Isolation of Monomeric Human VHs by a Phage Selection*

Human VH domains are promising molecules in applications involving antibodies, in particular, immunotherapy because of their human origin. However, they are, in general, prone to aggregation. Therefore, various strategies have been employed to acquire monomeric human VHs. We had previously discovered that filamentous phages displaying engineered monomeric VHs domains gave rise to significantly larger plaques on bacterial lawns than phages displaying wild type VHs with aggregation tendencies. Using plaque size as the selection criterion and a phage-displayed naïve human VH library we identified 15 VHs that were monomeric. Additionally, the VHs demonstrated good expression yields, good refolding properties following thermal denaturation, resistance to aggregation during long incubation at 37 °C, and to trypsin at 37 °C. These 15 VHs should serve as good scaffolds for developing immunotherapeutics, and the selection method employed here should have general utility for isolating proteins with desirable biophysical properties.

The immune repertoire of Camelidae is unique in that it possesses unusual types of antibodies referred to as heavy chain antibodies (1). These antibodies lack light chains and thus their combining sites consist of only one domain, termed V_{H}. Recombinant V_{H} single-domain antibodies (sdAbs) are comparable with their single-chain Fv (scFv) counterparts derived from conventional four-chain antibodies in terms of affinity, but outperform scFvs in terms of stability, resistance to aggregation, refolding properties, expression yield, and relative ease of DNA manipulation, library construction, and three-dimensional structural determination (2–6). Many of the aforementioned properties of V_{H} sdAbs are desired in applications involving antibodies. However, the non-human nature of V_{H}s limits their use in human immunotherapy because of the immunogenicity issue. In this respect, human V_{H} sdAbs are ideal candidates for therapeutic applications because they are expected to be least immunogenic. However, human V_{H}s are by and large prone to aggregation, a characteristic common to VHs derived from conventional antibodies (7–9). Thus, attempts were made previously to obtain monomers, i.e. human V_{H}s suitable for antibody applications. Such V_{H}s also displayed other useful properties typical of V_{H}s, such as high expression yield, high refoldability, and resistance to aggregation. Synthetic libraries built on these V_{H}s as library scaffolds should serve as a promising source of therapeutic proteins.

Camelization (7, 8) as well as lamination, which involves incorporating key solubilizing residues from cameldid sdAbs into human V_{H}s, have been employed to generate monomeric human V_{H}s. Synthetic sdAb libraries constructed based on these V_{H}s and generated by complementarity determining region (CDR) randomization were shown to yield binders to various antigens (8, 10). In another approach, fully monomeric human V_{H}s were isolated from human synthetic VH libraries without resorting to engineering of the sort mentioned above. In one experiment a monomeric human VH was discovered when a human V_{H} library was panned against hen egg white lysozyme (11). More recently, a selection method based on reversible unfolding and affinity criteria yielded many monomeric V_{H}s from synthetic human VH phage display libraries (12). This finding underlined the fact that an appropriate selection method is key to efficient capturing of rare human V_{H}s with desirable biophysical properties.

Here, we provide yet another approach for obtaining monomeric human V_{H}s. We report the isolation of 15 different V_{H}s originating from germlines DP-38, DP-47, V3-49, V3-53, YAC-5, and 8-1B from a phage-displayed naïve human VH repertoire by a selection method that is based on plaque size. These VHs, by and large, are also refoldable, retain their native fold following exposure to trypsin at 37 °C or long incubation at 37 °C, and are expressed in good yields in the Escherichia coli. When used as scaffolds, the diversity of the selected V_{H}s should allow for construction of more comprehensive libraries and provide flexibility in terms of choosing an optimal V_{H} scaffold for humanizing therapeutic camelid VHH binders. The current selection method permits high throughput identification of proteins with good biophysical properties by the naked eye, is very simple, eliminates affinity or stability selection steps, and is of general utility.

**MATERIALS AND METHODS**

**Phage Display Library Construction and Panning**—cDNA was synthesized from human spleen mRNA (Ambion Inc., Austin, TX) using random hexanucleotide primers and First Strand cDNA® kit (GE Healthcare, Baie d’Urqué, QC, Canada). Using the cDNA as template, V_{H} genes with flanking C_{H} sequences were amplified by polymerase chain reaction in nine separate reactions using VH framework region 1 (FR1)-specific primers and an immunoglobulin M-specific primer (13). The products were gel-purified and used as the template in the second round of PCR to construct V_{H} genes using the FR1- and FR4-specific primers (13) that also introduced flanking Apall and NotI restriction sites for cloning purposes. The resultant V_{H} repertoire DNA was cloned into fd-tetGIIIID phage vector and a V_{H} phage display library was constructed (8). Panning against protein A (GE Healthcare) was performed as described (8). Germline sequence assignment of the selected V_{H}s was...
performed using DNAPLOT software version 2.0.1 and V BASE version 1.0.4. Llama V<sub>λ</sub>H phage display library by panning against H11 scFv as described (5).

Protein Expression and Purification—Single-domain antibodies were cloned into pSf2 expression vector by standard cloning techniques (14). Periplasmic expression of sdAbs and subsequent purification by immobilized metal affinity chromatography were performed as described (15). Protein concentrations were determined by A<sub>280</sub> measurements using molar absorption coefficients calculated for each protein (16). Gel filtration chromatography of the purified sdAbs was performed on a Superdex 75 column (GE Healthcare) as described (17).

