A Novel Promiscuous Class of Camelid Single-Domain Antibody Contributes to the Antigen-Binding Repertoire

Nick Deschacht, Kurt De Groeve, Cécile Vincke, Geert Raes, Patrick De Baetselier and Serge Muyldermans

*J Immunol* published online 19 April 2010
http://www.jimmunol.org/content/early/2010/04/19/jimmunol.0903722

**Supplementary Material**
http://www.jimmunol.org/content/suppl/2010/04/19/jimmunol.0903722_2.DC1

**Subscription**
Information about subscribing to *The Journal of Immunology* is online at: http://jimmunol.org/subscription

**Permissions**
Submit copyright permission requests at: http://www.aai.org/About/Publications/JI/copyright.html

**Email Alerts**
Receive free email-alerts when new articles cite this article. Sign up at: http://jimmunol.org/alerts
A Novel Promiscuous Class of Camelid Single-Domain Antibody Contributes to the Antigen-Binding Repertoire

Nick Deschacht,*†,1 Kurt De Groeve,*†,1 Cécile Vincke,*† Geert Raes,*† Patrick De Baetselier,*† and Serge Muyldermans*†

It is well established that, in addition to conventional Abs, camelids (such as Camelus dromedarius and Lama glama) possess unique homodimeric H chain Abs (HCAbs) devoid of L chains. The Ag-binding site of these HCAbs consists of a single variable domain, referred to as VHH. It is widely accepted that these VHHs, with distinct framework-2 imprints evolved within the V(H) clan III-family 3, are exclusively present on HCAbs. In this study, we report the finding of a distinct leader signal sequence linked to variable genes displaying a high degree of homology to the clan II, human VH(4) family that contributes to the HCAb Ag-binding diversity. Although the VHH framework-2 imprints are clearly absent, their VH(4)-D-JH recombination products can be rearranged to the H chains of both classical and HCAbs. This suggests that for these V domains the presence of a L chain to constitute the Ag-binding site is entirely optional. As such, the capacity of this promiscuous VH(4) family to participate in two distinct Ab formats significantly contributes to the breadth of the camelid Ag-binding repertoire. This was illustrated by the isolation of stable, dendritic cell-specific VH(4) single domains from a VH(4)-HCAb phage display library. The high degree of homology with human VH(4) sequences is promising in that it may circumvent the need for “humanization” of such single-domain Abs in therapeutic applications.

The Journal of Immunology, 2010, 184: 000–000.

Camelidae are known to possess a dichotomous immune system that enables them to generate 1) classical Abs composed of two H and two L chains and 2) Abs composed of two H chains and no L chains. The latter Abs are referred to as H chain Abs (HCAbs) (1, 2), and their H chain also lacks the first constant domain (CH1), which in a classical Ab interacts with the constant domain of the L chain. The Ag-binding fragment of these HCAbs is reduced in size to a single variable domain (referred to as VHH) and is generated from a V-D-JH gene rearrangement for which a distinct set of V genes, the VHH germline genes, is available in the camelid genome (3–5). Different subfamilies have been distinguished for these VHH genes, but all are closely related to the human VH(3) family of clan III (4, 6). A VH(3) domain of dromedary or llama HCAbs differs from the V(H) domain of classical Abs only by a few crucial substitutions of amino acids normally involved in VL pairing (7, 8). These framework region 2 (FR2) mutations, Val37Phe/Tyr, Gly44Glu, Leu45Arg, and Trp47Gly [Kabat numbering system (9)], are encoded in a dedicated subset of V germline genes of camelids and are therefore not the result of somatic hypermutation (2). Obviously, such hallmark mutations in the FR2 will reshape the VL interface and abrogate a possible VH–VL association. These VHH hallmark residues are also considered to be important for the solubility of the autonomous VHH(3) domain (10, 11) and were demonstrated to affect its stability (12, 13). However, classical Ag-binding single-domain Ab (sdAb) fragments of family VH(3) (i.e., lacking the VHH characteristic hallmark residues in FR2) have been isolated occasionally from immune V libraries, although some controversy remains about the exact IgG isotype (HCAbs or classical) at its origin (14–17). Apart from the absence of the VH(3) hallmark residues, these VH(3) sdAbs also lack the typical interloop disulfide bridge and the long CDR3 loops, yet they remain highly soluble, stable, and functional. So far, it was assumed that all of these variable domains of HCAbs are part of the same clan III VH(3) family. In this study, we report the presence of novel variable domains of clan II and describe their ability to be part of classical as well as HCAbs. Indeed, although all of the V domains of this clan were classical Ab VH sequences, they are able to generate functional HCAbs, so-called VH-HCAbs. Analysis of these novel clan II VH-HCAbs and sequence comparison with human V sequences indicate a strong resemblance with the human VH(4) family. We also report that the same VH(4)-D-JH recombination products can be rearranged to both camelid classical and HCAbs. Therefore, the variable domain obtained after a VH(4)-D-JH recombination is functional as an Ag-binding entity as a single domain or after combining with a VL as a VH–VL pair. Hence, the use of a L chain is entirely superfluous to generate Abs with these VH(4) genes. That these V domains are functional in Ag binding was demonstrated by the retrieval of dendritic cell-specific VH(4) sdAbs from a llama sdAb library derived from the VH(4)-HCAbs. These llama sdAbs exhibit good solubility and are the first autonomous single-domain Ab fragments reported from the VH(4)-HCAbs repertoire.

Materials and Methods

Animals

Mice. Female C57BL/6d mice (6–8 wk old) were used from the in house breeding facility. Animal care and treatment were conducted in conformity with institutional guidelines of the Belgian Council for Laboratory Animal Care.

Abbreviations used in this paper: BM, bone marrow; BMDC, bone marrow-derived dendritic cell; CD, circular dichroism; CH1, first constant domain; DC, dendritic cell; FR, framework region; HCAb, H chain Ab; HIS, histidine; IMAC, immobilized metal ion affinity chromatography; LB, Luria-Bertani; LS, leader signal; rmGM-CSF, recombinant murine granulocyte macrophage colony-stimulating factor; sdAb, single-domain Ab; SEC, size exclusion chromatography.
Science as inspired by the Federation of European Laboratory Animal Science Associations recommendations.

