Solution Structure of a Llama Single-Domain Antibody with Hydrophobic Residues Typical of the VH/VL Interface[†]

Wim Vranken,[‡,§] Dmitri Tolkatchev,[‡] Ping Xu,[‡] Jamshid Tanha,[||] Zhigang Chen,[‡] Saran Narang,[||] and Feng Ni*,[‡]

*Biomolecular NMR Laboratory, Biotechnology Research Institute, National Research Council of Canada,
6100 Royalmount Avenue, Montreal, Quebec, Canada H4P 2R2, and Institute for Biological Sciences,
National Research Council of Canada, Ottawa, Ontario, Canada K1A 0R6*

ABSTRACT: The three-dimensional structure of a llama single-domain antibody BrucD4-4 was established by use of solution NMR spectroscopy. BrucD4-4 has Val, Gly, Leu, and Trp residues at positions 37, 44, 45, and 47, which are considered to be a hallmark to distinguish llama VH from V$_H$H fragments at the germline level. In contrast to the murine and human VHs, BrucD4-4 has sufficient solubility, is monomeric in solution, and displays high-quality NMR spectra characteristic of well-structured proteins. Amide proton/ deuterium exchange and the ^{15}N relaxation data showed that BrucD4-4 has a classic protein structure with a well-packed core and comparatively mobile surface loops. The three-dimensional architecture of BrucD4-4 is analogous to that of VHs from murine and human F$_v$s and camelid V$_H$Hs with two pleated β-sheets formed by four and five β-strands. A canonical and undistorted β-barrel exposes a number of hydrophobic residues into the solvent on the surface of the three-dimensional structure. The eight-residue H3 loop folds over the side chain of Val37 similarly to that in llama V$_H$Hs; however, this interaction may be transient due to the H3 conformational flexibility. Overall, the surface characteristics of BrucD4-4 with respect to hydrophobicity appear to lie between the human VH domain from Fv Pot and the llama V$_H$H fragment HC-V, which may explain its enhanced solubility allowing NMR structural analysis.

Antibody-derived antigen recognition units are markedly more advantageous for biotechnological, therapeutic, and diagnostic applications as compared to monoclonal antibodies (*1−3*). The most stable and commonly used antibody fragments are the so-called F$_{ab}$s, which consist of the light chain with the variable (VL)[1] and the constant (CL) domains plus the variable (VH) and the first constant (CH1) domains of the heavy chain (*4*). The F$_{ab}$ fragment can be reduced to a VH/VL heterodimer, or F$_v$, the smallest antibody fragment still conferring complete antigen-binding capacity (*5*). The stability of the noncovalently associated Fv fragments is often compromised somewhat due to the loss of the CH1 and CL domains and a disulfide linker (*6, 7*). The F$_v$s can be stabilized through the introduction of engineered disulfide bonds or polypeptide bridges linking VH and VL domains (*6, 8, 9*). The latter form of the F$_v$ fragments, referred to as single-chain antibodies, or scF$_v$s, has found numerous applications as therapeutic and/or diagnostic agents (*10−12*).

The antibody F$_v$ fragments can be further decreased in size by splitting them into the individual VH and VL domains. The concept of single-domain antibodies (dAbs) was first introduced by Ward et al. (*13*), who successfully expressed and isolated VH fragments from mice with good antigen-binding affinities. Unfortunately, the VH fragments out of the context of the full-length antibodies or at least separated from the VL fragments are often insoluble and prone to aggregation and display nonspecific binding properties (*13−15*). The discovery of camelid heavy-chain antibodies naturally devoid of light chains opened up a new opportunity to develop novel dAbs with improved solution properties (*16*). The single-domain antibody fragments derived from this class of antibodies (or V$_H$H fragments) are highly soluble and stable in solution (*17−21*). Their small size, stability, and ability to recognize novel epitopes, such as enzyme active sites (*22*), that are not accessible to conventional antibodies, offer additional prospects in biotechnology applications (*1, 23*).

