contribution in terms of hydrophobicity [122], which could benefit to bypass laborious works such as protein production and characterization that are needed to screen for antibodies with reduced aggregation propensity.

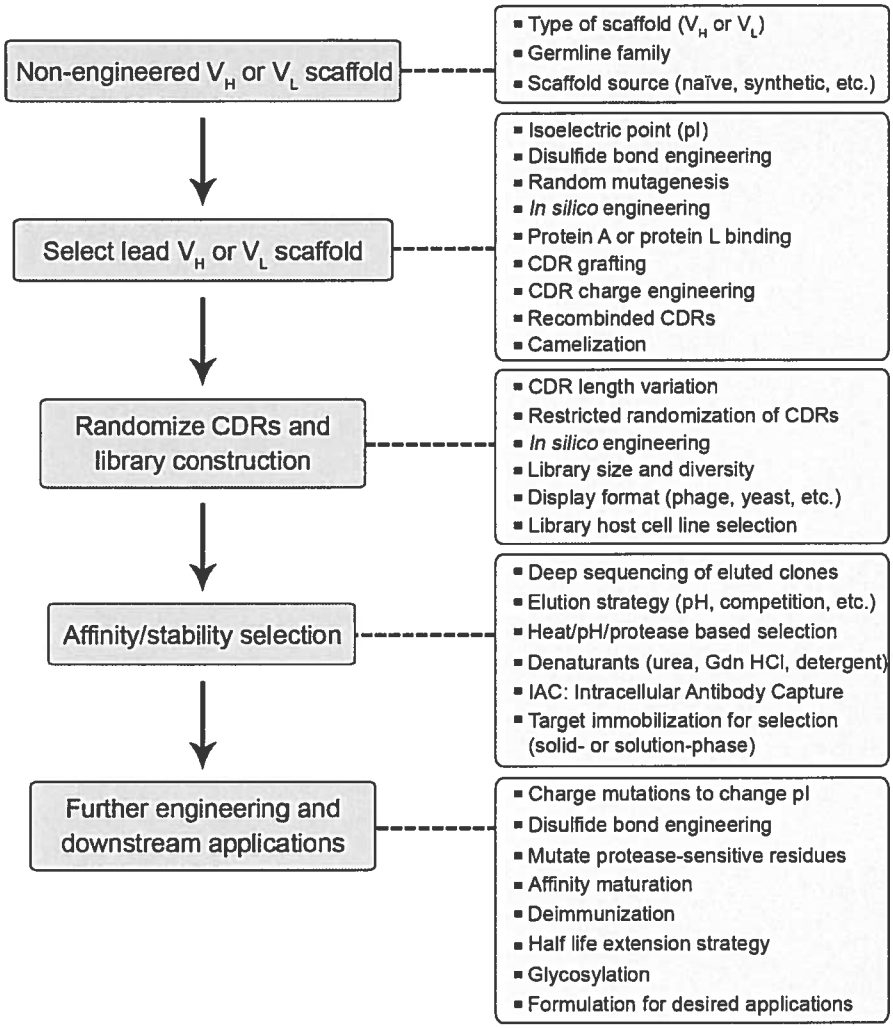
FIGURES

Figure 1



Type of sdAb:	V_H	V_L	V_HH	V_{NAR}
PDB reference:	3B9V	3LRH	1I3V	2COQ
Species:	Human	Human	Camelidae	Cartilaginous fish
Library source(s):	Naive Synthetic Transgenic mice	Naive Synthetic	Immunized Naive Synthetic	Immunized Naive Synthetic
Characteristics (non-engineered)	Poor solubility Aggregation prone Low expression Low immunogenicity	Moderate solubility Aggregation prone Moderate expression Low immunogenicity	High solubility Aggregation resistant High expression Possibly immunogenic	High solubility Aggregation resistant High expression Possibly immunogenic

Figure 2


Non-engineered V_H or V_L scaffold
Type of scaffold (V_H or V_L)
Germline family
Scaffold source (naïve, synthetic, etc.)
Select lead V_H or V_L scaffold
Isoelectric point (pI)
Disulfide bond engineering
Random mutagenesis
In silico engineering
Protein A or protein L binding
CDR grafting
CDR charge engineering
Recombinded CDRs
Camelization
Randomize CDRs and library construction
CDR length variation
Restricted randomization of CDRs
In silico engineering
Library size and diversity
Display format (phage, yeast, etc.)
Library host cell line selection
Affinity/stability selection
Deep sequencing of eluted clones
Elution strategy (pH, competition, etc.)
Heat/pH/protease based selection
Denaturants (urea, Gdn HCl, detergent)
IAC: Intracellular Antibody Capture
Target immobilization for selection (solid- or solution-phase)
Further engineering and downstream applications
Charge mutations to change pI
Disulfide bond engineering
Mutate protease-sensitive residues
Affinity maturation
Deimmunization
Half life extension strategy
Glycosylation
Formulation for desired applications

Supplementary data for:

Mutational approaches to improve the biophysical properties of human single-domain antibodies

Dae Young Kim[1*], Greg Hussack[1*], Hiba Kandalaft[1], and Jamshid Tanha[1,2,3,4]

[1]Human Health Therapeutics Portfolio, National Research Council Canada, Ottawa, ON, Canada K1A 0R6, [2]Department of Biochemistry, Microbiology and Immunology, University of Ottawa, Ottawa, ON, Canada K1H 8M5, [3]School of Environmental Sciences, Ontario Agricultural College, University of Guelph, Guelph, ON, Canada N1G 2W1

Supplementary Table 1: Effects of site-specific mutations on human sdAbs.

Wild-type sdAbs	Mutations[a]	Effects	References
V_H:			
VH-Ox13	VH-P1 (G44E/L45R/W47G/**I51V**)	• Decreased soluble expression yield, mg/L (6 → 1)	[1-3]
	VH-P8 (G44E/L45R/W47I/**I51V**)	• Decreased soluble expression yield, mg/L (6 → 4)	
	VH-Ox21-AVRL (G44E/L45R/W47G/***I51V/ L95K***)	• Decreased soluble expression yield, mg/L (6 → 1) • Decreased T_m, °C (71.2 → 56.6)	
VH-Ox21-AVRL	AFRL (V37F)	• Increased soluble expression yield, mg/L (1 → 2) • Increased T_m, °C (56.6 → 58.2)	
	AFAL (V37F/R94A)	• Increased soluble expression yield, mg/L (1 → 2) • Increased T_m, °C (56.6 → 63.9)	
	CVRC ***(A33C/L100bC)***	• Decreased soluble expression yield, mg/L (1 → 0.3) • Increased T_m, °C (56.6 → 57.9)	
	CFRC (A33C/V37F/***L100bC)***	• Increased T_m, °C (56.6 → 61.1)	
	CFAC (A33C/V37F/R94A/***L100bC)***	• Increased T_m, °C (56.6 → 58.5)	
	G47I	• Increased soluble expression yield, mg/L (1 → 4) • Increased T_m, °C (56.6 → 67.4)	
IcaH-01 VH	A16G	• Increased t_{50}, min (31.2 → 51.2) • Increased T_{50}, °C (36.4 → 38.1) • Increased C_m (Urea), M (3.63 → 3.85)	[4]
	A16G/***I58T***	• Increased t_{50}, min (31.2 → 63.2) • Increased T_{50}, °C (36.4 → 39.8) • Increased C_m (Urea), M (3.63 → 3.88)	
	A16G/***I58T***/R43Q	• Increased t_{50}, min (31.2 → 149) • Increased T_{50}, °C (36.4 → 40.1) • Increased C_m (Urea), M (3.63 → 4.11)	
	A16G/***I58T***/R43Q/P75S	• Increased t_{50}, min (31.2 → 184) • Increased T_{50}, °C (36.4 → 41.4) • Increased C_m (Urea), M (3.63 → 4.34)	
	A16G/***I58T***/R43Q P75S/P7S	• Increased t_{50}, min (31.2 → 606) • Increased T_{50}, °C (36.4 → 42.5) • Increased C_m (Urea), M (3.63 → 4.62) • Improved soluble expression yield, mg/L (0 → 1.2)	
DP47d VH	W47R	• Decreased T_m, °C (61.4 → 54.8) • Decreased ΔG_{N-U}, kJ/mol (34.7 → 24.0) • Improved Superdex G75 recovery (<5% → 80%) • Improved % SEC monomer (90% monomer) • Increased soluble expression yield, mg/mL (2.9 → 6.1)	[5]
	S35G	• No impact on T_m, °C (61.4 → 61.6) • Decreased ΔG_{N-U}, kJ/mol (34.7 → 29.1) • Improved Superdex G75 recovery (<5% → 60%) • Improved % SEC monomer (85% monomer) • Increased soluble expression yield, mg/mL (2.9 → 10.3)	
BT32/A6 VH	BT32/A6.L1: E6A/S23A/S82aN /V93A /T108Q	• No impact on soluble expression yield (5 mg/L) • Improved % SEC monomer (51.3% →87%) • Improved thermal refolding (13% →44% (20 µg/mL); 58% → 94% (2 µg/mL) • Increased T_m, °C (62.4 → 63.4)	[6]