Llama. The *Lama glama* were kept at the premises of the veterinary school of the University of Ghent. All of the animal experiments were conducted with the approval of the Ethical Committee of the Faculty of Veterinary Medicine (University of Ghent, Ghent, Belgium) (18).

Dromedary. The *Camelus dromedarius* was housed at the Central Veterinary Research Laboratory (Dubai, United Arab Emirates).

Cells

Splenic CD11c+ dendritic cells. Four female C57BL/6j mice were infected with 5000 *Trypanosoma brucei* brucei parasites (administered i.p. to recruit dendritic cells [DCs] to the spleen) and sacrificed 7 d postinfection. Splenocytes were isolated and treated with 400 U/ml collagenase III (Roche, Basel, Switzerland) for 30 min at 37˚C in HBSS (Gibco, Invitrogen, Paisley, U.K.) and neutralized by adding 2% heat-inactivated FCS (Hyclone, Logan, UT) and 2 mM EDTA (Invitrogen). Cells were treated with Trix-buffered ammonium chloride to lyse erythrocytes and passed through a nylon mesh to remove undigested material. The splenic myeloid cell fraction was enriched by centrifugation (2000 rpm for 25 min at room temperature) on a Nycodenz gradient (ODnycodenz = 0.660 ± 0.670). The cellular fraction of the interphase was purified by positive selection on magnetic separation columns (MACS) with CD11c+ beads (Miltenyi Biotech, Gladbach, Germany) according to manufacturers’ recommendations.

Marine bone marrow-derived CD11c+ DCs. Femurs and tibias were removed and mechanically purified from surrounding tissues. Subsequently, epiphyses were removed, and bone marrow (BM) was flushed using RPMI 1640 (Invitrogen). Erythrocytes in the single-cell suspensions were lysed as mentioned above. BM cells were cultured in 10 ml RPMI medium containing 10% FCS (Hyclone, Logan, UT), gentamycin, 100 U/ml penicillin, 100 U/ml streptomycin, 2% heat-inactivated FCS, and 2-ME (50 μM; Sigma-Aldrich, St. Louis, MO), and 5% heat-inactivated FCS (HyClone). At day 0, BM cells were seeded at 2 × 10^7 cells per 100-mm diameter bacteriological Petri dish (Becton Dickinson) in 10 ml RPMI medium containing 200 U/ml (20 ml per weak) recombinant murine granulocyte macrophage colony-stimulating factor (mGM-CSF) (gift from Dr. H. Wereley, University of Oxford, Oxford, U.K.). At day 3, another 10 ml RPMI medium containing 200 U/ml mGM-CSF was added to the plates. At day 6, half of the culture supernatant was collected and centrifuged, and the cell pellet was re-suspended in 10 ml fresh R10 medium containing 10% FCS and poured into the original plate. Reducing the concentration of the mGM-CSF during culture diminishes the granulocyte contamination. Harvesting the cells 1 d later yielded the immature DC population. For complete maturation of DCs, the 7-d-old non-adherent cells were collected by gentle pipetting, centrifuged at 1500 rpm for 5 min at room temperature, and resuspended in 6 ml fresh R10 medium to a fresh six-well plate containing 100 U/ml rmGM-CSF and 5 μg/ml LPS (Sigma-Aldrich). Cells were then cultured for an additional 1 or 2 d to obtain the mature DC population (19).

RACE of dromedary IgG

PBLS were isolated from dromedary or llama using Lymphoprep (Axis-Shield, Oslo, Norway), and mRNA was extracted (20). Then, with a SMART RACE cDNA amplification kit (BD Clontech, Heidelberg, Germany), cDNA was generated, and 5’ RACE PCR products were amplified by using a FR4-specific primer (5’-GGACTAIGCGCGCGGCAGACCTGGTACCGGTG-3’). These 5’ RACE products were cloned into the pCR2.1-TOPO vector using the TOPO TA cloning kit (Invitrogen), and insert-positive clones were identified for sequencing. A llama (L. glama) was injected s.c. with 10^9 immature mouse BM-derived DCs (BMDCs) without adjuvant. The immunization was repeated six times at weekly intervals. After the last boost, anticoagulated peripheral blood was taken from the jugular vein to prepare an immune library. Gene fragments from the V region leader sequence upstream until the CH2 domain of the IgGs were first amplified as described above. The PCR fragments containing the variable gene fragments of the IgG HCabs were purified from agarose gels. Subsequently, the gene fragments encoding the entire V region were amplified with one additional PCR with nested primers annealing at FR1 (5’-CCACGCCGCTGTCAGTGCAGCTGCTGTCATGGTCCG-3’) and FR4 (5’-GGACTAIGCGCGCGGCAGACCTGGTACCGGTG-3’). The PCR fragments were ligated into the phagemid vector pHEN4 (20), using the restriction sites NcoI and NotI (underlined). Ligated material was transformed in *Escherichia coli* cells (TG1) and plated on selective medium (16). The colonies were scraped from the plates and washed, and this cell library was stored at −80˚C in Luria-Bertani (LB) medium supplemented with glycerol (50% final concentration).