In contrast to V$_H$Hs, VH fragments, particularly those based on the human VH framework, are expected to have certain advantages in pharmaceutical applications. Therefore, superior performance of V$_H$H antibody libraries has led to the idea and practice of camelisation, i.e., replacement of residues Gly44, Leu45, and Trp47, located on the VL side of VHs, with Glu44, Arg45, and Gly47/Ile47, partially mimicking well-behaved V$_H$H fragments (*14*), to improve the solubility and prevent nonspecific binding of human VHs (*14, 24, 25*). It has been assumed that the procedure of camelisation increases the solubility of the VH fragments through increase of the hydrophilicity of the VL interface region or more subtly by induction of local conformational changes (*26, 27*). Overall, however, the camelisation approach remains somewhat limited, since it often does not eliminate tendencies toward dimerization and aggregation of normal VH fragments (*14, 28, 29*).

* Corresponding author. E-mail: Feng.Ni@nrc.ca.
[‡] Biotechnology Research Institute.
[||] Institute for Biological Sciences.
[§] WV and DT contributed equally to this work.
[1] Abbreviations: dAb, single-domain antibody; VH, variable domain from the heavy chain of a conventional antibody; VL, variable domain from the light chain of a conventional antibody; V$_H$H, the VH fragment from a camelid heavy chain antibody; Fv, an antibody fragment consisting of a VH and a VL domains; CDR, complementary determining region; DSC, differential scanning calorimetry.

The available information on the three-dimensional organization of dAb molecules is mostly represented by the X-ray structures of $V_H H$ fragments, free and in complex with antigens (*26, 27, 30−32*). The $V_H H$ scaffolds were shown to adopt a β-barrel fold similar to typical VH domains of conventional immunoglobulins, but the hypervariable regions H1 and H2 often deviate from the canonical predicted structures (*33*). In addition, H3 loops of $V_H H$ fragments partly fold over the side corresponding to the interacting surface between the VH and VL domains of conventional antibodies. On the other hand, much less is known about the solution structure of single-domain antibodies in general and the structural properties of isolated VH fragments in particular. To date, a single NMR solution structure of a camelised human VH fragment, or VH-P8, has been reported (*34*). The molecule displayed in aqueous solution limited solubility and partial aggregation as shown by rather small transverse proton relaxation times (∼16 ms), which was somewhat improved only by the addition of the detergent CHAPS (*14*). Another camelised human VH fragment, VH-P1, displayed better solution behavior, or prolonged transverse proton relaxation times in the absence of detergent, but was not available for structural characterization due to its low yield of expression (*14*). In the presence of the CHAPS detergent, the solution structure of VH-P8 is essentially similar to that of the noncamelised, VL-associated VH domains. However, the former VL interface in VH-P8 was distorted as compared to VL-associated VHs, causing the disruption of a regular β-structure at residues 37 and 38. Therefore, it appears that camelisation, while intended for improving VH solubility and binding specificity of VH molecules, might in some cases introduce structural deformations affecting the stability and performance of camelised VH libraries (*35*). In addition to this earlier study, partial NMR assignments and high-quality [^{15}N, ^{1}H]-HSQC spectra of a llama $V_H H$ (*19*) and members of a camelised human VH library (*25*) have been reported recently. The quality and resonance line widths of the presented [^{15}N, ^{1}H]-HSQC spectra are characteristic of well-structured proteins of this size.

Recently, we isolated from a llama dAb library a set of anti-idiotypic dAbs to two carbohydrate-specific antibodies, Yst9.1 and 1B1 (*36*). Panning the library against Yst9.1 and 1B1 yielded a number of anti-idiotypic dAbs enriched with the amino acid tetrad Val37, Gly44, Leu45, and Trp47 (Kabat numbering (*37*)), or the VGLW tetrad, which is considered to be a hallmark to distinguish the VH from the $V_H H$ imprinted at the germline level in camelidae (*38−41*). Three members of these anti-idiotypic "conventional VH" sequences, 1B1E10, BrucE3-1, and BrucB8, along with another representative from the same library, C44, were analyzed and confirmed to originate from the conventional llama H_2L_2 antibodies (*36*). Surprisingly, a number of anti-idiotypic dAbs, namely BrucB3, BrucB10, BrucD4-4, BrucC6, and 1B1E10, all with the VGLW tetrad, were produced in high yields in *Escherichia coli* and were found to be highly soluble. One of the anti-Yst9.1 dAbs, BrucD4-4, was chosen for structural characterization due to its solubility and *E. coli* production optimal for isotope labeling and NMR studies. BrucD4-4 also has a VGLW tetrad and is highly homologous to the BrucE3-1, BrucB8, C44, and 1B1E10 VH clones.