	BT32/A6.L2 BT32/A6.L1 +S74A/R83K/A84P	• No impact on soluble expression yield (5 mg/L) • Improved % SEC monomer (51.3% →90%) • Improved thermal refolding (13% →74% (20 μg/mL); 58% → 90% (2 μg/mL) • Increased T_m, °C (62.4 → 64.4)	
VH-4D5	VH-A1 (***H35G***/L45Y/S93R/R94T***/W95F/G96T/G97T/Y100aN/A100bS/M100cK/D101K/Y102A***)	• Increased soluble expression yield, mg/L (1 → 14) • Increased T_m, °C (58→ 73) • Increased monomer fraction (0.52 → 0.85) • Increased reversibility of thermal unfolding (α) (0.29 → 0.98)	[7]
	VH-A2 (***H35G***/V37F/L45M/W47M/S93T/R94S/***G96Y/G97K/Y100aN/A100bS/M100cT/D101V/Y102I***)	• Increased soluble expression yield, mg/L (1 → 13) • Increased T_m, °C (58→ 75) • Increased monomer fraction (0.52 → 0.60) • Increased reversibility of thermal unfolding (α) (0.29 → 0.70)	
	VH-A3 (***H35A***/L45V/W47M/S93T/R94S/***W95K/G96K/G97K/A100bS/M100cS/D101P/Y102I***)	• Increased soluble expression yield, mg/L (1 → 10) • Increased T_m, °C (58→ 68) • Increased monomer fraction (0.52 → 0.87) • Increased reversibility of thermal unfolding (α) (0.29 → 0.89)	
	VH-A4 (***H35S***/V37A/L45M/W47S/S93V/R94T***/W95G/G96N/G97R/A100bT/M100cL/D101K/Y102K***)	• Increased soluble expression yield, mg/L (1 → 8) • Increased T_m, °C (58→ 65) • Increased monomer fraction (0.52 → 0.88) • Increased reversibility of thermal unfolding (α) (0.29 → 0.91)	
	VH-A5 (***H35S***/L45R/W47E/S93I***/W95K/G96L/Y100aN/A100bR/M100cS/D101N/Y102A***)	• Increased soluble expression yield, mg/L (1 → 7) • Increased T_m, °C (58→ 68) • Increased monomer fraction (0.52 → 0.92) • Increased reversibility of thermal unfolding (α) (0.29 → 0.93)	
	VH-A6 (***H35G***/V37L/L45V/W47V/S93G/R94V***/W95S/G96I/G97N/A100bK/ D101F/Y102H***)	• Increased soluble expression yield, mg/L (1 → 6) • Increased T_m, °C (58→ 72) • Increased monomer fraction (0.52 → 0.72) • Increased reversibility of thermal unfolding (α) (0.29 → 0.77)	
VH-A1	VH-B1 (Q39R/Y45E/R50S/T94S***/F95L/N98D/K101T***/W103R)	• Increased soluble expression yield, mg/L (14 → 19) • Increased T_m, °C (58→ 73) • Increased monomer fraction (0.52 → 0.84) • Increased reversibility of thermal unfolding (α) (0.29 → 0.94)	
VH-4D5	VH-B1a ***(H35G***/Q39R/L45E/***R50S***/S93A***)***	• Increased T_m, °C (58→ 79) • Increased monomer fraction (0.52 → 0.86) • Increased reversibility of thermal unfolding (α) (0.29 → 0.93)	
VH-B1a	W47A	• Decreased T_m, °C (79 → 76) • No impact on monomer fraction (0.86 → 0.86) • Reduced elution time, min (46.3 → 30.1) • Increased reversibility of thermal unfolding (α) (0.93 → 0.98)	
	W47V	• Increased T_m, °C (79 → 83) • Increased monomer fraction (0.86 → 0.88) • Reduced elution time, min (46.3 → 33.5) • No impact on reversibility of thermal unfolding (α) (0.93 → 0.94)	
	W47L	• Increased T_m, °C (79 → 82) • Increased monomer fraction (0.86 → 0.91) • Reduced elution time, min (46.3 → 31.5) • No impact on reversibility of thermal unfolding (α) (0.93 → 0.95)	
	W47E	• Decreased T_m, °C (79 → 75) • Increased monomer fraction (0.86 →0.92) • Reduced elution time, min (46.3 → 29.8) • No impact on reversibility of thermal unfolding (α) (0.93 → 0.96)	
	W47T	• Decreased T_m, °C (79 → 75) • Increased monomer fraction (0.86 → 0.90)	

		• Reduced elution time, min (46.3 → 32.8) • Increased reversibility of thermal unfolding (α) (0.93 → 0.98)	
	W47L/V37S	• Decreased T_m, °C (82→ 72) • Increased monomer fraction (0.91 → 0.97) • Reduced elution time, min (31.5 → 30.9) • Increased reversibility of thermal unfolding (α) (0.93 → 0.99)	
	W47L/V37T	• No impact on monomer fraction (0.91 → 0.89) • No impact on elution time, min (31.5 → 31.3)	
	W47L/W103R	• Increased T_m, °C (82 → 83) • Decreased monomer fraction (0.91 → 0.56) • Reduced elution time, min (31.5 → 29.8) • Increased reversibility of thermal unfolding (α) (0.93 → 1.00)	
	W47L/W103S	• No impact on monomer fraction (0.91 → 0.92) • No impact on elution time, min (31.5 → 30.5)	
	W47L/W103T	• No impact on monomer fraction (0.91 → 0.88) • No impact on elution time, min (31.5 → 30.0)	
	W47L/R39S	• No impact on monomer fraction (0.91 → 0.90) • No impact on elution time, min (31.5 → 32.0)	
	W47L/R39T	• Decreased monomer fraction (0.91 → 0.84) • No impact on elution time, min (31.5 → 31.6)	
	W47L/R39H	• Decreased monomer fraction (0.91 → 0.85) • No impact on elution time, min (31.5 → 33.9)	
	W47L/R39K	• No impact on monomer fraction (0.91 → 0.90) • No impact on elution time, min (31.5 → 31.8)	
	W47L/R39D	• No impact on monomer fraction (0.91 → 0.88) • No impact on elution time, min (31.5 → 30.1)	
	W47L/R39E	• No impact on monomer fraction (0.91 → 0.88) • No impact on elution time, min (31.5 → 30.0)	
	W47L/E45H	• Decreased monomer fraction (0.91 → 0.86) • No impact on elution time, min (31.5 → 30.9)	
	W47L/E45S	• Decreased monomer fraction (0.91→ 0.83) • Reduced elution time, min (31.5 → 35.9)	
	W47L/E45T	• Increased aggregation (data not reported)	
	G35H	• Increased T_m, °C (79 → 80) • Decreased monomer fraction (0.86 → 0.57) • Reduced elution time, min (46.3 → 43.4) • Increased reversibility of thermal unfolding (α) (0.93 → 0.98)	
	R39Q	• Increased T_m, °C (79 → 82) • Decreased monomer fraction (0.86 → 0.79) • No impact on elution time, min (46.3 → 46.5) • No impact on reversibility of thermal unfolding (α) (0.93 → 0.95)	
	E45L	• Increased T_m, °C (79 → 83) • Decreased monomer fraction (0.86 → 0.80) • Increased elution time, min (46.3 → 74.8) • No impact on reversibility of thermal unfolding (α) (0.93 → 0.94)	
	S50R	• Decreased T_m, °C (79 → 77) • Increased monomer fraction (0.86 → 0.92) • Reduced elution time, min (46.3 → 30.1) • No impact on reversibility of thermal unfolding (α) (0.93 → 0.94)	