Selection of Ag-specific sdAb fragments

A representative aliquot of the library was cultured in LB medium with ampicillin and infected with M13K07 helper phages (16) to express the cloned V region at the tip of phage particles as fusion products with the gene III (bacillus phage coat protein). DC-specific V-region domains from this library were enriched using whole-cell panning on an in vivo cell population of CD11c+ monocytes from the spleens of mice (26). Freshly harvested splenic DCs (1 ml, 4 × 10^6 cells) were incubated with phage particles (1.5 × 10^10) for 1 h at room temperature. Cells were then pelleted by low-speed centrifugation and resuspended in 2 ml MACS buffer, and anti-CD11c beads (Miltenyi Biotech) were added to the suspension. Once mixed with the beads, the cells were then isolated via CD11c+ selection by MACS according to the manufacturer’s protocol (Miltenyi Biotech). After the final wash, the column was removed from the magnet, and the bead-coated/phage-coated/CD11c+ cells were eluted from the column with 5 ml sterile PBS. The CD11c+ cells were immediately centrifuged at 1800 rpm for 10 min and resuspended in 200 μl triethylamine (pH 10) to elute cell-bound phages and lyse cells. After 10 min incubation at room temperature, the solution of phage particle eluate and cellular debris was neutralized with 200 μl Tris-HCl (1 M, pH 7.4) and used to infect exponentially growing *E. coli* TG1 cells that were further cultured for the second round of selection. An aliquot of the neutralized phage particle eluate was also taken to infect bacteria and plated on LB agar plates supplemented with ampicillin. After each round of panning, individual colonies were picked randomly from the titration plate, and their sdAb gene was sequenced.

Purification and identification of DC-specific sdAb fragments

**Purification of the sdAb.** The variable regions of the clones that appeared multiple times after panning in our sequence analysis were recloned into the pCR2.1-TOPO vector using the restriction site Eco91I (Fermentas) and transformed into *E. coli* WK6 cells. The V gene, under control of the lac promoter, is preceded by a PelB sequence to direct the translation product to the periplasm and followed by a His6 tag to facilitate the purification of the sdAb protein by immobilized metal ion affinity chromatography (IMAC). Liter-scale production and purification of these sdAbs by IMAC and gel filtration (Sephadex 75 column) in PBS were performed as described previously (16).

Biophysical stability. To analyze the oligomeric state of the sdAbs, the VH4) sdAbs and two (family 3) sdAbs (28) were expressed and subsequently purified (16). The purified monomeric fractions of the sdAbs in PBS at two different concentrations (55 μM [0.5 mg/ml] and 175 μM [2.5 mg/ml]) were incubated at 37 and 4˚C for up to 14 d. Samples (0.5 mg of each sdAb) were taken at different time points (day 0, 1, 7, and 14), and their integrity was analyzed by size exclusion chromatography (SEC) on an A¨ KTA Explorer (GE Healthcare, Munich, Germany) using Sephadex 75 in PBS and by circular dichroism (CD) measurement. CD spectra were obtained with a Jasco J715 spectropolarimeter in the UV region (200–350 nm), using a protein concentration of 0.2 mg/ml and a 0.1-cm cell path length. Data were acquired with a reading frequency of 1/20 s, a 1 s integration time, and a 2 nm bandwidth. Data analysis was performed according to Durnouit et al. (13).

Temperature-induced denaturation. Thermal unfolding was followed by CD with a Jasco J715 spectropolarimeter in the UV region (190–350 nm). A total volume of 300 μl of each sample was heated in 50 mM sodium phosphate buffer (pH 7). The temperature was increased from 35 to 95˚C at a rate of 1˚C/min, and the signal intensity at 205 nm was recorded as a function of temperature.
Flow cytometry. For flow cytometric analysis, the DC population was stained with Abs (from BD Biosciences [San Jose, CA], unless otherwise stated) for 20 min at 4°C using standard protocols. Cells (10^6 per 100 μl) were incubated with 1 μg anti-FcγRI Ab (clone 2.4G2) before adding allophycocyanin-conjugated CD11c (HL3) Ab and 1 μg FITC-conjugated MHC II (clone M5/114.15.2), FITC-conjugated CD86 (Gl-1), or FITC-conjugated CD40 Ab (clone 3/23). Cell recognition of different sdAbs (at 1 μg) was monitored using 0.1 μg FITC-conjugated anti-histidine (HS) (Serocte, Oxford, U.K.) as a secondary Ab. Multiparameter acquisition was performed on an FACS Canto II (BD Biosciences), followed by analysis with FlowJo (TreeStar, Ashland, OR).

Alexa Fluor 488 labeling of sdAb. The sdAbs were labeled using the Alexa Fluor 488 protein labeling kit (Molecular Probes, Eugene, OR). Briefly, 100 μl of 1 M sodium bicarbonate buffer was added to 1 ml sdAb solution at a concentration of 1 mg/ml. The sdAb solution (either sdAb_31 or sdAb_32) was subsequently transferred to a vial of reactive dye (Alexa Fluor 488 carboxylic acid, tetrafluorophenyl ester) and incubated overnight at 4°C under mild agitation. Labeled sdAbs_31 and sdAbs_32 were separated from unreacted dye by SEC on Sephadex 75 in PBS. The protein concentration and the degree of Alexa Fluor 488 labeling of the sdAb was determined via spectrophotometry (Nanodrop ND-1000; Isogen Life Sciences, De Meern, The Netherlands).

In vitro competition study. Each BMDC population (10^7 cells in 100 μl) was preincubated with 40 μg unlabeled sdAb (sdAb-BCII10, sdAb_31 or sdAb_32) for 20 min at 4°C. The cells with bound sdAb were recovered in a pellet by low-speed centrifugation (supernatant containing the unbound sdAb was removed), resuspended in PBS, stained with 1 μg Alexa Fluor 488-labeled sdAb_31 or sdAb_32, and incubated for 20 min at 4°C. Multiparameter acquisition was performed on an FACS Canto II (Becton Dickinson), followed by analysis with FlowJo (Tree Star).

Results

Identification in cameldigs of a distinct LS linked to clan II variable genes

A 5’ RACE on cDNA of llama or dromedary PBLs with a JH-specific primer (i.e., FR4 of VH) and a 5’ universal primer was used to investigate the V repertoire. This 5’ RACE resulted in a PCR fragment of ~500 bp that was cloned into TOPO TA. Sequence analysis of these cloned inserts revealed a few clones encoding a LS sequence that deviates from the established LS sequence of cameldig VH(3) and VH(3) genes (3) (Supplemental Fig. 1). The V region downstream of the novel LS was aligned to the human IgV genes (23). However, previous reports assigned the dromedary V genes to the VH(3) family of clan III (3, 4), the International ImMunoGeneTics information system amino acid numbering/V-QUEST (www.imgt.org) and IgBlast (www.ncbi.nlm.nih.gov/igblast) search located the novel LS and its V sequences to the human V genes of clan II.