MATERIALS AND METHODS

Protein Preparation. Bacterial expression and purification of the dAb fragments were essentially the same as described by Tanha et al. (*25*). Single clones were used to inoculate 25 mL of LB medium containing 100 μg/mL of ampicillin, and the culture was shaken overnight at 240 rpm at 37 °C. The grown cells were transferred into 1 L of M9 medium supplemented with 5 μg/mL of vitamin B1, 0.4% casamino acid, and 100 μg/mL of ampicillin. The cell culture was shaken at room temperature for 30 h at 180 rpm and subsequently supplemented with 100 mL of 10X induction medium (Terrific Broth containing no salts) and 100 μL of 1 M isopropylthio-β-D-galactoside. After induction, the culture was shaken for another 60 h, and the periplasmic fraction was extracted by the osmotic shock method (*42*). The presence of dAb fragments in the extract was confirmed by Western blotting (*43*). The dAb fragments were affinity-purified using HiTrap Chelating column (Amersham Pharmacia Biotech) as described previously (*43*). The dAb molecules were eluted in 10 mM HEPES buffer, 500 mM NaCl, pH 7.0, with a 10−500 mM imidazole gradient and extensively dialyzed against 10mM HEPES buffer, 150mM NaCl, 3.4 mM EDTA, pH 7.4.

For NMR experiments, the BrucD4-4 sample was uniformly labeled with the ^{15}N and ^{13}C isotopes by growing cells in the ^{15}N/^{13}C-labeled rich medium Bio-Express-1000 (Cambridge Isotope Labs). Six milliliters of LB medium were inoculated with a single colony and shaken at 260 rpm and 37 °C until an OD_{600} of 5 was reached. The cells were collected by centrifugation and resuspended in 3 mL of sterile PBS. Aliquots were transferred to 25 mL of Bio-Express medium supplied with 100 μg/mL of ampicillin in 125 mL flasks to give an OD_{600} of 0.06 and incubated at 37 °C and 200 rpm for 9−10 h. Single-domain antibody fragments were purified as described above. For the NMR samples, purified unlabeled or ^{15}N/^{13}C labeled BrucD4-4 in 10mM HEPES buffer, 150mM NaCl, 3.4 mM EDTA at pH 7.4, were transferred into 10 mM sodium phosphate buffer, 150 mM NaCl, 0.2 mM EDTA at pH 6.8, by means of buffer exchange using an Amicon Centriprep YM-10 concentration cell. Amide proton/deuterium exchange was followed after dilution of the NMR sample with an equal volume of the same sodium phosphate buffer in D_2O at room temperature. If not specified otherwise, the protein concentration was ∼0.5 mM.

DSC Measurements. Purified dAb fragments were dialyzed overnight against 500 volumes of 10mM HEPES buffer, 150mM NaCl, 3.4 mM EDTA, pH 7.4, with one buffer change. Dialyzed dAb fragments were diluted with the same buffer to specified protein concentrations. DSC measurements were carried out on VP-DSC (sample volume 0.5197 mL) and MCS-DSC (sample volume 1.4 mL) calorimeters (MicroCal) at a scan rate of 60°/hour.

NMR Spectra and Assignments. NMR spectra were recorded on Bruker Avance 800 and 500 MHz spectrometers at 303 K. The complete data set included HNCA, HNcoCA, HBHAcoNH, HNCACB, CBCAcoNH with gradient enhancement, an HCCH-TOCSY, 3D ^{15}N-edited NOESY-HSQC, 3D ^{13}C-edited NOESY-HSQC, side-chain ^{13}C-edited 3D NOESY-HSQC, and a 3D ^{15}N/^{13}C-edited NOESY-HSQC (*44*). The 3D ^{15}N/^{13}C-edited NOESY-HSQC was recorded with spectral widths of 9980 Hz (F1, ^{1}H), 6640 Hz (F2, ^{15}N/