Binding and Refolding Efficiency Experiments—Equilibrium dissociation constants (K<sub>D</sub>) and refolding efficiencies (REs) of V<sub>λ</sub>S/V<sub>λ</sub>Hs were derived from surface plasmon resonance (SPR) data collected with the BIACORE 3000 biosensor system (Biacore Inc., Piscataway, NJ). To measure the binding of V<sub>λ</sub>S to protein A, 2000 resonance units of protein A or a reference antigen-binding fragment (Fab) were immobilized on research grade CM5 sensor chips (Biacore Inc.). Immunobilizations were carried out at concentrations of 25 μg/ml (protein A) or 50 μg/ml (Fab) in 10 mM sodium acetate buffer, pH 4.5, using the amine coupling kit provided by the manufacturer. To measure the binding of the anti-idiotypic llama V<sub>λ</sub>S to H11 scFv (18), 4100 resonance units of 50 μg/ml H11 scFv or 3000 resonance units of 10 μg/ml Se155-4 immunoglobulin G (19) were immobilized as described above. In all instances, analyses were carried out at 25 °C in 10 mM HEPES, pH 7.4, containing 150 mM NaCl, 3 mM EDTA, and 0.005% P20 surfactant at a flow rate of 60 μl/min. Digestion reactions were carried out in a total volume of 1.0.4 Llama V<sub>λ</sub>Hs H11C7, H11F9, and H11B2 were isolated from a llama go fVH in 100 mM Tris-HCl, pH 8.0, or V<sub>λ</sub>Ss were denatured by incubation at 85 °C for 20 min at 10 μg/ml concentrations. The protein samples were then cooled down to room temperature for 30 min to refold and were subsequently centrifuged in a microcentrifuge at 14,000 × g for 5 min at room temperature to remove any protein precipitates. The supernatants were recovered and analyzed for binding activity by SPR as described above. For both folded and refolded protein, data were fit to a 1:1 interaction model simultaneously using BIAevaluation 4.1 software (Biacore Inc.) and K<sub>D</sub> values were subsequently determined. REs were determined from RE = (K<sub>P</sub>/K<sub>R</sub>)ref × 100, where K<sub>P</sub>ref is the K<sub>P</sub> of the native protein and K<sub>R</sub>ref is the K<sub>R</sub> of the refolded protein.

Tryptic Digest Experiments—Three μl of a freshly prepared 0.1 μg/μl sequencing grade trypsin (Hoffmann-La Roche Ltd., Mississauga, ON, Canada) in 1 mM HCl was added to 60 μg of V<sub>λ</sub>S in 100 mM Tris-HCl buffer, pH 7.8. Digestion reactions were carried out in a total volume of 60 μl for 1 h at 37 °C and stopped by adding 5 μl of 0.1 μg/μl trypsin inhibitor (Sigma). Following completion of digestion, 5 μl was removed and analyzed by SDS-PAGE; the remaining was desalted using ZipTip<sub>C18</sub> (Millipore, Ontario, Canada), eluted with 1% acetic acid in 50:50 methanol:water and subjected to V<sub>λ</sub>S mass determination by matrix-assisted laser desorption ionization mass spectrometry.

Protein Stability Experiments at 37 °C—Single-domain antibodies at 0.32–3.2 mg/ml concentrations were incubated at 37 °C in phosphate-buffered saline buffer for 17 days. Following incubation, the protein samples were spun down in a microcentrifuge at maximum speed for 5 min even in the absence of any visible aggregate formation. The samples were then applied onto a Superdex 75 size exclusion column and the monomeric peaks were collected for SPR analysis of binding to protein A. SPR analyses were performed as described above except that 500 resonance units of protein A or reference Fab was immobilized and immobilizations were carried out at a concentration of 50 μg/ml.

NMR Experiments—V<sub>λ</sub>S samples for NMR analysis were dissolved in 10 mM sodium phosphate, 150 mM NaCl, 0.5 mM EDTA, and 0.02% NaN<sub>3</sub> at pH 7.0. The protein concentrations were 40 μM to 1.0 mM. All NMR experiments were carried out at 298 K on a Bruker Avance-800 or a Bruker Avance-500 NMR spectrometer. One-dimensional 1H NMR spectra were recorded with 16,384 data points and the spectral widths were 8,992.81 Hz at 500 MHz and 17,605.63 Hz at 800 MHz, respectively. Two-dimensional 1H-1H NOESY spectra of 2,048 × 400 data points were acquired on a Bruker Avance-800 NMR spectrometer with a spectral width of 11,990.04 Hz and a mixing time of 120 ms. In all NMR experiments, water suppression was achieved using the WATERGATE method implemented through the 3–9–19 pulse train (20, 21). NMR data were processed and analyzed using the Bruker XWINNMR software package. All PFG-NMR diffusion measurements were carried out with the water-suppressed LED sequence (22), on a Bruker Avance-500 NMR spectrometer equipped with a triple-resonance probe with three-axis gradients. One-dimensional proton spectra were processed and analyzed using Bruker XWINNMR software package. NMR signal intensities were obtained by integrating NMR spectra in the methyl and methylene proton region (2.3 to −0.3 ppm) where all NMR signals were attenuated uniformly at all given PFG strengths.