	W47L/***G35H***	• Decreased T_m, °C (82 → 77) • Decreased monomer fraction (0.91 → 0.79) • Increased elution time, min (31.5 → 33.5) • No impact on reversibility of thermal unfolding (α) (0.93 → 0.97)	
	W47L/R39Q	• Decreased T_m, °C (82 → 77) • Decreased monomer fraction (0.91 → 0.85) • No impact on elution time, min (31.5 → 31.4) • No impact on reversibility of thermal unfolding (α) (0.93 → 0.96)	
	W47L/E45L	• Increased T_m, °C (82 → 84) • Decreased monomer fraction (0.91 → 0.87) • Increased elution time, min (31.5 → 36.9) • No impact on reversibility of thermal unfolding (α) (0.93 → 0.96)	
	W47L/***S50R***	• Decreased T_m, °C (82 → 78) • Increased monomer fraction (0.91 → 0.96) • Reduced elution time, min (31.5 → 30.0) • No impact on reversibility of thermal unfolding (α) (0.93 → 0.96)	
	W47T/***G35H***	• Decreased T_m, °C (75 → 65) • Decreased monomer fraction (0.90 → 0.79) • Reduced elution time, min (32.8 → 31.3) • No impact on reversibility of thermal unfolding (α) (0.93 → 0.94)	
	W47T/R39Q	• Decreased T_m, °C (75 → 73) • No impact on monomer fraction (0.90 → 0.88) • No impact on elution time, min (32.8 → 32.7) • No impact on reversibility of thermal unfolding (α) (0.93 → 0.96)	
	W47T/E45L	• Increased T_m, °C (75 →80) • Decreased monomer fraction (0.90 → 0.79) • No impact on elution time, min (32.8 → 32.7) • No impact on reversibility of thermal unfolding (α) (0.93 → 0.96)	
	W47T/***S50R***	• No impact on T_m, °C (75 → 75) • Increased monomer fraction (0.90 → 0.96) • Reduced elution time, min (32.8 → 30.2) • No impact on reversibility of thermal unfolding (α) (0.93 → 0.97)	
DP47d VH	T28R	• Increased T_m, °C (47 → 48), at pH 3.2 • No impact on ΔG_{N-U}, kcal/mol (16 → 16), at pH 3.2 • Less aggregation prone (80°C and pH 3.2)	[8]
HEL4 VH	R28T	• No impact on T_m, °C (55 → 55), at pH 3.2 • Increased ΔG_{N-U}, kcal/mol (21 → 24), at pH 3.2 • More aggregation prone (80°C and pH 3.2)	
PH3-10 VH	R28T	• Decreased T_m, °C (60 → 58), at pH 3.2 • Increased ΔG_{N-U}, kcal/mol (27 → 30), at pH 3.2 • More aggregation prone (80°C and pH 3.2)	
DP47d VH	E1R/L11K/P41D/***D62K***/S112R	• Decreased ΔG_{N-U}, kJ/mol (32 ± 1.7 → 31 ± 1.7) • Decreased T_m, °C (66 ± 0.1 → 61 ± 0.2) • Increased % soluble after melting (27 → 42)	[9]
	L11D/P41E/G66E/S75D/ S112E	• No impact on ΔG_{N-U}, kJ/mol (32 ± 1.7 → 32 ± 1.2) • Decreased T_m, °C (66 ± 0.1 → 61 ± 0.1) • Increased % soluble after melting (27 → 50)	
	W47R	• Decreased ΔG_{N-U}, kJ/mol (32 ± 1.7 → 25 ± 2.0) • Decreased T_m, °C (66 ± 0.1 → 58 ± 0.2) • No impact on % soluble after melting (27 → 30)	
	W47L	• Decreased ΔG_{N-U}, kJ/mol (32 ± 1.7 → 26 ± 1.2) • Decreased T_m, °C (66 ± 0.1 → 58 ± 0.3) • Increased % soluble after melting (27 → 32)	

	T28R	• Decreased ΔG_{N-U}, kJ/mol (32 ± 1.7 → 31 ± 1.9) • Decreased T_m, °C (66 ± 0.1 → 64 ± 0.3) • No impact on % soluble after melting (27 → 30)	
	F29I	• Decreased ΔG_{N-U}, kJ/mol (32 ± 1.7 → 21 ± 2.0) • Decreased T_m, °C (66 ± 0.1 → 58 ± 0.8) • Increased % soluble after melting (27 → 34)	
	S31D	• Decreased ΔG_{N-U}, kJ/mol (32 ± 1.7 → 23 ± 1.7) • Decreased T_m, °C (66 ± 0.1 → 61 ± 1.2) • Increased % soluble after melting (27 → 37)	
	Y32E	• Decreased ΔG_{N-U}, kJ/mol (32 ± 1.7 → 28 ± 1.5) • Decreased T_m, °C (66 ± 0.1 → 61 ± 0.4) • Increased % soluble after melting (27 → 40)	
	S35G	• Decreased ΔG_{N-U}, kJ/mol (32 ± 1.7 → 28 ± 2.4) • Decreased T_m, °C (66 ± 0.1 → 61 ± 0.1) • Increased % soluble after melting (27→ 32)	
	31*SYA*33 → 31*DED*33	• Decreased ΔG_{N-U}, kJ/mol (32 ± 1.7 → 24 ± 1.2) • Decreased T_m, °C (66 ± 0.1 → 63 ± 0.4) • Increased % soluble after melting (27 → 98)	
	F29D	• Decreased ΔG_{N-U}, kJ/mol (32 ± 1.7 → 25 ± 1.3) • No impact on T_m, °C (66 ± 0.1 → 66 ± 0.2) • Increased % soluble after melting (27→ 100)	
	$CDR1_{Hel4}$	• Decreased ΔG_{N-U}, kJ/mol (32 ± 1.7 → 16 ± 0.5) • Decreased T_m, °C (66 ± 0.1 → 65 ± 0.2) • Increased % soluble after melting (27 → 97)	
	$CDR2_{Hel4}$	• Decreased ΔG_{N-U}, kJ/mol (32 ± 1.7 → 22 ± 1.2) • Decreased T_m, °C (66 ± 0.1 → 65 ± 0.1) • No impact on % soluble after melting (27 → 30)	
	$CDR3_{Hel4}$	• Decreased ΔG_{N-U}, kJ/mol (32 ± 1.7 → 16 ± 1.7) • No impact on T_m, °C (66 ± 0.1 → 65 ± 0.7) • No impact on % soluble after melting (27 → 28)	
huVHAm302	S49C/I69C	• Increased T_m, °C (53.8 ± 0.1 →71.4 ± 0.2) • Decreased reversibility of thermal unfolding (α) (0.60 ± 0.12 →0.46 ± 0.07) • Decreased affinity to protein A (K_D) (3 → 10) • Decreased % aggregate (13.9 ± 0.02 →4.1 ± 0.15) • Increased % recovery (69 ± 0.6 →76 ± 0.1)	[10]
huVHAm427	S49C/I69C	• Increased T_m, °C (73.0 ± 0.3 →89.2 ± 2.6) • Decreased reversibility of thermal unfolding (α) (0.84 ± 0.04 →0.75 ± 0.11) • Decreased affinity to protein A (K_D) (1.6 → 4) • No impact on % aggregate (3.8 ± 2.9 →6.6 ± 0.76)	
huVHAm431	S49C/I69C	• Increased T_m, °C (71.4 ± 0.4→87.5 ± 0.2) • Increased reversibility of thermal unfolding (α) (0.45 ± 0.02 →0.89 ± 0.07) • Decreased affinity to protein A (K_D) (4 → 8) • No impact on % aggregate (5.5 ± 0.08 →4.3 ± 0.14) • Increased % recovery (82 ± 0.5 →100 ± 1.9)	
huVHPC235	S49C/I69C	• Increased T_m, °C (59.1 ± 0.2 →73.0± 0.0) • Increased reversibility of thermal unfolding (α) (0.58 ± 0.05 →0.75 ± 0.08) • Decreased affinity to protein A (K_D) (0.3 → 3) • No impact on % aggregate (0 ± 0.02 →0.6 ± 0.09) • Increased % recovery (70 ± 3.7 →105 ± 3.6)	