^{13}C), and 9980 Hz (F3, ^{1}H) and with 154, 74, and 2048 sampling points along the t_1, t_2, and t_3 time dimensions, respectively. The NOE mixing time was 120 ms. Side-chain ^{13}C-edited 3D NOESY-HSQC with a mixing time of 200 ms was carried out as a $80 \times 21 \times 512$ complex data matrix. The ^{13}C carrier frequency was set to 121.5 ppm, and the spectral width in the ^{13}C dimension was 24 ppm. Cosine-squared window functions were used along all time dimensions for transformation of the data matrixes into the spectral domain. Prior to Fourier transformation, linear prediction was performed to 256 and 148 points for the t_1 and t_2 dimensions, respectively. After Fourier transformation, the baseline of the F3 dimension was corrected by a polynomial function. Spectra were processed with NMRPipe (*45*) and converted to the format used by NMRView (*46*) for further analysis.

The backbone ^{1}H, ^{15}N, and ^{13}C chemical shift assignments were derived from 2-dimensional [^{15}N,^{1}H]- and [^{13}C,^{1}H]-HSQC spectra, combined with the set of three-dimensional HNCA, HNcoCA, HNCACB, and CBCAcoNH spectra. Further assignments of the side-chain resonances were obtained through the analysis of the HBHAcoNH and HCCH-TOCSY spectra. The ^{1}H assignments of the aromatic side chains were derived from the side-chain ^{13}C-edited 3D NOESY-HSQC spectrum, combined with a 3D^{15}N/^{13}C-edited NOESY spectrum and a [^{13}C,^{1}H]-HSQC spectrum of the aromatic region. During the structure calculation process, several assignments were added or changed on the basis of the side-chain ^{13}C-edited 3D NOESY-HSQC and 3D^{15}N/^{13}C-edited NOESY spectra. The complete list of assignments has been deposited to BMRB (deposition No. BMRB-4969).

Heteronuclear {^{1}H}-^{15}N NOE intensities and ^{15}N T_1 and T_2 relaxation rates were measured using pulse sequences from Farrow et al. (*47*). Decoupling of ^{13}C$^\alpha$ and ^{13}CO during the evolution time was achieved by a 180° hard pulse and a 600 Hz SEDUCE-1 sequence, respectively. To measure ^{15}N T_1, we carried out 14 experiments with relaxation delays of 105, 35, 280, 5, 810, 70, 145, 640, 205, 380, 500, 1010, 1250, and 70 ms in the specified sequence. A 400 μs REBURP 180° selective pulse was used to saturate the NH proton resonances. To measure ^{15}N T_2, we carried out 11 experiments with relaxation delays of 14.4, 57.6, 100.8, 43.2, 86.4, 28.8, 129.6, 43.2, 72.0, 158.4, and 144.0 ms. To measure heteronuclear NOE intensities, a relaxation delay of 2 s prior to a 3 s of saturation of the NH proton resonances was used, while a relaxation delay of 5 s was used to record the reference spectrum.

Structure Calculations. Structure calculation was performed in two stages. The first round of calculations was focused on the generation of global topology via automatic NOE peak assignment using ARIA (*48*). In the second round, manual assignment of NOE peaks with low ambiguity was performed and used along with 108 dihedral angle and 56 hydrogen bond constraints in the subsequent ARIA runs.

The peak positions in the [^{15}N,^{1}H]- and [^{13}C,^{1}H]-HSQC spectra enabled hand-driven peak-picking of the combined ^{15}N/^{13}C 3D NOESY and side chain ^{13}C-edited 3D NOESY-HSQC. The resulting peak list was then filtered on the basis of possibilities and probabilities to obtain an initial list of NOE peaks with low ambiguity (less than four assignment possibilities for each overlapped NOE peak). Assignment probabilities of ambiguous restraints were divided by the total number of assignment possibilities for that restraint. Probable