RESULTS

During the course of the construction of fully human and laminated human V<sub>λ</sub>S libraries, we learned that the phages displaying monomeric laminated V<sub>λ</sub>S formed larger plaques on bacterial lawns than phages displaying fully human V<sub>λ</sub>S with aggregation tendencies. We thus decided to use plaque size as a means of identifying rare, naturally occurring monomer V<sub>λ</sub>Ss from the human V<sub>λ</sub>S repertoire (Fig. 1). To this end, a phage library displaying human V<sub>λ</sub>Ss with a size of 6 × 10<sup>6</sup> was constructed and propagated as plaques on agar plates. On the titer plates, the library consisted essentially of small plaques interspersed with some large ones. PCR on 20 clones revealed that the small plaques corresponded to the V<sub>λ</sub>S-displaying phages, whereas the large ones represented the wild type phages, i.e. phages lacking V<sub>λ</sub>S sequence inserts. None of the V<sub>λ</sub>S-displaying phages were found with large plaque morphology. This was not unexpected because of the paucity of the monomeric V<sub>λ</sub>S in the human repertoire and the large size of the library. To facilitate the identification of monomeric V<sub>λ</sub>Ss, it was decided to reduce the library size to a manageable one and remove interfering wild type phage with large plaque-size morphology by panning the library against protein A, which binds to a subset of human V<sub>λ</sub>S from the V<sub>λ</sub>3 family.

Following a few rounds of panning, the library became enriched for phage producing large plaques, and PCR and sequencing of more than 110 such plaques showed that all had complete V<sub>λ</sub>S open reading frames. The size of the large plaques that were picked for analysis is represented in Fig. 1. Sequencing revealed 15 different V<sub>λ</sub>Ss that belonged to the V<sub>λ</sub>3 family and utilized DP-38, DP-47, V3-49, V3-53, YAC-5, or 8-1B germline V segments (TABLE ONE; Fig. 2). The DP-38 and DP-47 germline sequences have been previously implicated in protein A binding (12, 23). In addition, all V<sub>λ</sub>Ss had a Thr residue at position 57 (Fig. 2), consistent with their protein A binding activity (24, 25). The most frequently utilized germline V segment was DP-47 with a relative frequency of 46% (Fig. 2). The V<sub>λ</sub>3 CDR3 lengths ranged from 4 amino acids for HVHB82 to 16 amino acids for HVHP430, with HVHP430...
having a pair of Cys residues in CDR3. Amino acid mutations with respect to the parental germline V segment (residues 1–94) and FR4 (residues 103–113) sequences were observed in all V_{H\delta} and ranged from two mutations for HVHP44 (L5V and Q105R) and HVHB82 (E1Q and L5Q) to 16 mutations for HVHP426 (TABLE ONE). Mutations were concentrated in the V segments; only two mutations were detected in all the 15 FR4s, at positions 105 and 108. HVHP44 and HVHB82 differed from other V_{H\delta} in that they both had a positively charged amino acid at position 105 instead of a Gln (TABLE ONE, Fig. 2). However, whereas the positively charged amino acid in HVHP44 was acquired by mutation, the one in HVHB82 was germline-encoded. Except for HVHP423 and HVHP44B, the remaining V_{H\delta} had the germline residues at the key solubility positions (4): 37V/44G/45L/47W or 37F/44G/45L/47W (HVHP428), HVHP423 and HVHP44B had a V37F mutation. Mutations at other positions, which have been shown or hypothesized to be important in V_{H} solubility, included seven E6Q, three S35T/H, one R83G, one K83R, one A84P, one T84A, and one M108L mutation (5, 11, 26). Frequent mutations were also observed at positions 1 and 5 that included 11 E1Q, eight L5V/Q, and one V5Q mutations.

All V_{H\delta} except HVHP44B, which was essentially the same as HVHP423, were expressed in 1-liter culture volumes in E. coli strain TG1 in fusion with a c-Myc-His_{6} tag and purified to homogeneity from periplasmic extracts by immobilized metal affinity chromatography. The expression yields ranged from 1.8 to 62.1 mg of purified protein per liter of bacterial culture in shake flasks with the majority having yields of several milligrams (TABLE TWO). In the instances of HVHP423 and HVHP430, another trial under "apparently" the same expression conditions gave yields of 2.4 and 6.4 mg as opposed to 62.1 and 23.7 mg, respectively. This implies that for many of the V_{H\delta} described here, optimal expression conditions should be achieved, without much effort, resulting in expression yields significantly higher than the values reported in TABLE TWO. As expected, all the V_{H\delta} bound to protein A in SPR analyses, with K_{s} values of 0.2–3 μM, a range and magnitude comparable with affinities reported previously for llama V_{H}H variants with protein A binding activity (24). None of the V_{H\delta} bound to the Fab reference surface.

The aggregation tendency of the human V_{H\delta} was assessed in terms of their oligomerization state by gel filtration chromatography and NMR (TABLE TWO). All V_{H\delta} were subjected to Superdex 75 gel filtration chromatography. Similar to a llama V_{H}H, H11C7, all V_{H\delta} gave a symmetric single peak at the elution volume expected for a monomer, and were essentially free of any aggregates (see the example for HVHP428 in Fig. 3A). In contrast, a typical human V_{H} (i.e. BT32/A6) formed a considerable amount of aggregates. For three of the V_{H\delta}, a minor peak with a mobility expected for a V_{H} dimer was also observed. SPR analyses of the minor peaks gave off-rate values that were significantly slower than those for the monomer V_{H\delta}, consistent with them being V_{H} dimers. The dimer peak was also observed in the case of the llama V_{H}H, H11C7. The folding and oligomerization states of the V_{H\delta} at high concentrations were further studied by NMR spectroscopy. As shown in TABLE TWO, all the V_{H} proteins studied appeared to be relatively soluble and assumed a well folded three-dimensional structure. One-dimensional NMR spectra of the V_{H} fragments (Fig. 3B) showed structure folds characteristic of V_{H} domains. The state of protein aggregation was also assessed by use of an PFGE-NMR diffusion experiment for the HVHP414 fragment and two isoforms, VH14 and VH14-cMyc, with and without the c-Myc sequence, of the HVHP414. VH14 is a modified version of HVHP414 with a c-Myc N132E mutation and with an additional methionine residue at the N terminus. In brief, the PFGE-NMR data (not shown) indicated that all the protein samples had expected monomeric molecular weights even at the relatively high protein concentrations used for NMR experiments.