Note that on previous occasions we always amplified the VH1 genes with LS-specific primers of the VH(3) family that would fail to amplify these new V genes, and therefore these V genes were eliminated from all of our immune V libraries (15). Hence, a new primer (CALLS) based on this novel LS (Supplemental Fig. 1) was designed. Interestingly, the PCR amplification with CALLS and a CH2-IgG-specific primer generated two distinct amplicons (Fig. 1). From their sizes, it can be inferred that the longest fragment (~1000 bp) corresponds to the H chain fragment of classical IgG Abs, whereas the shorter fragment (~700 bp) probably encodes a fragment of the H chain of the HCAbs. The difference in size accounts for the absence or presence of the CH1 region that occurs only in H chains of classical IgGs (29). Fragments of both, the HCAbs and the classical Abs, were gel-purified, cloned independently in TOPO TA vector, and sequenced. These particular clan II variable genes occurred in the HCAb of all cameldig samples tested, although the band intensity varied and the fraction in the classical H chain was always more pronounced (Fig. 1). Surprisingly, a 5’ RACE IgG analysis of the classical H chains demonstrated that the VH(4) genes contributed up to 50% of the classical Ab repertoire.

Phylogenetic sequence analysis

A total of ~150 individual clones, randomly chosen over both clonings (i.e., classical and HCAb-derived), were sequenced. After removing ambiguous sequence readings, nonfunctional sequences (containing a reading frame shift), or incomplete sequences, we assembled two sets of 58 V gene sequences for the H chain of the classical and HCAb IgGs, respectively (Supplemental Figs. 2, 3). Sequence alignment demonstrated that the V regions of both Ab classes belong to the same V gene family and both groups of V regions encode classical VH elements (not VHH). Furthermore, a neighboring-joining analysis was performed with the V-encoding cDNA sequences and the functional human IgV germlines. Analysis of these llama V elements (of classical Abs and HCAbs) showed that they all cluster in one branch that is most closely related to the branch of the human VH (4 family), which forms part of clan II (Fig. 2). The llama VH(4) and human Ig VH(4) domains display high sequence similarity, yet some framework residues (Thr20, Ile24, Thr33, Arg39, Gln41, Pro57, and Glu65) deviate regularly from the human VH(4) sequences (23). The impact of these deviations is visualized by the hydrophobicity plot (Fig. 3). The decreased hydrophobicity profile at the outer framework-3 (asterisks in Fig. 3) for the llama VH(4)s compared with that of human VH(4) is suggestive for a reduced aggregation-prone behavior for llama VH(4) relative to that reported for human VH(4) (30).

VH(4) enlarges both the classical and the HCAb diversity

Analysis of VH(4) sequences from both isotypes (i.e., classical Abs and HCAbs) showed that sequence variation was the highest in the CDRs of the V regions (Supplemental Figs. 2, 3). Variability was also noticed in the FR3 regions, although to a lesser extent compared with that of the CDRs. No obvious differences in CDR1 sequence are observed between VH(4)s from classical Abs and HCAbs (Supplemental Figs. 2, 3). With the notable exception of three clones (conv12, 32, and HCAb28), both VH(4) groups contain, similar to the human VH(4), 7 aa in their CDR1. Such a CDR1 loop of 7 aa is predicted to adopt canonical structure type-3 (31). The HCAb28 clone possesses an enlarged CDR1 and forms potentially a novel loop structure, whereas the other two classical VH(4) clones (clones 12 and 32) probably adopt canonical loop structure type-1 and type-2, respectively. Likewise, no major differences are detected for CDR2 between VH(4)s of both groups (Supplemental Figs. 2, 3). The vast majority of the CDR2 loops display a length of 16 aa, except clones conv56 and conv57 (Supplemental Fig. 2). Therefore, with the exception of these clones, the CDR2 regions of llama VH(4) are predicted to fold into the known canonical loop structures of the human VH(4) domains (31). In summary, the CDR1 and CDR2 Ag-binding loops of llama VH(4) apparently adopt the canonical loop structures observed for human VH(4) domains. In sharp contrast, these loops are routinely deviating from the canonical conformations in the case of VHH of V family 3 (11, 32). Nonetheless it is obvious that the mere presence of these VH(4) domains within the
HCAb type enlarges the known structural repertoire of these unique camelid Abs (4). In contrast to CDR1 and CDR2 that is imprinted in the V germline gene, CDR3 arises from an imprecise V-D-JH recombination, which creates a much larger variability in loop length and sequence. Within the VH(4) dataset, CDR3 ranges from 7 to 22 aa (Supplemental Figs. 2, 3), although the average length of CDR3 from classical Abs (12.8 aa) does not deviate significantly from that of HCAbs (13.5 aa). In contrast, for dromedary VH(3) and VHH(3), a net difference in CDR3 length is observed (9 and 15 aa, respectively) (33), although this difference is less pronounced in llama (9 and 13 aa) (8).

From inspection of the CDR3 sequences, it appears that the VH(4) sequences of classical Abs and HCAbs are recombined to a common set of D and JH genes, although such analysis is complicated by imprecise recombination and somatic hypermutation. A minor difference is that VH(4) of the classical Ab group lacks inter- or intraloop disulfide bridges (Supplemental Figs. 2, 3). In contrast, for the HCAb VH(4) sequences, clones 36 and 12 possess two cysteine residues in their CDR3 loop that probably form an intraloop disulfide bond.