DP47d VH	***S31D***	• No impact on ΔG_{N-U}, kJ/mol (40.4 → 40.2)	[11]
	Y32E	• Decreased ΔG_{N-U}, kJ/mol (40.4 → 24.7)	
	A33D	• Decreased ΔG_{N-U}, kJ/mol (40.4 → 34.3)	
	31-31DED	• Decreased ΔG_{N-U}, kJ/mol (40.4 → 26.8)	
	Single mutations : ***S31D*** ***Y32E*** ***A33D***	• Increased soluble expression yield, mg/L (1.7 → 4.7) • Increased % heat refolding (4 → 69.3) • Decreased Gel filtration elution volume, ml (24.9 → 23.1)	
	Double mutations : ***31/32DE*** ***31/33DD*** ***32/33DD***	• Increased soluble expression yield, mg/L (1.7 → 13.3) • Increased % heat refolding (4 → 82.3) • Descreased Gel filtration elution volume, ml (24.9 → 21.3)	
	Triple mutations : ***31-31DED***	• Increased soluble expression yield, mg/L (1.7 → 11) • Increased % heat refolding (4 → 88) • Decreased Gel filtration elution volume, ml (24.9 → 19.7)	
Aβ15-24 VH	Charged Aβ15-24 (Aβ15-24 + L11D/P41D/K43E/S74D/ S112D)	• No impact on soluble expression yield (5 – 7 mg/L) • No impact on secondary structure and stability • Modestly increased % soluble after melting above 70°C & soluble at 25°C (non-native and native solubility, respectively)	[12, 13]
	D-Charged Aβ15-24-***D***	• Increased soluble expression yield, mg/L (6 → 16) • Modestly increased non-native and native solubility • Modestly improved SEC behavior • No impact on ΔG_{N-U}, kJ/mol (49 ± 0.8 → 49 ± 1.2) • No impact on T_m, °C (78 → 78) • No impact on dAb binding, K_D, nM (370 ± 45 → 375 ± 45) • No impact on secondary structure	
	DD- Charged Aβ15-24-***DD***	• Increased soluble expression yield, mg/L (6 → 16) • Greatly improved non-native and native solubility • Modestly improved SEC behavior • No impact on ΔG_{N-U}, kJ/mol (49 ± 0.8 → 48 ± 1.5) • No impact on T_m, °C (78 → 78) • No impact on dAb binding, K_D, nM (370 ± 45 → 365 ± 35) • No impact on secondary structure	
	DED- Charged Aβ15-24-***DED***	• Increased soluble expression yield, mg/L (6 → 16) • Greatly improved non-native and native solubility • Greatly improved SEC behavior • No impact on ΔG_{N-U}, kJ/mol (49 ± 0.8 → 47 ± 1.6) • No impact on T_m, °C (78 → 77) • Increased soluble expression yield, mg/L (6 → 18) • No impact on dAb binding, K_D, nM (370 ± 45 → 370 ± 30) • No impact on secondary structure	
	DED- Charged Aβ15-24	• Increased soluble expression yield, mg/L (6 → 14) • Greatly improved SEC behavior • No impact on ΔG_{N-U}, kJ/mol (49 ± 0.8 → 48 ± 1.0) • No impact on T_m, °C (78 → 78) • No impact on secondary structure	
	Charged Aβ15-24-***DED***	• Increased soluble expression yield, mg/L (6 → 16) • Greatly improved SEC behavior • No impact on ΔG_{N-U}, kJ/mol (49 ± 0.8 → 49 ± 1.5) • No impact on T_m, °C (78 → 78) • No impact on secondary structure	
Aβ33-42 VH	Charged Aβ33-42 (Aβ33-42 + L11D/P41D/K43E/S74D/ S112D)	• No impact on soluble expression yield (5 – 7 mg/L) • No impact on secondary structure and stability • Modestly increased % soluble after melting above 70°C & soluble at 25°C (non-native and native solubility, respectively)	

	D- Charged Aβ33-42-***D***	• Increased soluble expression yield, mg/L (6 → 16) • Modestly increased non-native and native solubility • Modestly improved SEC behavior • No impact on ΔG_{N-U} kJ/mol (49 ± 1.4 → 49 ± 1.8) • No impact on T_m, °C (78 → 77) • No impact on K_D, nM (335 ± 20 → 340 ± 25) • No impact on secondary structure	
	DD- Charged Aβ33-42-***DD***	• Increased soluble expression yield, mg/L (6 → 16) • Greatly increased non-native and native solubility • Modestly improved SEC behavior • No impact on ΔG_{N-U} kJ/mol (49 ± 1.4 → 48 ± 1.0) • No impact on T_m, °C (78 → 78) • No impact on K_D, nM (335 ± 20 → 335 ± 45) • No impact on secondary structure	
	DED- Charged Aβ33-42-***DED***	• Increased soluble expression yield, mg/L (6 → 18) • Greatly increased non-native and native solubility • Greatly improved SEC behavior • No impact on ΔG_{N-U} kJ/mol (49 ± 1.4 → 48 ± 1.2) • No impact on T_m, °C (78 → 77) • No impact on K_D, nM (335 ± 20 → 340 ± 35) • No impact on secondary structure	
	DED-ChargedAβ33-42	• No impact on soluble expression yield, mg/L (6 → 6) • Modestly improved SEC behavior • Modestly increased non-native and native solubility • No impact on ΔG_{N-U} kJ/mol (49 ± 1.4 → 47 ± 1.7) • No impact on T_m, °C (78 → 77) • No impact on secondary structure	
	Charged Aβ33-42-***DED***	• Increased soluble expression yield, mg/L (6 → 15) • Greatly improved SEC behavior • Greatly increased non-native and native solubility • No impact on ΔG_{N-U} kJ/mol (49 ± 1.4 → 48 ± 1.2) • No impact on T_m, °C (78 → 78) • No impact on secondary structure	
Aβ12-21 VH	***DED***-Aβ12-21-***DED***	• Increased soluble expression yield, mg/L (5 → 17) • Greatly increased non-native and native solubility • No impact on ΔG_{N-U} kJ/mol (48 ± 1.5 → 47 ± 1.1) • Decreased T_m, °C (79 → 77) • No impact on secondary structure	
	DED-Aβ12-21	• No impact on soluble expression yield, mg/L (5 → 7) • Modestly increased non-native and native solubility • No impact on ΔG_{N-U} kJ/mol (48 ± 1.5 → 49 ± 1.7) • Decreased T_m, °C (79 → 78) • No impact on secondary structure	
	Aβ12-21-***DED***	• Increased soluble expression yield, mg/L (5 → 18) • Greatly increased non-native and native solubility • No impact on ΔG_{N-U} kJ/mol (48 ± 1.5 → 48 ± 1.6) • Decreased T_m, °C (79 → 78) • No impact on secondary structure	
	DED-Aβ18-27-***DED***	• Increased soluble expression yield, mg/L (6 → 18) • Greatly increased non-native and native solubility • No impact on ΔG_{N-U} kJ/mol (48 ± 1.2 → 47± 1.6) • No impact on T_m, °C (78 → 78) • No impact on secondary structure	
Aβ18-27 VH	***DED***-Aβ18-27	• Increased soluble expression yield, mg/L (6 → 18) • Greatly increased non-native and native solubility • No impact on ΔG_{N-U} kJ/mol (48 ± 1.2 → 49 ± 1.3) • Decreased T_m, °C (78 → 77)	