We further investigated the stability of the V_{H\delta} in terms of their resistance to trypsin at 37 °C and integrity following long incubations at 37 °C. Trypsin cleaves polypeptide amide backbones at the C terminus of an Arg and a Lys residue. There are 9–13 Arg and Lys residues in the human V_{H\delta} (Fig. 2). There is also an additional Lys residue in the C-terminal c-Myc tag (27), which is susceptible to digestion by trypsin. Fig. 4A is an SDS-PAGE analysis of HVHP414 during trypsin digestion. Within 1 h the original band was completely converted to a single product that had a mobility expected for the V_{H} without the c-Myc-His_{6} tag. The same result was obtained for 12 other V_{H\delta} following a 1-h incubation with trypsin. Mass spectrometry on a randomly selected sample of the trypsin-treated V_{H\delta} (HVHP414, HVHP419, HVHP420, HVHP423, HVHP429, HVHP430, and HVHM81) confirmed that in every case the molecular mass of the digested product corresponded to a V_{H} with the c-Myc Lys as the C-terminal residue (Fig. 4B). HVHM41 gave a significantly shorter fragment than the rest upon digestion, and in this case mass spectrometry experiments mapped the cleavage site to Arg^{29} in CDR3 (data not shown).

**FIGURE 1.** The contrast in plaque size between phages displaying a soluble V_{H} obtained in this study (HVHP428), and those displaying an insoluble V_{H} (BT32/A6). The photo depicts a part of the bacterial lawn agar plate that was magnified to enhance plaque visualization. Although the plate contained an equal number of each of the two plaque types, the photo essentially shows only the large, HVHP428 plaques. The majority of the BT32/A6 plaques were too small to produce clear, well-defined images in this photo. The plaques marked by arrows, thus, represent a minor proportion of BT32/A6 phages that were large enough to be visible in this image. Asterisks mark representative plaque sizes for HVHP428 phages. The identities of plaques were determined by DNA digestion, and in this case mass spectrometry experiments mapped the NMR spectra of the VH fragments (Fig. 3). The contrast in plaque size between phages displaying a soluble VH, i.e., BT32/A6 phages that were large enough to be visible in this image, and those displaying an insoluble VH (BT32/A6).
Monomeric Human V_{H}\text{S}

TABLE ONE

| V_{H} | V/J germline | Amino acid deviation from V and FR4 germline sequences |
|-------|--------------|------------------------------------------------------|
| HVHP44 | DP47/JH4b | L5V, Q105R |
| HVHP82 | DP47/JH6c | E1Q, L5Q |
| HVHP421 | DP47/JH4b | E1Q, V2L, L5Q, L11V, G16R |
| HVHP419 | DP47/JH4b | E1Q, V2L, L5Q, T77S, R83G, K94R |
| HVHP430 | DP47/JH3b | E1Q, L5V, V12L, Q13K, S31N, G52AS, L78V, A93V, K94R |
| HVHP429 | DP47/JH4 | L5V, G10T, S30I, S31N, G42D, E46D, A50T, G52aN, S33N, S56A K75N, A84P, E85D |
| HVHM41 | DP47/JH3a | E1Q, L5V, E6Q, G16R, T28A, S53G, G55D, S56H, M108L |
| HVHM81 | DP47/JH3a | L5V, E6Q, G16R, S30D, S31D, S35H, A50D, G55A, E85G, V89L, K94R |
| HVHP428 | V3–49/JH4b | E1Q, V2L, V5Q, R16G, T60A, G73D, K83R, T84A, V89M, T93A |
| HVHP420 | DP–38/JH4b | E1Q, S35T, S52A |
| HVHP414 | DP–38/JH3b | E1D, E6Q, A23T, T28P, K52T, A60V |
| HVHP423 | V3–53/JH1 | E1Q, V2M, E6Q, L11V, L12V, N32S, Y33R, V37F, K43M, K64R, T68S, V89L |
| HVHP44B | V3–53/JH1 | E1Q, E6Q, N32S, Y33R, V37F, K43M, Y58S, K64R, T68S, V89L |
| HVHP413 | YAC–5/JH3b | E1Q, E6Q, Q105R, V92F, S31D, N32Y, V89L |
| HVHP426 | 8–1B/JH3b | E1Q, E6Q, L11V, G16R, T28I, S30D, N32A, Y33R, V37F, K43M, T68S, V89M, K94R |

Eleven V_{H}\text{s} ranging in concentration from 0.32 (HVHP428) to 3.2 (HVHP420) mg/ml were incubated at 37 °C for 17 days. Their stability was subsequently determined in terms of oligomerization state and protein A binding. As shown by gel filtration chromatography, treatment of VHs did not affect their native fold. The possibility that the VHs were resistant to trypsin at 37 °C (see above), a property typically associated with well folded native proteins.

To ensure that the VHs maintained their native fold following 37 °C treatment, two VHs, namely, HVHP414 (1.2 mg/ml) and HVHP420, were selected at random and the KD values for binding to protein A were determined by SPR (data shown for HVHP420, Fig. 4C, inset) and compared with the KD values obtained for untreated VHs (TABLE TWO). The calculated KD values for the treated VHs were 1.4 and 1.0 μM for HVHP414 and HVHP420, respectively. These values are essentially identical to the values for the corresponding untreated VHs (TABLE TWO), demonstrating that 37 °C treatment of VHs did not affect their native fold. The possibility that the VHs may have been in a less compact, non-native fold during the 37 °C incubation periods and resumed their native fold upon return to room temperature during gel filtration and SPR experiments is unlikely in light of the fact that the VHs were resistant to trypsin at 37 °C (see above), a property typically associated with well folded native proteins.