intraresidue NOEs were removed from the list, and all kept NOE peaks were converted to a set of loose distance restraints with lower and upper bounds of 1.8−5.0 Å. ARIA (*48*) runs using the Crystallography and NMR System (CNS) (*49*) with these restricted set of distance restraints along with chemical shift index-based dihedral angle constraints (see below) resulted in a set of structures defining the global fold of the protein. These global folds served as an initial structure-based assignment filter for the complete list of NOE peaks. A second set of ARIA runs, also using the dihedral restraints, was then performed to remove assignment errors, noise, and other artifacts from the complete NOE peak list. The first iterations had a high ambiguous cutoff to allow greater conformational variability and to decrease the probability of misassignments. The ambiguous cutoff was lowered to increase structure convergence in the later runs when an improved restraint list became available. The backbone dihedral angles for the regions determined as β-sheet were restrained to (-135 ± 50) for the φ angle and (135 ± 50) for the ψ angle in the initial runs. In the final runs, the internal ARIA routine for setting the distance limits based on NOE peak intensities was enabled. Also, somewhat more stringent restraints of (-135 ± 30) for the φ angle, and (135 ± 30) for the ψ angle, were applied for some of the residues involved in the regular β-sheets identified on the basis of chemical shift index, manually assigned backbone-backbone NOE connectivities, pattern of NH proton exchange rates, and ^{15}N relaxation parameters. A disulfide bond constraint between residues Cys22 and Cys92 was employed throughout the calculations. This final computation resulted in the set of structures presented here. The final list of distant restraints contained 1471 unambiguous and 944 ambiguous restraints divided into 618 sequential, 549 medium range ($|i - j| = 3$), and 1248 long-range contacts. Analysis of the ARIA results was carried out using the in-house developed aria.overview.pl perl script, which is publicly available and can be obtained from Wim Vranken (wim@ebi.ac.uk) or from Feng Ni (Feng.Ni@nrc.ca). The final structures (PDB entry 1IEH) were analyzed and visualized with the Sybyl, InsightII (Biosym) and Molmol software programs (*50*).

RESULTS

Thermal Unfolding and Transition Temperature (T_m) for Anti-Idiotypic dAbs. Five anti-idiotypic dAb molecules, BrucD4-4, 1B1E10, BrucB3, BrucC6, and BrucC7.2, were chosen to assess their thermal stability (Table 1). The 1B1E10, BrucD4-4, BrucB3, and BrucC6 molecules contain Val, Gly, Leu, and Trp at positions 37, 44, 45, and 47, respectively. In BrucC7.2, residues Val37 and Trp47 are replaced with Phe37 and Tyr47. These five dAbs also vary somewhat in the length of the CDR3 loops, containing 8−10 amino acid residues.

The five protein samples were tested at three different concentrations, high (ranging from 0.31 to 0.47 mg/mL), medium (approximately 0.11 mg/mL), and low (approximately 38 μg/mL) using differential scanning calorimetry (DSC). At high concentrations, all protein samples except BrucC7.2 were found to aggregate and precipitate upon thermal unfolding, as judged by a sharp drop in the post-transition baseline. At a concentration of approximately 38 μg/mL, the post-transition baseline of all samples was normal, suggesting that protein aggregation was negligible,

Table 1: Sequence Comparison and Thermal Transition Temperatures (T_m) of Some Anti-Idiotypic dAbs[a]

	1(1)	10(10)	20(20)	30(30)	40(40)	50(50)
BrucD4-4	DVQLQASGGGLVQPGGSLRVSCAAS**GFTFS**SYHMAWVRQAPGKGLEWVSTINPGDGSTYY
1B1E10	DVQLQASGGGLVQPGGSLRVSCAASGFTFSNYHMVWVRQAPGKGLEWVSTINIDGGTTYY
BrucB3	DVQLQASGGGLVQPGGSLRLSCAASGFTFSSYAMSWVRQAPGKGLEWVSGIEGGGGITRY
BrucC6	DVQLQASGGGSVQPGGSLRVSCAASGFTFSTYAMTWVRQAPGKGLEWVSTINTSGRGTYY
BrucC7.2	DVQLQASGGGSVQAGGSLRLSCAASGFTFDEHAIGWFRQAPGKGLEYVSHID-TGGSTWY