	Aβ18-27-***DED***	• No impact on secondary structure • No impact on soluble expression yield, mg/L (6 → 6) • Modestly increased non-native and native solubility • No impact on ΔG_{N-U}, kJ/mol (48 ± 1.2 → 48 ± 1.4) • No impact on T_m, °C (78 → 78) • No impact on secondary structure	
Aβ18-27 VH	***D***-Aβ18-27	• Increased soluble expression yield, mg/L (21 → 28) • Modestly increased reversibility of thermal unfolding • Modestly increased non-native & native solubility • No impact on ΔG_{N-U}, kJ/mol (49 ± 1.4 → 48 ± 1.3) • No impact on T_m, °C (77 ± 1.3 → 77 ± 0.9) • Modestly improved SEC behavior • No impact on secondary structure	[14]
	DD-Aβ18-27	• Increased soluble expression yield, mg/L (21 → 33) • Greatly increased reversibility of thermal unfolding • Greatly increased non-native & native solubility • No impact on ΔG_{N-U}, kJ/mol (49 ± 1.4 → 48 ± 2.1) • No impact on T_m, °C (77 ± 1.3 → 79 ± 0.8) • Greatly improved SEC behavior • No impact on secondary structure	
	DDD-Aβ18-27	• Increased soluble expression yield, mg/L (21 → 45) • Greatly increased reversibility of thermal unfolding • Greatly increased non-native & native solubility • Decreased ΔG_{N-U}, kJ/mol (49 ± 1.4 → 47 ± 1.5) • No impact on T_m, °C (77 ± 1.3 → 78 ± 0.4) • Increased reversibility of thermal unfolding • Increased non-native & native solubility • Improved SEC behavior • No impact on antigenic binding activity, IC_{50}, nM (891 ± 15 → 870 ± 12) • No impact on secondary structure	
	E-Aβ18-27	• Increased soluble expression yield, mg/L (21 → 28) • Modestly increased non-native & native solubility • No impact on ΔG_{N-U}, kJ/mol (49 ± 1.4 → 48 ± 1.7) • No impact on T_m, °C (77 ± 1.3 → 78 ± 1.6) • Modestly increased reversibility of thermal unfolding • Modestly increased non-native & native solubility • Modestly improved SEC behavior • No impact on secondary structure	
	EE-Aβ18-27	• Increased soluble expression yield, mg/L (21 → 31) • Greatly increased non-native & native solubility • No impact on ΔG_{N-U}, kJ/mol (49 ± 0.8 → 49 ± 2.4) • No impact on T_m, °C (77 ± 1.3 → 78 ± 0.8) • Increased reversibility of thermal unfolding • Increased non-native & native solubility • Improved SEC behavior • No impact on secondary structure	
	EEE-Aβ18-27	• Increased soluble expression yield, mg/L (21 → 41) • Greatly increased non-native & native solubility • No impact on ΔG_{N-U}, kJ/mol (49 ± 0.8 → 47 ± 1.8) • No impact on T_m, °C (77 ± 1.3 → 78 ± 0.5) • Increased reversibility of thermal unfolding • Increased non-native & native solubility • Improved SEC behavior • No impact on antigenic binding activity, IC_{50}, nM (891 ± 15 → 898 ± 17) • No impact on secondary structure	

	A-Aβ18-27	• No impact on soluble expression yield, mg/L (21 → 20) • No impact on non-native & native solubility • No impact on ΔG_{N-U}, kJ/mol (49 ± 0.8 → 49 ± 2.4) • No impact on T_m, °C (77 ± 1.3 → 77 ± 0.6) • No impact on secondary structure	
	AA-Aβ18-27	• No impact on soluble expression yield, mg/L (21 → 21) • No impact on non-native & native solubility • No impact on ΔG_{N-U}, kJ/mol (49 ± 0.8 → 49 ± 2.7) • No impact on T_m, °C (77 ± 1.3 → 78 ± 0.6) • No impact on secondary structure	
	AAA-Aβ18-27	• No impact on soluble expression yield, mg/L (21 → 20) • No impact on non-native & native solubility • No impact on ΔG_{N-U}, kJ/mol (49 ± 0.8 → 48 ± 1.9) • No impact on T_m, °C (77 ± 1.3 → 77 ± 0.4) • No impact on antigenic binding activity, IC_{50}, nM (891 ± 15 → 870 ± 12) • No impact on secondary structure	
	R-Aβ18-27	• No impact on soluble expression yield, mg/L (21 → 19) • Modestly increased non-native & native solubility • No impact on ΔG_{N-U}, kJ/mol (49 ± 1.4 → 46 ± 4.5) • No impact on T_m, °C (77 ± 1.3 → 76 ± 0.5) • No impact on secondary structure	
	RR-Aβ18-27	• Decreased soluble expression yield, mg/L (21 → 15) • Modestly increased non-native & native solubility • No impact on ΔG_{N-U}, kJ/mol (49 ± 1.4 → 47 ± 4.5) • No impact on T_m, °C (77 ± 1.3 → 77 ± 0.3) • No impact on secondary structure	
	RRR-Aβ18-27	• Decreased soluble expression yield, mg/L (21 → 9) • Greatly increased non-native & native solubility • Poor SEC behavior • No impact on ΔG_{N-U}, kJ/mol (49 ± 1.4 → 44 ± 4.4) • No impact on T_m, °C (77 ± 1.3 → 76 ± 0.7) • No impact on antigenic binding activity, IC_{50}, nM (891 ± 15 → 902 ± 24) • No impact on secondary structure	
	K-Aβ18-27	• Decreased soluble expression yield, mg/L (21 → 18) • No impact on non-native & native solubility • No impact on ΔG_{N-U}, kJ/mol (49 ± 1.4 → 45 ± 3.2) • No impact on T_m, °C (77 ± 1.3 → 77 ± 1.3) • No impact on secondary structure	
	KK-Aβ18-27	• Decreased soluble expression yield, mg/L (21 → 12) • Modestly increased non-native & native solubility • No impact on ΔG_{N-U}, kJ/mol (49 ± 0.8 → 46 ± 4.3) • No impact on T_m, °C (77 ± 1.3 → 76 ± 0.6) • No impact on secondary structure	
	KKK-Aβ18-27	• Decreased soluble expression yield, mg/L (21 → 10) • Greatly increased non-native & native solubility • Poor SEC behavior • No impact on T_m, °C (77 ± 1.3 → 76 ± 0.2) • No impact on antigenic binding activity, IC_{50}, nM (891 ± 15 → 872 ± 23) • No impact on secondary structure	
Aβ33-42 VH	Aβ33-42-***D***	• Increased soluble expression yield, mg/L (20 → 25) • Modestly increased non-native & native solubility • No impact on ΔG_{N-U}, kJ/mol (48 ± 1.7 → 48 ± 1.3) • No impact on T_m, °C (78 ± 1.5 → 78 ± 0.3) • No impact on secondary structure	