We also investigated the RE of the human VHs by comparing the KD values for the binding of the native (K_{D, n}) and heat-treated, refolded (K_{D, ref}) VHs to protein A (5). If a fraction of the VHs is inactivated by heat treatment, the measured K_{D, n} would increase, because this parameter is based on the concentration of folded, i.e. active, antibody fragment. Thus, the ratio of K_{D, n}/K_{D, ref} gives a measure of the VH RE. Fig. 5 compares sensorgrams for HVHP423 binding to immobilized protein A in native (thick lines) and refolded (thin lines) states at several selected VH concentrations. As can be seen, binding of the refolded VH to protein A is less in all instances, indicating that the unfolding is not fully reversible. For each of the 14 VHs, protein A binding in both native and refolded states was measured at several concentrations and the KD values, and subsequently the REs, were determined (TABLE TWO, K_{D, ref} values are not shown). The KD values and REs of two anti-idiotypic llama V_{H}\text{s}, H11F9 and H11B2, which were used as references, were also determined. Four of the 14 VHs had REs in the range of 92–95% and were similar to the REs for H11F9 and H11B2, which were 95 and 100%, respectively. Another five had REs in the range of 84–88% and three were over 70%. Only two had significantly lower REs, HVHP413 (52%) and HVHP421 (14%). Several llama V_{H}\text{s} examined previously had REs of ~50% (2).

DISCUSSION

Synthetic fully human V_{H} libraries, which are made by CDR randomization on human V_{H} scaffolds, are a promising source of immunotherapeutics. However, suitable, i.e. monomeric, human V_{H} scaffolds are not readily available and thus, acquiring them constitutes the first step toward constructing synthetic fully human V_{H} libraries. Here we isolated 15 monomeric human V_{H}s with different germline and overall
Monomeric Human \( V_{H} \)s

Table Two

| VH/VHH | Expression* | \( K_{D} \) | Trypsin resistance | RE | Oligomerization state (by GFC) | Folding and Oligomerization State (by NMR) |
|--------|-------------|------------|--------------------|----|------------------------------|----------------------------------------|
|        | mg          | \( \mu \text{M} \) |                   |    |                             |                                        |
| HVHP44 | 8.2         | 1.3        | √                  | 93 | Monomer                     | s/f/ps\(^{a}\)                            |
| HVHB82 | 5.9         | 0.2        | √                  | 71 | Monomer                     | s/f/ps\(^{a}\)                            |
| HVHP421| 5.5         | 1.0        | √                  | 14 | Monomer                     | s/f/ps\(^{a}\)                            |
| HVHP419| 3.4         | 1.6        | √                  | 84 | Monomer                     | ND                                     |
| HVHP430| 6.4, 23.7   | 2.3        | √                  | 88 | Monomer                     | s/f/ps\(^{a}\)                            |
| HVHP429| 3.4         | 1.3        | √                  | 86 | Monomer                     | s/f/ps\(^{a}\)                            |
| HVHM41 | 1.8         | 0.5        | X                  | 92 | Monomer                     | ND                                     |
| HVHM81 | 4.3         | 1.3        | √                  | 87 | Monomer                     | ND                                     |
| HVHP428| 3.1         | 1.8        | √                  | 95 | Monomer                     | s/f/ps\(^{a}\)                            |
| HVHP420| 59.0        | 1.2        | √                  | 92 | Monomer                     | ND                                     |
| HVHP414| 11.8        | 1.6        | √                  | 73 | Monomer                     | s/f/ps\(^{a}\)                            |
| HVHP423| 2.4, 62.1   | 3.0        | √                  | 86 | Monomer                     | s/f/ps\(^{a}\)                            |
| HVHP413| 5.8         | 0.3        | √                  | 52 | Monomer                     | s/f/ps\(^{a}\)                            |
| HVHP426| 6.3         | 0.8        | √                  | 70 | Monomer                     | s/f/ps\(^{a}\)                            |
| H11F9′ | ND\(^{f}\)  | 3.5        | ND                 | 95 | Monomer                     | ND                                     |
| H11B2′ | ND\(^{f}\)  | 2.0        | ND                 | 100| Monomer                     | ND                                     |

\(^{a}\) Expression yield per liter of bacterial culture.
\(^{b}\) GFC, gel filtration chromatography.
\(^{c}\) The solution properties of the various \( V_{H} \) molecules are characterized as soluble (s) and structurally folded (f) with broad (b), sharp (s), and partially sharp (ps) NMR spectra.
\(^{d}\) Folding was determined by one-dimensional \(^{1}\)H spectra recorded at 500 MHz.
\(^{e}\) Folding was confirmed by one-dimensional \(^{1}\)H and two-dimensional \(^{1}\)H-\(^{1}\)H NOESY spectra recorded at 800 MHz. No concentration dependent line broadenings were observed from \(~30\) to \(300\ \mu \text{M} \).
\(^{f}\) ND, not determined.
\(^{g}\) Sharp resonances at a concentration of \(40\ \mu \text{M} \) and partial line-broadening were observed at \(\sim1.0\ \text{mM} \).
\(^{h}\) Folding was confirmed by one-dimensional \(^{1}\)H spectrum recorded at 800 MHz.