	60(61)	70(71)	80(81)	90(94)	100(104)	110(114)	Tm, C
	ADSVKGRFTISRDNAKNTLYLQMNSLKSEDTAVYYCAKYSGGA--LDAWGQGTQVTVSS	58.5
	ADSVKGRFTISRDNAKNTLYLQMNSLRPEDTALYYCVRDQGGTR-YDYWGQGTQVTVSS	57.0
	ADSVKGRFTISRDNARNTLYLQMNSLKPEDTALYYCARAHGGYGAFGSWGQGTQVTVSS	56.7
	ADSVKGRFTASRDNAKNTLYLQMNSLKPEDTAVYYCAAQGYAGS-YDYWGQGTQVTVSS	58.5
	AASVKGRFTVSRDDAKNTLYLQMNSLKPEDTGLYYCARLSQGA--MDYWGKGTQVTVSS	74.2

[a] CDR residues are in bold, and H-loop residues are underlined (*37*, *63*). The hallmark residues at positions 37, 44, 45, and 47 are shaded. The sequence numbering is from Kabat et al. (*37*). The numbering in the parentheses follows the sequential order of BrucD4-4 used for the purpose of NMR spectral analyses (see Figures 1, 2, and 5).

and it was possible to establish the temperature (T_m) values of unfolding transition (Table 1). The thermal unfolding for these proteins could be considered as reversible at concentrations of 38 μg/mL. At the protein concentration of 0.108 mg/mL, only BrucB3 displayed a sharp drop in the post-transition baseline typical of aggregation upon unfolding.

The transition temperatures of dAb molecules correlate with the protein concentrations allowing reversible unfolding. BrucC7.2 has the highest unfolding transition temperature, and the unfolding was reversible up to the concentration of 0.47 mg/mL. BrucB3 displays the lowest transition temperature, and marked aggregation was observed at the concentration of 0.108 mg/mL. BrucC6, 1B1E10, and BrucD4-4 molecules demonstrated an intermediate tendency toward aggregation upon thermal unfolding. The thermal transition curve of BrucC7.2 at 0.118 mg/mL was subjected to further analysis and was found to fit to a single two-state unfolding mechanism (data not shown).

Hydrodynamic Radius of BrucD4-4. We measured the molecular dimensions of the BrucD4-4 molecule in solution to characterize the folding and possible aggregation states of the molecule. We used pulse field gradient (PFG) NMR techniques (*51*) to establish the effective hydrodynamic dimension of the BrucD4-4 dAb (*52*). A series of one-dimensional proton spectra were recorded by the PG-SLED pulse sequence, and the observed hydrodynamic radius was calculated as described by Wilkins et al. (*52*). Values of the hydrodynamic radius were calculated in reference to acetate anion, which was included in the protein sample. The effective hydrodynamic radius of 2.3 Å for acetate was determined from that of hen lysozyme at different concentrations (*51*, *52*).

The hydrodynamic radius for BrucD4-4 was slightly dependent on the sample concentration and was determined to be 22.9, 20.2, and 19.2 Å at protein concentrations of 0.90, 0.45, and 0.1 mM, respectively, and with identical concentrations for all other solution components. These values for the hydrodynamic radii are close to the empirically expected value of 19.7 Å for a native protein of the size of BrucD4-4 (*52*). The slight concentration dependence of the observed hydrodynamic radius indicates that BrucD4-4 is primarily monomeric in solution with only some residual aggregation at concentrations equal to or above ∼0.5 mM.
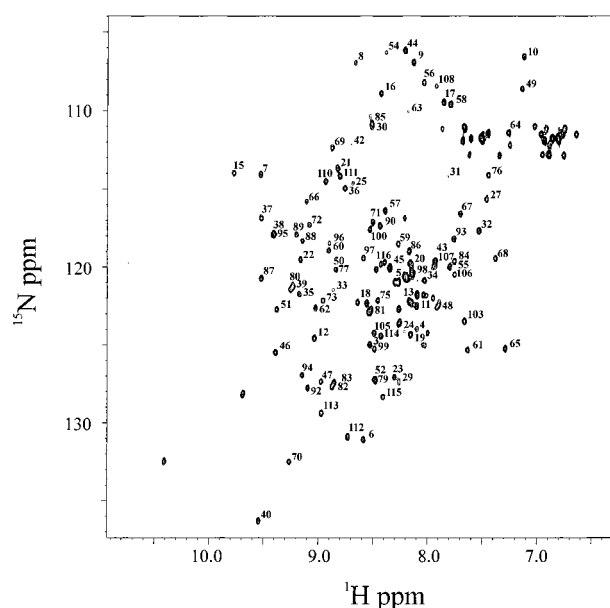