In conjunction, no major difference in sequence or loop structure was discerned between the VH(4) from classical Abs and HC Abs (Supplemental Figs. 2, 3). Interestingly, three different VH(4) sequences were found twice, once in the classical Ab group and once in the HCAb group (clones HCAb10/conv46, HCAb8/conv19, and HCAb52/conv7). When aligned pairwise (Fig. 4A), these VH(4)s contained a minimal amount of point mutations and a homologous CDR3 sequence, indicating that each pair probably originates from an identical V-D-JH recombination. One sequence of each pair was linked with (or properly spliced to) an intact CH1 region and will therefore be part of an Ab with an Ag-binding fragment comprising a VH and VL domain. The twin sequence partner definitely lacked the CH1 region, and the V region was immediately followed by the hinge sequence, which demonstrates that this V domain was derived from a HCAb devoid of L chains. Therefore, this latter VH(4) domain will function as a single-domain Ag-binding fragment in a HCAb. This indicates that a VH(4) gene involved in a V-D-JH recombination in a pre-B cell generates a VH domain that might function either independent from a VL in a HCAb or paired with a VL domain in a classical Ab.

**Generation of VH(4) library and selection of Ag-specific single-domain Abs**

To verify that the isolated VH(4) domains from HC Abs (i.e., in the absence of a VL partner) are indeed functional in Ag binding, we generated a VH(4)-HCAb–derived single-domain phage display library of a llama immunized with immature BMDCs. Two rounds of panning were performed on purified splenic CD11c+ cells (26). After the second round of panning, the number of phage particles eluted from CD11c+ cells increased 100-fold relative to the first round using the same amount of input phage particles and identical washing steps. Individual bacterial colonies (∼100) from the second round were randomly picked and sequenced. Two sequences (referred to as sdAb_31 and sdAb_32) were represented repeatedly, although several point mutations throughout the framework or CDR occurred in the homologous clones of the dominating sdAb_31 sequence. The alignment of the deduced amino acid sequences, against the human VH consensus sequence (Fig. 4B) confirmed their VH(4) signature. The CDR3 lengths of 14 and 15 aa for sdAb_31 and sdAb_32, respectively, contrast with the shorter average CDR3 length of llama VH(3) (8). In addition, cysteine residues were absent in the CDR, and therefore these long loops cannot be tethered by an inter- or intraloop disulfide bond that might assist in the stabilization of the sdAb (34).
A

Biochemical properties of Ag-specific VH(4) sdAb

The two VH(4) sdAbs were recloned in the pHE6 expression vector. After induction, the periplasmic proteins were extracted, and the recombinant protein was purified by IMAC and SEC (Supplemental Fig. 4). A yield of 5–7 mg protein per liter of culture was obtained for these sdAbs, which is within the range (5–10 mg/l culture) normally obtained for sdAbs of the VHH(3) family (15). The sdAb_31 eluted from a Superdex 75 gel filtration column as a single symmetrical peak as expected for a soluble, monomeric protein, whereas the elution profile of the sdAb_32 showed an additional minor fraction of dimmerized material. The thermal stability of the sdAbs was measured by CD, and T_m values of 62.4 and 63.5°C were recorded for sdAb_31 and sdAb_32, respectively (Supplemental Fig. 5). Although these T_m values are similar to the values measured for VHH(3) sdAbs (13), they are high for sdAbs belonging to the VH(4) family, because V domains of the human VH(4) family form consistently less stable domains (30). Earlier studies proposed for sdAbs a direct correlation between their net acidic charge and aggregation resistance (35, 36). Therefore, we calculated the theoretical pI of VH(4)s of both Ab groups. The distribution of the individual pIs of VH(4)s from classical Abs or HCAbs as well as the pI ranges are indistinguishable (Supplemental Fig. 6A). Remarkably, in this current dataset of VH(4)s of unknown specificity, no pI values between 7.0 and 7.5 were recorded. However, the DC-enriched sdAb_32 and sdAb_31 display a near neutral and basic pI, respectively, and are soluble and monomeric entities, in sharp contrast to predictions from previous work (35, 36).

We next analyzed the integrity and aggregation tendency of our two DC-specific VH(4) sdAbs over time, relative to the properties of VHH(3)-derived sdAbs and DC-specific VH(3) sdAbs (28). These four sdAbs at two different concentrations (35 and 175 μM) were incubated at 4 and 37°C for different time intervals up to 14 d. Gel filtration was used to profile the quaternary structure of the sdAbs, and the CD signal was taken to ascertain the maintenance of their structural integrity. The VH(4) sdAbs eluted as a single symmetric peak at all incubation times, and their chromatographic profiles remained functional (data not shown). Their epitope is present on human IgHV4_30 and human IgHV4_33, and the VH(4) sdAbs appear robust.

The Ag-binding profiles of sdAb_31 and sdAb_32 on mature and immature BMDCs were analyzed by flow cytometry (Fig. 5A). The results of this in vitro binding study demonstrate that sdAb_31 and sdAb_32 are functional in Ag binding, and as expected from the results above, even after 2 wk incubation at 37°C, our sdAbs remained functional (data not shown). Their epitope is present on CD11c+ cells, including both mature and immature DCs (Fig. 5A). The observed signals are due to the sdAbs recognizing specifically their cognate cell surface markers, because control experiments of BMDCs with sdAb_BCII10, a sdAb with specificity for β-lactamase, gave background signals that were indistinguishable from those obtained with cells mixed with FITC-labeled anti-HIS secondary Ab. In addition, the experiments with other non-DC-specific VH(4) sdAbs all failed to give a fluorescence intensity shift (Fig. 5A). This precludes that the signal originates from a nonspecific adsorption of VH(4) sdAbs on the BM-DCs. Furthermore, the flow cytometry signal intensity with the sdAbs directed against the immature DCs is slightly higher than that for mature DCs, and the relative shift of the fluorescence intensity peak is more pronounced for sdAb_32 than that for sdAb_31. Finally, experiments with Alexa Fluor 488-labeled sdAb tested in competition with an excess of unlabeled sdAb confirmed the specific recognition of their cognate epitope and indicated that sdAb_31 and sdAb_32 recognize different (i.e., noncompeting) epitopes on their target cells (Fig. 5B).