\( V_{H} \) sequences from a phage-displayed naive human \( V_{H} \) repertoire by a selection method that is based on phage plaque size. The \( V_{H} \)s remain functional and monomeric following trypsin treatment at \(37^\circ \text{C} \), weeks of incubations at \(37^\circ \text{C} \) or months of storage at \(4^\circ \text{C} \), have high thermal refolding efficiencies, are produced in good yields in \( E. \ coli \), and possess protein A binding activity. Such properties would also be manifested by these libraries to situations where \( V_{H} \) binders are required to maintain characteristics would further extend the biotechnological application of and cost-effective production. High thermal refolding efficiency characteristics would further extend the biotechnological application of these libraries to situations where \( V_{H} \) binders are required to maintain their activity after transient exposure to high temperatures. The \( V_{H} \)s should also be very suitable for intrabody applications because of their desirable biophysical properties. The protein A binding property would simplify \( V_{H} \) purification and detection in biotechnological application, immunoblotting and immunocytochemistry, and can be exploited to enhance library performance by removing nonfunctional \( V_{H} \)s from the libraries (28).

Previously reported fully human \( V_{H} \)s with favorable biophysical properties were based on one \( V \) germline sequence, DP-47 (11, 12). The observation that the monomeric human \( V_{H} \)s in this study stem from six different germline sequences, including DP-47, demonstrates that stable \( V_{H} \)s are not restricted in terms of germline gene usage. In fact, it is very likely that we would have isolated monomeric \( V_{H} \)s with family and germline origins different from the ones we describe here had we not restricted our selection to a subset of \( V_{H} \)3 family \( V_{H} \)s with protein A binding activity. The appearance of DP-47 germline in over 50% of the \( V_{H} \)s may be because of its over-representation in the expressed \( V_{H} \) repertoire (29). It is not possible to pinpoint amino acid mutations (TABLE ONE) responsible for the observed biophysical behavior of the present \( V_{H} \)s because of the occurrence of multiple mutations in the \( V_{H} \)s and the fact that CDR3 is also known to be involved in shaping the biophysical profiles of sdAbs (30–32). It may be, however, that mutations at positions known to be important for sdAb stability and solubility, e.g. V37F in HVHP423 and HVHP44B, or mutations occurring multiple times at the same position, e.g. L5V/Q and V5Q in nine \( V_{H} \)s, have a role in determining \( V_{H} \) biophysical properties. In terms of library construction, it would be desirable that the monomericity of the present \( V_{H} \)s not be dependent on CDRs, in particular CDR3, so that CDR randomization be performed without the worry of jeopardizing library stability. In this regard, the \( V_{H} \)s with smaller CDR3s, e.g. HVHB82, may be preferred scaffolds because there would be less dependence on CDR3 for stability (30, 31).

The diversity of the present \( V_{H} \)s in terms of overall sequence and CDR3 length should allow the construction of better performing libraries. Synthetic \( V_{H} \) libraries have been constructed on single scaffolds (8, 10, 12, 33). Such an approach to repertoire generation is in sharp contrast to the natural, \textit{in vivo} “approach” that utilizes a multiplicity of scaffolds. Based on the sequences reported here one can take advantage of the availability of the diverse set of \( V_{H} \)s to create libraries that are based on multiple \( V_{H} \) scaffolds. Such libraries would be a better emulation of \textit{in vivo} repertoires and, therefore, would have a more optimal complexity. Of the three CDRs in sdAbs, CDR3 generally contributes most significantly to repertoire diversity and for this reason CDR3 randomization has always been included in library construction strategies. CDR3 randomizations on sdAb scaffolds are typically accompanied by concomitant varying CDR3 lengths. Whereas this significantly improves library complexity, it may also compromise library stability by disrupting the length of the parental scaffold CDR3 (5). The heterogeneity of our \( V_{H} \)s in terms of CDR3 length will permit us to create librar-
ies with both good complexity and good stability. Such a library would consist of sublibraries, where each sublibrary is created by CDR3 randomization (and CDR1 and/or CDR2 randomization, if desired) on a single V_{H} scaffold without disrupting the parental CDR3 length. The versatility of the present V_{H}s is also beneficial in terms of choosing an optimal V_{H} framework for humanizing well characterized camelid V_{HH} binders against therapeutic targets (34–36). High affinity camelid V_{HH}s against therapeutic targets can be obtained from immune V_{HH} libraries with relative ease (4, 37) and be subsequently subjected to humanization to remove possible V_{HH} immunogenicity, hence providing an alternative to human V_{H} library approach for the production of therapeutic V_{H}s. Generating high affinity therapeutic V_{H}s by the latter approach may often require additional tedious and time consuming in vitro affinity maturation of lead binder(s) selected from the primary synthetic human V_{H} libraries.