	Aβ33-42-***DD***	• Increased soluble expression yield, mg/L (20 → 36) • Greatly increased non-native & native solubility • No impact on $\Delta G_{N\text{-}U}$ kJ/mol (48 ± 1.7 → 48 ± 2.1) • No impact on T_m, °C (78 ± 1.5 → 79 ± 0.7) • No impact on secondary structure	
	Aβ33-42-***DDD***	• Increased soluble expression yield, mg/L (20 → 47) • Greatly increased non-native & native solubility • Greatly improved SEC behavior • No impact on $\Delta G_{N\text{-}U}$ kJ/mol (48 ± 1.7 → 48 ± 1.5) • No impact on T_m, °C (78 ± 1.5 → 78 ± 0.6) • No impact on antigenic binding activity, IC_{50}, nM (747 ± 11 → 752 ± 20) • No impact on secondary structure	
	Aβ33-42-***R***	• Decreased soluble expression yield, mg/L (20 → 17) • GreaModestly increased non-native & native solubility • No impact on $\Delta G_{N\text{-}U}$ kJ/mol (48 ± 1.7 → 45 ± 3.6) • No impact on T_m, °C (78 ± 1.5 → 77 ± 0.2) • No impact on secondary structure	
	Aβ33-42-***RR***	• Decreased soluble expression yield, mg/L (20 → 13) • Greatly increased non-native & native solubility • No impact on $\Delta G_{N\text{-}U}$ kJ/mol (48 ± 1.7 → 46 ± 2.9) • No impact on T_m, °C (78 ± 1.5 → 76 ± 0.2) • No impact on secondary structure	
	Aβ33-42-***RRR***	• Decreased soluble expression yield, mg/L (20 → 10) • Greatly increased non-native & native solubility • Modestly improved SEC behavior • No impact on $\Delta G_{N\text{-}U}$ kJ/mol (48 ± 1.7 → 44 ± 5.4) • No impact on T_m, °C (78 ± 1.5 → 77 ± 0.2) • No impact on antigenic binding activity, IC_{50}, nM (747 ± 11 → 746 ± 20) • No impact on secondary structure	
	Aβ33-42-***A***	• No impact on soluble expression yield, mg/L (20 → 21) • No impact on non-native & native solubility • No impact on $\Delta G_{N\text{-}U}$ kJ/mol (48 ± 1.7 → 48 ± 2.4) • No impact on T_m, °C (78 ± 1.5 → 77 ± 0.8) • No impact on secondary structure	
	Aβ33-42-***AA***	• No impact on soluble expression yield, mg/L (20 → 19) • No impact on non-native & native solubility • No impact on $\Delta G_{N\text{-}U}$ kJ/mol (48 ± 1.7 → 48 ± 2.7) • No impact on T_m, °C (78 ± 1.5 → 78 ± 0.9) • No impact on antigenic binding activity, IC_{50}, nM (747 ± 11 → 741 ± 22) • No impact on secondary structure	
	Aβ33-42-***AAA***	• No impact on soluble expression yield, mg/L (20 → 20) • No impact on non-native & native solubility • No impact on T_m, °C (78 ± 1.5 → 77 ± 0.3) • No impact on antigenic binding activity, IC_{50}, nM (747 ± 11 → 736 ± 31) • No impact on secondary structure	
Aβ33-42 VH	Aβ33-42-***DDD*** + L11D/G26E/K43D/K64D/Q105E/S112D	• Increased soluble expression yeld, mg/L (20 → 26) • Greatly improved non-native & native solubility • Greatly improved SEC behavior • Decreased T_m, °C (78 ± 1.5 → 71 ± 0.9) • No impact on antigenic binding activity, IC_{50}, nM (747 ± 11 → 737 ± 22) • No impact on secondary structure	
		• Decreased soluble expression yield, mg/L (20 → 10)	

Aβ33-42 VH	Aβ33-42-***RRR***+ L11D/G26E/K43D/K64D/Q105E/ S112D	• Modestly improved non-native & native solubility • Modestly improved SEC behavior • Decreased $\Delta G_{N\text{-}U}$, kJ/mol (48 ± 1.7 → 35 ± 5.3) • Decreased T_m, °C (78 ± 1.5 → 70 ± 0.7) • No impact on secondary structure	
	Aβ33-42-***AAA***+ L11D/G26E/K43D/K64D/Q105E/ S112D	• No impact on soluble expression yield, mg/L (20 → 20) • No impact on non-native & native solubility • No impact on SEC behavior • Decreased $\Delta G_{N\text{-}U}$, kJ/mol (48 ± 1.7 → 35 ± 3.9) • Decreased T_m, °C (78 ± 1.5 → 71 ± 0.6) • No impact on antigenic binding activity, IC_{50}, nM (747 ± 11 → 732 ± 16) • No impact on secondary structure	
	Aβ33-42-***DDD*** + L11R/G26R/S74R/E85R/A94R/Q 105K/S112R	• Decreased soluble expression yield, mg/L (20 → 15) • Modestly improved non-native & native solubility • Modes improved SEC behavior • Decreased $\Delta G_{N\text{-}U}$, kJ/mol (48 ± 1.7 → 32 ± 3.8) • Decreased T_m, °C (78 ± 1.5 → 69 ± 1.2) • No impact on antigenic binding activity, IC_{50}, nM (747 ± 11 → 749 ± 17) • No impact on secondary structure	
	Aβ33-42-***RRR***+ L11R/G26R/S74R/E85R/A94R/Q 105K/S112R	• Decreased soluble expression yield, mg/L (20 → 10) • Greatly improved non-native & native solubility • Greatly improved SEC behavior • Decreased $\Delta G_{N\text{-}U}$, kJ/mol (48 ± 1.7 → 33 ± 5.1) • Decreased T_m, °C (78 ± 1.5 → 70 ± 0.6) • No impact on secondary structure	
	Aβ33-42-***AAA***+ L11R/G26R/S74R/E85R/A94R/Q 105K/S112R	• Decreased soluble expression yield, mg/L (20 → 16) • No impact on non-native & native solubility • No impact on SEC behavior • Decreased $\Delta G_{N\text{-}U}$, kJ/mol (48 ± 1.7 → 32 ± 4.2) • Decreased T_m, °C (78 ± 1.5 → 70 ± 0.9) • No impact on secondary structure	
V_L:			
REI VL	Y49F	• No impact on C_m (Gdn HCl)(1.55 M → 1.59 M) • No impact on $\Delta\Delta G_{unf}$(-0.2 kcal/mol) • Increased Congo red binding (mol/mol) (0.07 → 0.14)	[15]
	G57E	• Decreased C_m (Gdn HCl)(1.55 M → 1.30 M) • Increased $\Delta\Delta G_{unf}$(1.0 kcal/mol) • Increased Congo red binding (mol/mol) (0.07 → 0.32)	
	R61N	• Decreased C_m (Gdn HCl)(1.55 M → 0.75 M) • Increased $\Delta\Delta G_{unf}$(3.1 kcal/mol) • Increased Congo red binding (mol/mol) (0.07 → 0.46)	
	G68D	• Decreased C_m (Gdn HCl)(1.55 M → 1.22 M) • Increased $\Delta\Delta G_{unf}$(1.8 kcal/mol) • Increased Congo red binding (mol/mol) (0.07 → 0.20)	
	L78T	• Decreased C_m (Gdn HCl)(1.55 M → 1.43 M) • Increased $\Delta\Delta G_{unf}$(0.5 kcal/mol) • No impact on Congo red binding (mol/mol) (0.07 → 0.07)	
	A84T	• Decreased C_m (Gdn HCl)(1.55 M → 1.08 M) • Increased $\Delta\Delta G_{unf}$(2.1 kcal/mol) • Increased Congo red binding (mol/mol) (0.07 → 0.49)	
REI-VL	$T39K^{SS}$	• Increased ΔG_{unf}, kJ/mol (24.6 ± 0.4 →29.9 ± 0.5)	[16]
	$T39K^{(SH)2}$	• Decreased ΔG_{unf}, kJ/mol (24.6 ± 0.4 →12.1 ± 0.9)	
	Y32H^{SS}		