Discussion

The Ab generation process in Camelidae should deviate, at least in some details, from the general scheme in other vertebrates because camels are actively producing unique, functional HCAbs. For the production of HCAbs, the camels use dedicated C genes and JH genes (4, 5, 37). In addition, the camelid VH(3), VHH(3), and VH(4) sdAbs appear robust.

The Ag-binding profiles of sdAb_31 and sdAb_32 on mature and immature BMDCs were analyzed by flow cytometry (Fig. 5A). The results of this in vitro binding study demonstrate that sdAb_31 and sdAb_32 are functional in Ag binding, and as expected from the results above, even after 2 wk incubation at 37°C, our sdAbs remained functional (data not shown). Their epitope is present on CD11c+ cells, including both mature and immature DCs (Fig. 5A). The observed signals are due to the sdAbs recognizing specifically their cognate cell surface markers, because control experiments of BMDCs with sdAb_BCII10, a sdAb with specificity for β-lactamase, gave background signals that were indistinguishable from those obtained with cells mixed with FITC-labeled anti-HIS secondary Ab. In addition, the experiments with other non-DC-specific VH(4) sdAbs all failed to give a fluorescence intensity shift (Fig. 5A). This precludes that the signal originates from a nonspecific adsorption of VH(4) sdAbs on the BM-DCs. Furthermore, the flow cytometry signal intensity with the sdAbs directed against the immature DCs is slightly higher than that for mature DCs, and the relative shift of the fluorescence intensity peak is more pronounced for sdAb_32 than that for sdAb_31. Finally, experiments with Alexa Fluor 488-labeled sdAb tested in competition with an excess of unlabeled sdAb confirmed the specific recognition of their cognate epitope and indicated that sdAb_31 and sdAb_32 recognize different (i.e., noncompeting) epitopes on their target cells (Fig. 5B).
a homodimeric Ig. In contrast, VH(3)-D-JH is cotranslated with the classical C\(_{\text{g}1}\)a or C\(_{\text{g}1}\)b, and the polypeptide associates with a L chain to form a classical Ig (Fig. 6). However, it has been reported that occasionally VH(3)D-JH products lacking the CH1 exon can be cloned from cDNA, indicating that they probably are part of the HCAb pool (14, 16). The majority of such classical V domains of a HCAb are (most often) generated after an unusual D-JH gene recombination whereby the codon for the conserved Trp103 (a residue that is essential for VL pairing in classical Abs) is substituted mostly by an arginine codon (6, 14, 16). In this study, we demonstrate that the V repertoire of camelids also contains functional V genes that belong to clan II. These V genes cluster with the Ig VH(4) family in the phylogenetic tree and the IgBlast search (Fig. 2). The occurrence in alpaca of such genes was briefly mentioned recently (5). Interestingly, in llama, the VH(4) gene is transcribed in nearly 50% of the expressed classical IgGs (possessing CH1 and by default also the L chain). This percentage comes from a 5' RACE cloning of the H chain of classical Abs and shotgun sequencing that was not biased for any particular V element. The abundant usage of VH(4) within this pool suggests that these V genes fulfill a central role in forming the Ag-binding repertoire of classical Abs in llama. The reason why their presence was overlooked until now is dual: first, most work on camelid Abs focuses on the HCAbs and neglects the classical Abs (14, 15). Secondly, the V genes are normally amplified after PCR with primers designed to anneal at the LS sequence or at the FR1, and these primers fail to hybridize to the corresponding regions in the VH(4) family. In this study, we demonstrate that this VH(4) family contributes significantly to the classical Ig Ab repertoire and, more importantly, is numerously present as part of HCAbs as well.

Nonetheless, the frequency of the VH(4) occurrence within the HCAb pool is lower than that in the classical Ab group. Interestingly, the very same VH(4)-D-JH rearrangement can occasionally be found within both classical Abs and HCAbs. In the former case, it is coexpressed with a CH1 domain, and probably the VH domain will pair with a VL domain to constitute a functional Ag-binding site. In the latter case, the CH1 exon is definitely removed during the splicing reaction within the mRNA of the HCAb-specific Cy2 or Cy3 genes and as such impeding a possible association with

**FIGURE 5.** Binding profile of VH(4) sdAb_31 and sdAb_32 on BMDCs. The activation state of the DC populations was evaluated using different cell surface markers (CD40, CD86, and MHC II). A, DC populations were stained with 1 \(\mu\)g sdAb-BCII10 [VHH(3); negative control], a nonspecific VH(4) sdAb, or VH(4) sdAb_31 or sdAb_32 visualized by 0.1 \(\mu\)g anti–HIS-FITC Ab gated on CD11c\(^+\) BMDCs (solid black line and open profiles). The profiles in dashed lines and gray filling are from cells treated with 0.1 \(\mu\)g anti–HIS-FITC only. B, Immature DCs preincubated with 40 \(\mu\)g unlabeled sdAb-BCII10, sdAb_31, or sdAb_32 (gray-filled, dashed-lined contoured profiles) followed by staining with 1 \(\mu\)g sdAb_32-Alexa Fluor 488 (open, solid-line contoured profiles).
FIGURE 6. Schematic representation of the H locus with V genes of VH(4) and VH(3) family followed by D and JH gene clusters and constant genes (only Cy genes are shown). Note that the current organization of V genes (clustered or interspersed) is currently unknown, and therefore the order of VH(4), VH(3), and VH(3) is entirely tentative. The dedicated HCAb Cy2 and Cy3 have a deficient CH1 splicing site denoted by a star. A, (i) Generation of classical IgG is obtained after rearrangement of a classical VH(3) gene and coexpression with classical Cy(1a,1b) genes, (ii) generation of HCAbs is possible when a classical VH(3) gene recombines with a D-JH that encodes an Arg103 mutation (although sometimes Arg103 is absent) and coexpresses with the dedicated HCAb Cy2 and Cy3 genes, and (iii) the majority of HCAbs are formed when a VH(3) gene that differs from the VH(3) genes in its hallmark FR2 amino acids is linked after a D-JH rearrangement to the dedicated Cy2, Cy3 constant genes. B, Generation of HCAbs with a VH(4) gene without the Trp103 substitution can be joined to both sets of Cy genes to produce classical Abs or to produce HCAbs.

a L chain. Because the VH(4)-D-JH translation product seems to be competent to function on both types of Abs, the role of the L chain (and the VL in particular) appears to be superfluous for the expression of the BCR and for proper Ag recognition. Consequently, the VH(4) domain is forced to act as an autonomous, single-domain Ag-binding fragment in a HCAb molecule. Hereina, the stability and solubility of the VH(4) domain and its secretion from the B cell should be assured, and the Ag recognition of the autonomous domain should be guaranteed.