A number of evolutionary approaches for the selection of proteins with improved biophysical properties have been described (12, 38–41). Typically, stability pressure is required to ensure preferential selection of stable variants over unstable or less stable ones from a library population. For example, in a related study, heat treatment of V_{H} phage display libraries was required to select for aggregation resistant V_{H}s (12). Examples of evolutionary selection approaches involving phage...
FIGURE 4. Stability of the human V_{H}s in terms of their resistance to trypsin and integrity following long incubation at 37 °C. A, SDS-PAGE comparing the mobilities of the untreated and trypsin-treated HVHP414 V_{H} at 15, 30, and 60 min relative to a 21-kDa marker. HVHP414-cMyc denotes HVHP414 V_{H} lacking the c-Myc. B, molecular mass profiles obtained by mass spectrometry of untreated and trypsin-treated (60 min) HVHP414 V_{H}. The mass spectrometry profile of the treated V_{H} is superimposed onto that for the untreated one to provide a better visual comparison. The experimental molecular mass of the untreated V_{H} is 14,967.6 Da, which is essentially identical to the expected molecular mass, 14,967.7 Da. The observed molecular mass of the trypsin-treated V_{H} (13,368.5 Da) indicates loss of 13 amino acids at the C terminus by cleavage at Lys in the c-Myc tag to give an expected molecular mass of 13,368.0 Da. The trypsin cleavage site is shown by a vertical arrow above the amino acids sequence of HVHP414. C, gel filtration chromatograms comparing the oligomerization state of the 37 °C-treated HVHP420 V_{H} (upper profile) to that of untreated V_{H} (lower profile). The chromatograms were shifted vertically because they were indistinguishable when superimposed. The major and minor peaks in each chromatogram correspond to monomeric and dimeric V_{H}s, respectively. The dimeric V_{H} constitutes 3% of the total protein. The inset shows the sensorgram overlays for the binding of 37 °C-treated HVHP420 to protein A at various concentrations. The V_{H}s used for temperature stability studies were from stocks that had already been at 4 °C for several months.
display include conventional phage display, selectively infective phage, and the proteolysis approaches. In the first two approaches affinity selection is used to select stable species from a library, based on the assumption that stable proteins possess better binding properties for their ligand than unstable ones. However, even with the additional inclusion of a stability selection step, these approaches may primarily enrich for higher affinity rather than for higher stability (39). A binding step requirement also limits the applicability of these approaches to proteins with known ligands. The third, proteolysis selection, approach is based on the fact that stable proteins are generally compact and therefore are resistant to proteases, whereas unstable ones are not. The phage display format is engineered in such a way that the protease stability of the displayed protein translates to phage infectivity. Thus, when a variant phage display library is treated with a protease, only the phages displaying stable proteins retain their infectivity and can subsequently be selected by infecting an *E. coli* host. Because this approach is independent of ligand binding, it has general utility. However, even stable and well folded proteins have protease-sensitive sites, e.g. loops and linkers, and this could sometimes hinder the selection of stable species in a proteolysis approach (42).

In the present evolutionary approach, proteins with superior biophysical properties are simply identified by the naked eye. The approach does not require ligand binding, proteolysis, or destabilization steps and, thus, avoids complications that may be encountered in previous reported selection approaches. No requirement for a binding step also means that our approach has general utility. As an option, a binding step may be included to ensure that the selected proteins are functional. However, the dependence of the present approach on plating (for plaque visualization) introduces a possible logistical limitation in terms of the number of plates that can be handled and thus limits its application to smaller libraries. Nonetheless, the utility of the current approach can be extended to large libraries, if the library size is first reduced by, for example, incorporating a step that removes a large population of unstable species, e.g. library adsorption on a protein A surface as described here, or on a hydrodynamic interaction column to remove poorly folded proteins (41). Here, our approach was used to select *V_{H}* of good biophysical properties from a background of very unstable *V_{H}*s. However, it may be more difficult to select the “best” species from a mutant library that is populated with proteins with reasonably good stabilities. In this case, the best variants may be identified based on the rate of plaque formation by using shorter incubation times, or plaque size and frequency criteria.

The present selection approach can be extended to the identification of stable and well folded antibody fragments such as human *V_{H}*s, scFvs, Fabs, with the optional inclusion in the selection system of a binding step involving protein L, protein A, or other ligands, and non-antibody scaffolds. Moreover, the observed correlation between phage plaque size and *V_{H}* expression yield means that one can utilize the present approach for acquiring high-expressing versions of proteins with poor or unsatisfactory expression from mutant phage display libraries. This application would be particularly appealing where boosting expression of therapeutic proteins or expensive poor-expressing protein reagents would significantly reduce protein production cost.

Acknowledgments—We thank Sonia Leclere for oligonucleotide synthesis and DNA sequencing and Lisa Morrison and Kenneth H. N. Chan for performing mass spectrometry experiments. We are indebted to Shakeeb Waseh and Andy Ng for helping with protein expression.