	Y32H $^{(SH)2}$	• Increased ΔG_{unf}, kJ/mol (24.6 ± 0.4 →30.2 ± 0.5)	
	F73L^{SS}	• Decreased ΔG_{unf}, kJ/mol (24.6 ± 0.4 →11.7 ± 1.3)	
	Y71F^{SS}	• Increased ΔG_{unf}, kJ/mol (24.6 ± 0.4 →28.7 ± 0.5)	
	Y71F $^{(SH)2}$	• Increased ΔG_{unf}, kJ/mol (24.6 ± 0.4 →30.0 ± 0.5) • Decreased ΔG_{unf}, kJ/mol (24.6 ± 0.4 →11.5 ± 1.8)	
	Y32H/T39K^{SS}	• Increased ΔG_{unf}, kJ/mol (24.6 ± 0.4 →35.5 ± 0.6)	
	Y32H/T39K $^{(SH)2}$	• Decreased ΔG_{unf}, kJ/mol (24.6 ± 0.4 →17.4 ± 0.5)	
	T39K/Y71F^{SS}	• Increased ΔG_{unf}, kJ/mol (24.6 ± 0.4 →36.0 ± 0.6)	
	T39K/Y71F $^{(SH)2}$	• Decreased ΔG_{unf}, kJ/mol (24.6 ± 0.4 →18.3 ± 0.9)	
	T39K/F73L^{SS}	• Increased ΔG_{unf}, kJ/mol (24.6 ± 0.4 →34.0 ± 0.8)	
	T39K/F73L $^{(SH)2}$	• Decreased ΔG_{unf}, kJ/mol (24.6 ± 0.4 →14.6 ± 0.9)	
	Y71F/F73L^{SS}	• Increased ΔG_{unf}, kJ/mol (24.6 ± 0.4 →34.5 ± 1.0)	
	Y71F/F73L$^{(SH)2}$	• Decreased ΔG_{unf}, kJ/mol (24.6 ± 0.4 →14.2 ± 0.9)	
	T39K/Y71F/F73L^{SS}	• Increased ΔG_{unf}, kJ/mol (24.6 ± 0.4 →40.6 ± 0.7)	
	T39K/Y71F/F73L $^{(SH)2}$	• Decreased ΔG_{unf}, kJ/mol (24.6 ± 0.4 →17.9 ± 1.0)	
	C23V/***Y32H***	• Increased soluble expression yield, mg/L (0 → 1) • Decreased ΔG_{unf}, kJ/mol (24.6 ± 0.4 →12.6 ± 0.5)	
	C23V/***Y32H***/T39K	• Decreased ΔG_{unf}, kJ/mol (24.6 ± 0.4 →13.5 ± 0.6)	
VL-LEN	***Y27dD/N29D***	• Increased C_m (Gdn HCl)(1.76 M → 2.42 M) • Increased ΔG_{unf}, kcal/mol (7.7 ± 0.4 →9.1 ± 0.5) • No fibril formation	[17]
	Y27dD	• Increased C_m (Gdn HCl)(1.76 M → 2.37 M) • Increased ΔG_{unf}, kcal/mol (7.7 ± 0.4 →9.4 ± 0.5) • No fibril formation	
	Y27dD/N29D/K30T	• Increased C_m (Gdn HCl)(1.76 M → 2.19 M) • Increased ΔG_{unf}, kcal/mol (7.7 ± 0.4 →8.3 ± 0.2) • No fibril formation	
	V27bL	• Increased C_m (Gdn HCl)(1.76 M → 2.11M) • Increased ΔG_{unf}, kcal/mol (7.7 ± 0.4 →9.5 ± 0.8) • No fibril formation	
	S97T	• Increased C_m (Gdn HCl)(1.76 M → 1.89 M) • No impact on ΔG_{unf}, kcal/mol (7.7 ± 0.4 →7.4 ± 0.5) • No fibril formation	
	Y96P	• Decreased C_m (Gdn HCl)(1.76 M → 1.49 M) • No impact on ΔG_{unf}, kcal/mol (7.7 ± 0.4 →6.1 ± 0.8) • Fibril formation	
	K30T	• Decreased C_m (Gdn HCl)(1.76 M → 1.45 M) • Decreased ΔG_{unf}, kcal/mol (7.7 ± 0.4 →6.4 ± 0.4) • Fibril formation	
	L15P	• Decreased C_m (Gdn HCl)(1.76 M → 1.37 M) • Decreased ΔG_{unf}, kcal/mol (7.7 ± 0.4 →6.2 ± 0.2) • Fibril formation	

	N28F	• Decreased C_m (Gdn HCl)(1.76 M → 1.37 M) • Decreased ΔG_{unf}, kcal/mol (7.7 ± 0.4 →4.6 ± 0.6) • Fibril formation	
	Len***CDR1$_{REC}$***	• Decreased C_m (Gdn HCl)(1.76 M → 1.09 M) • Decreased ΔG_{unf}, kcal/mol (7.7 ± 0.4 →3.4 ± 0.4) • Fibril formation	
	Len***CDR1$_{REC}$Y96P***	• Decreased C_m (Gdn HCl)(1.76 M → 0.92 M) • Decreased ΔG_{unf}, kcal/mol (7.7 ± 0.4 →5.1 ± 0.4) • No fibril formation	
	Len***CDR1$_{REC}$T30K***	• Decreased C_m (Gdn HCl)(1.76 M → 1.51 M) • Decreased ΔG_{unf}, kcal/mol (7.7 ± 0.4 →5.4 ± 0.1) • Fibril formation	
	S29N	• Increased C_m (Gdn HCl)(1.76 M → 1.98 M) • Decreased ΔG_{unf}, kcal/mol (7.7 ± 0.4 →7.5 ± 0.3) • No fibril formation	
	T94H	• Increased C_m (Gdn HCl)(1.76 M → 1.93 M) • Increased ΔG_{unf}, kcal/mol (7.7 ± 0.4 →8.4 ± 0.7) • No fibril formation	
	S97T	• Increased C_m (Gdn HCl)(1.76 M → 1.89 M) • Decreased ΔG_{unf}, kcal/mol (7.7 ± 0.4 →7.4 ± 0.5) • No fibril formation	
	I106L	• Increased C_m (Gdn HCl)(1.76 M → 1.80 M) • Decreased ΔG_{unf}, kcal/mol (7.7 ± 0.4 →6.5 ± 0.4) • No fibril formation	
	K30R	• Decreased C_m (Gdn HCl)(1.76 M → 1.74 M) • Decreased ΔG_{unf}, kcal/mol (7.7 ± 0.4 →6.7 ± 0.4) • No fibril formation	
	P40L	• Decreased C_m (Gdn HCl)(1.76 M → 1.59 M) • Decreased ΔG_{unf}, kcal/mol (7.7 ± 0.4 →6.0 ± 0.5) • Fibril formation	
	Q89H	• Decreased C_m (Gdn HCl)(1.76 M → 1.04 M) • Decreased ΔG_{unf}, kcal/mol (7.7 ± 0.4 →6.9 ± 0.9) • No fibril formation	
	Y96Q	• Decreased C_m (Gdn HCl)(1.76 M → 1.04 M) • Decreased ΔG_{unf}, kcal/mol (7.7 ± 0.4 →6.7 ± 0.5) • Fibril formation	
VL-REC	Rec***T30K***	• Decreased C_m (Gdn HCl)(1.76 M → 1.23 M) • Decreased ΔG_{unf}, kcal/mol (7.7 ± 0.4 →5.6 ± 0.4) • No fibril formation	
VL-Len	***Y27dD***	• Increased C_m (Gdn HCl), M (1.76 → 2.37) • Increased ΔG_{N-U}, kcal/mol (7.7 ± 0.6 → 9.4 ± 0.5)	[18]
	M4L/***Y27dD/T94H***	• Increased C_m (Gdn HCl), M (1.76 → 2.71) • Increased ΔG_{N-U}, kcal/mol (7.7 ± 0.6 → 12.1 ± 0.2)	
	Q38D	• Decreased C_m (Gdn HCl), M (1.76 → 1.3) • Decreased ΔG_{N-U}, kcal/mol (7.7 ± 0.6 → 3.8 ± 0.3)	
	Q38N	• Decreased C_m (Gdn HCl), M (1.76 → 1.51) • Decreased ΔG_{N-U}, kcal/mol (7.7 ± 0.6 → 5.6 ± 0.1)	
	Q38D/***Y27dD***	• Increased C_m (Gdn HCl), M (1.76 → 2.1) • No impact on ΔG_{N-U}, kcal/mol (7.7 ± 0.6 → 7.4 ± 0.4)	
		• Increased C_m (Gdn HCl), M (1.76 → 2.2)	