The dual usage of VH(4) raises two questions. First, is VH(4) of the HCAbs a particular subset of the VH(4)-D-JH rearrangement products or are all of the VH(4) domains capable to appear on either Ab type? Second, what makes these camelid VH(4) different from the human VH(4) so that they can function as an isolated sdAb whereas a human VH(4) is much less adapted to fulfill this role (30)?

With regard to the first question, VH(4) on HCAbs and classical Abs seems to be indistinguishable. There is apparently no difference in usage of the VH(4) or JH genes. In both cases, they rearrange with the D gene to generate CDR3s of equal lengths. Also, the amino acid content and hydrophobicity of CDR3 within the VH(4) domains of both Abs reveal no significant bias (Supplemental Fig. 6). Additionally, the distribution of the pl of the VH(4) domains from HCAbs or classical Abs is identical. Remarkably, few domains are formed with a pl between 7.0 and 7.5, which might reflect a counterselection against such domains because these would be more prone to aggregation at a physiological pH. We also analyzed areas on the VH(4) domain, such as FR2 and Trp103 (first amino acid of FR4), where VH(3)s and VH(3)s of HCAbs and classical Abs are distinct. Yet again, FR2 of VH(4)s from HCAbs and classical Abs are identical. Although a few sequences from the HCAb pool show some amino acid substitutions in this conserved VL interface region, similar substitutions were noted for the classical Abs as well (Supplemental Figs. 2, 3; HCAb23, HCAb24, and conv33, respectively). The Trp103 (encoded by the 5’ end of the JH gene) is another major component of the -contacting side of a VH domain. Indeed, for the VH(3)-derived Ag-binding fragments of HCAbs, we observe regularly an arginine at this position, which will abrogate the VL pairing (7, 38).

Remarkably, the Trp103Arg substitution is highly employed in VH(3)-HCAbs having a short CDR3, whereas a small number of VH(3)-HCAbs have a long CDR3 loop but lack this Trp103Arg substitution. This indicates that the CDR3 amino acid content in conjunction with the long sequence that collapses on the erstwhile VL interface mimics the Trp103Arg effect. The situation for VH (4) is different. The VH(4) of HCAbs and classical Abs possess a CDR3 loop of equal average length, and in contrast with the VH (3)-HCAbs the majority of the VH(4)-HCAbs do normally not use Arg103 substitution to impede the VH–VL association.

Concerning the second question, we searched for possible determinants that enhance the solubility and good biophysical properties of the isolated camelid VH(4) domain compared with the poor behavior of the isolated human VH(4) (30). These determinants should belong to the CDRs or to the framework, and in the latter case preferably the VL binding interface (39–42). Although the substitution of Trp103, Leu45, or both of FR2 by a charged residue (arginine) may be favorable for the solubility of sdAbs, previous studies clearly demonstrate that it is not a prerequisite (17, 36, 41, 43). Despite the high sequence similarity between human and llama VH(4), profound sequence analysis of the camelid VH(4)s highlighted a few residues that deviate subtly but consistently from the human VH(4) framework. Remarkably, these residues are mainly present in FR3 (Supplemental Fig. 4) outside the VL binding face of the domain and render this area more hydrophilic than the human VH(4)s. Because framework residues influence the biochemical properties of sdAbs (44–46), it is conceivable that the unfavorable properties described for human single domains of family VH(4) (30) are attenuated through these substitutions. However, it would be surprising that these differences are the sole determinant for the autonomous character of the camelid VH(4)-HCAbs. Additional key residues are probably located in the CDR loops because these regions are reported to participate in the soluble behavior of autonomous human VH(3) domains (36, 41, 43, 47). Furthermore, the
nature of the CDR3 loop has been shown to contribute to the sta-
bilization and solubility of sdAbs (30, 40). However, the inherent hy-
pervariable character (in length and sequence) of the CDR re-
gions precludes the identification of the critical residues to convert
VH–VL domains into functional monomeric sdAbs.

That the VH(4) sdAbs are indeed functional in Ag binding was
shown by immunizing a llama with DCs, cloning the VH(4) rep-
ertoire from the HCAb-specific pool, and retrieving DC-specific
sdAbs after phage display. Of note, in the absence of any information
on a possible Ag preference of the VH(4) of HCAbs, we preferred to
immunize with cells as opposed to individual proteins, because theo-
retically these harbor an infinite number and complex type of epitopes,
which maximizes our chances to retrieve binders afterward. In addition, the retrieved VH(4) sdAbs produce readily in
bacterial expression systems, display good solubility, thermal sta-
nability, and shelf life, and are functional in the recognition of CD1c+ cells as measured by flow cytometry. Furthermore, the VH(4) of
camelids are closely related and share a high degree of identity with
human VH(4). It is therefore conceivable that for human therapy one
would prefer to select specifically VH(4)-derived sdAbs instead of
VHH(3)-based sdAbs because the latter might require more
drastic humanization efforts to minimize immunogenicity (12).