REFERENCES

1. Hamers-Casterman C., Aderhouch, T., Muyldermans, S., Robinson, G., Hamers, C., Songa, E. B., Bendaianman, N., and Hamers, R. (1993) Nature 363, 446–448
2. van der Linden, R. H., Frenken, L. G., de Geus, B., Harmse, M. M., Ruuls, R. C., Stok, W., de Ron, L., Wilson, S., Davis, P., and Verrips, C. T. (1999) Biochim. Biophys. Acta 1431, 37–46
3. Frenken, L. G., van der Linden, R. H., Hermans, P. W., Bos, J. W., Ruuls, R. C., de Geus, B., and Verrips, C. T. (2000) J. Biotechnol. 78, 11–21
4. Nguyen, V. K., Desmyter, A., and Muyldermans, S. (2001) Adv. Immunol. 79, 261–296
5. Tanha, J., Dubuc, G., Hiraama, T., Narang, S. A., and MacKenzie, C. R. (2002) J. Immunol. Methods 263, 97–109
6. Zhang, J., Tanha, J., Hiraama, T., Khieu, N. H., To, R., Tong-Sevin, H., Stone, E., Brisson, J. R., and MacKenzie, C. R. (2004) J. Mol. Biol. 335, 49–56
7. Davies, J., and Riechmann, L. (1994) FASEB J. 8, 285–290
8. Tanha, J., Xu, P., Chen, Z. G., Ni, F., Kaplan, H., Narang, S. A., and MacKenzie, C. R. (2001) J. Biol. Chem. 276, 24774–24780
9. Ward, E. S., Gussow, D., Griffiths, A. D., Jones, P. T., and Winter, G. (1989) Nature 341, 544–546
10. Davies, J., and Riechmann, L. (1995) Biotechnology 13, 475–479
11. Jespers, L., Schon, O., James, L. C., Veprintsev, D., and Winter, G. (2004) J. Mol. Biol. 337, 893–903
12. Jespers, L., Schon, O., Famm, K., and Winter, G. (2004) Nat. Biotechnol. 22, 1161–1165
13. de Haard, H. J., van Nee, N., Reurs, A., Hufton, S. E., Roovers, R. C., Henderiks, P., de Bruine, A. P., Arends, J. W., and Hoogenboom, H. R. (1999) J. Biol. Chem. 274, 18218–18230
14. Sambrook, J., Fritsch, E. F., and Maniatis, T. (1989) *Molecular Cloning: A Laboratory Manual*, 2nd Ed., Cold Spring Harbor Laboratory, Cold Spring Harbor, NY
15. Tanha, J., Muruganandam, A., and Stanimirovic, D. (2003) Mol. Med. 9, 839–843
16. Pace, C. N., Vajdos, F., Fee, L., Grimsley, G., and Gray, T. (1995) Protein Sci. 4, 2391–2423
17. Deng, S. J., MacKenzie, C. R., Hiraama, T., Brousseau, R., Lowary, T. L., Young, N. M., Bundle, D. R., and Narang, S. A. (1995) Proc. Natl. Acad. Sci. U. S. A. 92, 4992–4996
18. Reilly, R. M., Maiti, P. K., Kiarash, R., Prashar, A. K., Fast, D. G., Entwistle, J., Dan, Narang, S. A., Foote, S., and Kaplan, H. A. (2001) Nucleic Acid Mol. Commun. 22, 587–595
19. Bundle, D. R., Eichler, E., Gidney, M. A., Meldal, M., Ragauskas, A., Sigurskjold, B. W., Sinnott, B., Watson, D. C., Yaguchi, M., and Young, N. M. (1999) Biochemistry 38, 5172–5182
20. Piotto, M., Saudek, V., and Sklenar, V. (1992) J. Biomed. NMR 2, 661–665
21. Sklenar, V., Piotto, M., Leppik, R., and Saudek, V. (1993) J. Magn. Reson. 102, 241–245
22. Altieri, A. S., Hilton, D. P., and Byrd, R. A. (1995) J. Am. Chem. Soc. 117, 7566–7567
23. Akerstrom, B., Nilson, B. H., Hoogenboom, H. R., and Bjorck, L. (1994) J. Immunol. Methods 177, 151–163
24. Bond, C. J., Marsteers, J. C., and Sidhu, S. S. (2003) J. Mol. Biol. 322, 643–655
25. Randen, I., Potter, K. N., Li, Y., Thompson, K. M., Pascual, V., Forre, O., Natvig, J. B., and Capra, J. D. (1993) Eur. J. Immunol. 23, 2682–2686
26. Spinnelli, S., Frenken, L., Bourgeois, D., de Ron, L., Bos, W., Verrips, T., Anguille, C., Cambillau, C., and Tegoni, M. (1996) Nat. Struct. Biol. 3, 752–757
27. Terpe, K. (2003) *Appl. Microbiol. Biotechnol.* 60, 523–533
28. Luetz, G. A., Loberski, I., Kavlie, A., Stave, J. E., Borg, T., Kausmal, L., Hvitvatn, E., Simonsen, B., Hovda, M. B., and Brekke, O. H. (2005) J. Immunol. Methods 299, 47–62
29. Cook, G. P., and Tomlinson, I. M. (1995) Immunol. Today 16, 237–242
30. Desmyter, A., Transue, T. R., Arbab-Ghahroudi, M., Thi, M. H., Poortmans, F.
Hamers, R., Muyldermans, S., and Wyns, L. (1996) *Nat. Struct. Biol.* 3, 803–811
31. Ewert, S., Cambillau, C., Conrath, K., and Pluckthun, A. (2002) *Biochemistry* 41, 3628–3636
32. Martin, F., Volpari, C., Steinkuhler, C., Dimasi, N., Brunetti, M., Biasiol, G., Altamura, S., Cortese, R., De Francesco, R., and Sollazzo, M. (1997) *Protein Eng.* 10, 607–614
33. Reiter, Y., Schuck, P., Boyd, L. F., and Plaksin, D. (1999) *J. Mol. Biol.* 290, 685–698
34. Baca, M., Presta, L. G., O’Connor, S. J., and Wells, J. A. (1997) *J. Biol. Chem.* 272, 10678–10684
35. Rader, C., Cheresh, D. A., and Barbas, C. F., III (1998) *Proc. Natl. Acad. Sci. U. S. A.* 95, 8910–8915
36. Roguska, M. A., Pedersen, J. T., Keddy, C. A., Henry, A. H., Searle, S. J., Lambert, J. M., Goldmacher, V. S., Blattler, W. A., Rees, A. R., and Guild, B. C. (1994) *Proc. Natl. Acad. Sci. U. S. A.* 91, 969–973
37. De Genst, E., Silence, K., Ghahroudi, M. A., Decanniere, K., Loris, R., Kinne, J., Wyns, L., and Muyldermans, S. (2005) *J. Biol. Chem.* 280, 14114–14121
38. Forrer, P., Jung, S., and Pluckthun, A. (1999) *Curr. Opin. Struct. Biol.* 9, 514–520
39. Jung, S., Honegger, A., and Pluckthun, A. (1999) *J. Mol. Biol.* 294, 163–180
40. Waldo, G. S. (2003) *Curr. Opin. Chem. Biol.* 7, 33–38
41. Matsuura, T., and Pluckthun, A. (2003) *FEBS Lett.* 539, 24–28
42. Bai, Y., and Feng, H. (2004) *Eur. J. Biochem.* 271, 1609–1614