	Q38N/***Y27dD***	• No impact on ΔaG_{N-U}, kcal/mol (7.7 ± 0.6 → 7.6 ± 0.1)	
	Q89D/Y27dD	• Decreased C_m (Gdn HCl), M (1.76 → 1.17) • Decreased ΔG_{N-U}, kcal/mol (7.7 ± 0.6 → 4.5 ± 0.3)	
	Q89N/Y27dD	• Increased C_m (Gdn HCl), M (1.76 → 1.98) • Increased ΔG_{N-U}, kcal/mol (7.7 ± 0.6 → 7.9 ± 0.1)	
	Q89D/M4L/***Y27dD/T94H***	• Decreased C_m (Gdn HCl), M (1.76 → 1.48) • Decreased ΔG_{N-U}, kcal/mol (7.7 ± 0.6 → 6.5 ± 0.2)	
	Q89N/M4L/***Y27dD/T94H***	• Increased C_m (Gdn HCl), M (1.76 → 2.52) • Increased ΔG_{N-U}, kcal/mol (7.7 ± 0.6 → 11.5 ± 0.5)	
	Q89D	• Decreased C_m (Gdn HCl), M (1.76 → 0.55)	
	Q89N	• Decreased C_m (Gdn HCl), M (1.76 → 1.39) • Decreased ΔG_{N-U}, kcal/mol (7.7 ± 0.6 → 6.8 ± 0.3)	
DPK9 VL	***A50D***	• Increased ΔG_{N-U}, kJ/mol (28.9 → 29.9)	[11]
	S52D	• Decreased ΔG_{N-U}, kJ/mol (28.9 → 24.7)	
	S53D	• Decreased ΔG_{N-U}, kJ/mol (28.9 → 26.8)	
	50,52-53DDD	• Decreased ΔG_{N-U}, kJ/mol (28.9 → 24.0)	
	Single mutations: ***A50D*** ***S52D*** ***S53D*** ***A51D***	• Increased soluble expression yield, mg/L (47 → 78.6) • Increased % heat refolding (53.6 → 72.5) • No impact on Gel filtration elution volume, ml (13.6 → 13.3)	
	Double mutations: ***50/52DD*** ***50/53DD*** ***51/53DD*** ***52/53DD***	• Increased soluble expression yield, mg/L (47 → 103.7) • Increased % heat refolding (53.6 → 67.1) • No impact on Gel filtration elution volume, ml (13.6 → 13.0)	
	Triple mutations: ***50,52-53DDD***	• Increased soluble expression yield, mg/L (47 → 104.6) • Increased % heat refolding (53.6 → 91.4) • No impact on Gel filtration elution volume, ml (13.6 → 12.9)	
HVLP324	I48C/G64C	• Decreased soluble expression yield, mg/mL (2.5, 7 → 0.5, 1.1) • Increased T_m, °C (66.1 → 73.4) • Increased pepsin resistance (22% → 46%)	[58][1-3]
HVLP325	I48C/G64C	• Decreased soluble expression yield, mg/mL (0.5, 6.2 → 2.2, 3.1) • Increased T_m, °C (68.5 → 82.5) • Increased pepsin resistance (34% → 100%)	
HVLP335	I48C/G64C	• Decreased soluble expression yield, mg/mL (73.5 → 5.5) • Increased T_m, °C (61.7 → 79.0) • Increased pepsin resistance (0% → 62%)	
HVLP342	I48C/G64C	• Decreased soluble expression yield, mg/mL (1, 7.7 → 1.7, 10.8) • Increased T_m, °C (58.4→ 63.8) • Increased pepsin resistance (3% → 5%)	
HVLP351	I48C/G64C	• Decreased soluble expression yield, mg/mL (1.2, 8.9 → 1.9, 4.8) • Increased T_m, °C (62.0 → 71.9) • Increased pepsin resistance (31% → 94%)	
HVLP364	I48C/G64C	• Decreased soluble expression yield, mg/mL (0.3, 77 → 4.7) • Increased T_m, °C (57.0 → 72.3) • No impact on pepsin resistance (0% → 0%)	
HVLP3103	I48C/G64C	• Decreased soluble expression yield, mg/mL (0.7, 19 → 6.5) • Increased T_m, °C (65.7 → 76.4) • Increased pepsin resistance (0% → 89%)	

HVLP389	I48C/G64C	• Decreased soluble expression yield, mg/mL (3, 16.7 → 6.5) • Increased T_m, °C (51.9 → 66.3) Increased pepsin resistance (0% → 11%)	

[a]Residue numbering is as defined by Kabat *et al.* [19]. SEC, size exclusion chromatography; TRE, thermal refolding efficiency; CDR, complementarity determining region. CDR mutations are bolded and italicized, while the rest of the mutations are framework region (FR) sequence-related.

References

1. Davies, J. and L. Riechmann, *'Camelising' human antibody fragments: NMR studies on VH domains.* FEBS letters, 1994. **339**(3): p. 285-90.
2. Davies, J. and L. Riechmann, *Antibody VH domains as small recognition units.* Bio/technology, 1995. **13**(5): p. 475-9.
3. Davies, J. and L. Riechmann, *Single antibody domains as small recognition units: design and in vitro antigen selection of camelized, human VH domains with improved protein stability.* Protein engineering, 1996. **9**(6): p. 531-7.
4. Wirtz, P. and B. Steipe, *Intrabody construction and expression III: engineering hyperstable V(H) domains.* Protein science : a publication of the Protein Society, 1999. **8**(11): p. 2245-50.
5. Jespers, L., et al., *Aggregation-resistant domain antibodies selected on phage by heat denaturation.* Nature biotechnology, 2004. **22**(9): p. 1161-5.
6. Tanha, J., et al., *Improving solubility and refolding efficiency of human V(H)s by a novel mutational approach.* Protein Eng Des Sel, 2006. **19**(11): p. 503-9.
7. Barthelemy, P.A., et al., *Comprehensive analysis of the factors contributing to the stability and solubility of autonomous human VH domains.* The Journal of biological chemistry, 2008. **283**(6): p. 3639-54.
8. Famm, K., et al., *Thermodynamically stable aggregation-resistant antibody domains through directed evolution.* Journal of molecular biology, 2008. **376**(4): p. 926-31.
9. Perchiacca, J.M., M. Bhattacharya, and P.M. Tessier, *Mutational analysis of domain antibodies reveals aggregation hotspots within and near the complementarity determining regions.* Proteins, 2011. **79**(9): p. 2637-47.
10. Kim, D.Y., et al., *Disulfide linkage engineering for improving biophysical properties of human VH domains.* Protein Eng Des Sel, 2012. **25**(10): p. 581-9.
11. Dudgeon, K., et al., *General strategy for the generation of human antibody variable domains with increased aggregation resistance.* Proceedings of the National Academy of Sciences of the United States of America, 2012. **109**(27): p. 10879-84.
12. Perchiacca, J.M., et al., *Aggregation-resistant domain antibodies engineered with charged mutations near the edges of the complementarity-determining regions.* Protein engineering, design & selection : PEDS, 2012. **25**(10): p. 591-601.
13. Ladiwala, A.R., et al., *Rational design of potent domain antibody inhibitors of amyloid fibril assembly.* Proceedings of the National Academy of Sciences of the United States of America, 2012. **109**(49): p. 19965-70.
14. Perchiacca, J.M., C.C. Lee, and P.M. Tessier, *Optimal charged mutations in the complementarity-determining regions that prevent domain antibody aggregation are dependent on the antibody scaffold.* Protein engineering, design & selection : PEDS, 2014. **27**(2): p. 29-39.
15. Hurle, M.R., et al., *A role for destabilizing amino acid replacements in light-chain amyloidosis.* Proceedings of the National Academy of Sciences of the United States of America, 1994. **91**(12): p. 5446-50.

16. Frisch, C., et al., *Contribution of the intramolecular disulfide bridge to the folding stability of REIv, the variable domain of a human immunoglobulin kappa light chain.* Folding & design, 1996. **1**(6): p. 431-40.
17. Raffen, R., et al., *Physicochemical consequences of amino acid variations that contribute to fibril formation by immunoglobulin light chains.* Protein science : a publication of the Protein Society, 1999. **8**(3): p. 509-17.
18. Pokkuluri, P.R., et al., *Factors contributing to decreased protein stability when aspartic acid residues are in beta-sheet regions.* Protein science : a publication of the Protein Society, 2002. **11**(7): p. 1687-94.
19. Kabat, E., et al., *Sequences of Proteins of Immunological Interest . NIH Publication No. 91-3242. US Department of Health and Human Services.* Public Health Service, National Institutes of Health, Bethesda, MD, 1991.