We also foresee two extra possible developments whereby our insights on camelid autonomous VH(4)s could become practical in the future. In one example, we propose to remediate an “unstable”
human VH(4) that was identified (either as an autonomous sdAb or paired with L domains in a scFv or Fab format) by substituting its
FR3 beginning and ending sequences with the more hydrophilic amino acids as those of the llama VH(4). For the second possibility,
a strategy could be developed whereby a synthetic library would be
constructed by randomizing the CDR3 sequence on a llama VH(4)
as amino acids as those of the llama VH(4). For the second possibility,
a strategy could be developed whereby a synthetic library would be
constructed by randomizing the CDR3 sequence on a llama VH(4)
as amino acids as those of the llama VH(4). For the second possibility,
a strategy could be developed whereby a synthetic library would be
constructed by randomizing the CDR3 sequence on a llama VH(4)
as amino acids as those of the llama VH(4). For the second possibility,
a strategy could be developed whereby a synthetic library would be
constructed by randomizing the CDR3 sequence on a llama VH(4)
as amino acids as those of the llama VH(4). For the second possibility,
a strategy could be developed whereby a synthetic library would be
constructed by randomizing the CDR3 sequence on a llama VH(4)
as amino acids as those of the llama VH(4). For the second possibility,
a strategy could be developed whereby a synthetic library would be
constructed by randomizing the CDR3 sequence on a llama VH(4)
as amino acids as those of the llama VH(4). For the second possibility,
a strategy could be developed whereby a synthetic library would be
constructed by randomizing the CDR3 sequence on a llama VH(4)
as amino acids as those of the llama VH(4). For the second possibility,
a strategy could be developed whereby a synthetic library would be
constructed by randomizing the CDR3 sequence on a llama VH(4)
as amino acids as those of the llama VH(4). For the second possibility,
a strategy could be developed whereby a synthetic library would be
constructed by randomizing the CDR3 sequence on a llama VH(4)
as amino acids as those of the llama VH(4). For the second possibility,
a strategy could be developed whereby a synthetic library would be
constructed by randomizing the CDR3 sequence on a llama VH(4)
as amino acids as those of the llama VH(4). For the second possibility,
a strategy could be developed whereby a synthetic library would be
constructed by randomizing the CDR3 sequence on a llama VH(4)
as amino acids as those of the llama VH(4). For the second possibility,
a strategy could be developed whereby a synthetic library would be
constructed by randomizing the CDR3 sequence on a llama VH(4)
as amino acids as those of the llama VH(4). For the second possibility,
a strategy could be developed whereby a synthetic library would be
constructed by randomizing the CDR3 sequence on a llama VH(4)
as amino acids as those of the llama VH(4). For the second possibility,
a strategy could be developed whereby a synthetic library would be
constructed by randomizing the CDR3 sequence on a llama VH(4)
as amino acids as those of the llama VH(4). For the second possibility,
a strategy could be developed whereby a synthetic library would be
constructed by randomizing the CDR3 sequence on a llama VH(4)
as amino acids as those of the llama VH(4). For the second possibility,
a strategy could be developed whereby a synthetic library would be
constructed by randomizing the CDR3 sequence on a llama VH(4)
as amino acids as those of the llama VH(4). For the second possibility,
a strategy could be developed whereby a synthetic library would be
constructed by randomizing the CDR3 sequence on a llama VH(4)
as amino acids as those of the llama VH(4). For the second possibility,
a strategy could be developed whereby a synthetic library would be
constructed by randomizing the CDR3 sequence on a llama VH(4)
as amino acids as those of the llama VH(4). For the second possibility,
a strategy could be developed whereby a synthetic library would be
constructed by randomizing the CDR3 sequence on a llama VH(4)
as amino acids as those of the llama VH(4). For the second possibility,
a strategy could be developed whereby a synthetic library would be
constructed by randomizing the CDR3 sequence on a llama VH(4)
as amino acids as those of the llama VH(4). For the second possibility,
a strategy could be developed whereby a synthetic library would be
constructed by randomizing the CDR3 sequence on a llama VH(4)
as amino acids as those of the llama VH(4). For the second possibility,
a strategy could be developed whereby a synthetic library would be
constructed by randomizing the CDR3 sequence on a llama VH(4)
as amino acids as those of the llama VH(4). For the second possibility,
a strategy could be developed whereby a synthetic library would be
constructed by randomizing the CDR3 sequence on a llama VH(4)
as amino acids as those of the llama VH(4). For the second possibility,
a strategy could be developed whereby a synthetic library would be
constructed by randomizing the CDR3 sequence on a llama VH(4)
as amino acids as those of the llama VH(4). For the second possibility,
a strategy could be developed whereby a synthetic library would be
constructed by randomizing the CDR3 sequence on a llama VH(4)
as amino acids as those of the llama VH(4). For the second possibility,
a strategy could be developed whereby a synthetic library would be
constructed by randomizing the CDR3 sequence on a llama VH(4)
as amino acids as those of the llama VH(4). For the second possibility,
a strategy could be developed whereby a synthetic library would be
constructed by randomizing the CDR3 sequence on a llama VH(4)
as amino acids as those of the llama VH(4). For the second possibility,
a strategy could be developed whereby a synthetic library would be
constructed by randomizing the CDR3 sequence on a llama VH(4)
as amino acids as those of the llama VH(4). For the second possibility,
a strategy could be developed whereby a synthetic library would be
constructed by randomizing the CDR3 sequence on a llama VH(4)
as amino acids as those of the llama VH(4). For the second possibility,
a strategy could be developed whereby a synthetic library would be
constructed by randomizing the CDR3 sequence on a llama VH(4)
as amino acids as those of the llama VH(4). For the second possibility,
a strategy could be developed whereby a synthetic library would be
constructed by randomizing the CDR3 sequence on a llama VH(4)
as amino acids as those of the llama VH(4). For the second possibility,
a strategy could be developed whereby a synthetic library would be
constructed by randomizing the CDR3 sequence on a llama VH(4)
as amino acids as those of the llama VH(4). For the second possibility,
a strategy could be developed whereby a synthetic library would be
constructed by randomizing the CDR3 sequence on a llama VH(4)
as amino acids as those of the llama VH(4). For the second possibility,