FIGURE 1: [^{15}N,^{1}H]-HSQC spectrum of ^{15}N/^{13}C-labeled BrucD4-4. ^{15}N/^{13}C-labeled BrucD4-4 was prepared in 10 mM sodium phosphate buffer, 150 mM NaCl, 0.2 mM EDTA, and 10% D$_2$O at pH 6.8. Protein concentration was approximately 0.5 mM. The NMR spectrum was recorded on a Bruker Avance 800 MHz spectrometer at 303 K. The assignments of the backbone HSQC peaks are indicated by the residue numbers following the sequence of BrucD4-4 (Table 1).

NMR Spectra and Assignments. At approximately 0.3–0.4 mM, llama single domain antibodies 1B1E10, BrucB3, BrucD4-4, and BrucC7.2 exhibit similar 1D proton NMR spectra with comparable line widths in 10 mM sodium phosphate buffer, pH 6.8, 150 mM NaCl, 0.2 mM EDTA, at 303 K, suggesting little protein aggregation in these conditions (data not shown). Nearly complete assignments for all the ^{1}H, ^{15}N, and ^{13}C chemical shifts (except for the backbone carbonyl atoms) were obtained from the set of three-dimensional NMR spectra recorded for two samples of ^{15}N/^{13}C-labeled BrucD4-4, first at a concentration of ∼0.7 mM and the second at ∼0.5 mM (see Materials and Methods) (Figure 1). The high quality of the NMR spectra also allowed the identification of most resonances belonging to residues in CDR and other loops, with the exception of the HN/N atoms for Gly26, Thr28, Asn73, Asn76, and Gly97. In

addition, it was not possible to assign the resonances of the Gln114 and His129-His131 residues because of heavy resonance overlaps.

Hydrogen−Deuterium Exchange. We determined most of the slowly exchanging amide protons through dilution of the BrucD4-4 sample with buffered D_2O at pH 6.8, 298 K. The D_2O dilution method is a milder and safer alternative to the D_2O redissolution of a lyophilized protein sample since the risk of solubilization artifacts or protein aggregation is minimized. Immediately after dilution, the course of hydrogen−deuterium exchange was followed by recording a series of [^{15}N,^{1}H]-HSQC spectra. We found that amide protons of the residues Val12, Leu18, Val20-Cys22, Ala35-Gln39, Glu46, Ser49-Ile51, Ala60, Phe67, Thr68, Ser70, Tyr79, Gln81, Lys83, Asp87, Ala88-Tyr91, Ala93, Thr107, and Val109-Val111 exchange with the D_2O solvent slowly. The list of slowly exchanging protons is not limited to the indicated residues, but spectral overlaps prevented identifying a few other solvent-protected amide protons. The identified slowly exchanging protons correspond well to the ones forming hydrogen bonds expected for the β-structure typical of the VH domain architecture (*53*). The clearly resolved HSQC peaks of the amino acid residues located at the C-terminal tail and loops connecting β-strands were exchanged even before the first NMR spectrum could be recorded.

^{15}N NMR Relaxation Measurements of BrucD4-4. To assess the motional properties of the BrucD4-4 backbone, we recorded a set of ^{15}N relaxation data including T_1 and T_2 relaxation rates and heteronuclear {^{1}H}-^{15}N NOE intensities (Figure 2). Since in the ^{15}N- and ^{13}C-labeled sample the relaxation properties of the amide ^{15}N nuclei are affected by their interaction with the adjacent C^α and CO, we did not perform a detailed analysis of the obtained data, and extracted only general trends for dynamic behaviors of the protein. Specifically, {^{1}H}-^{15}N NOE intensities of the expected β-strand residues fluctuate around the value of ~0.8 typical of a well-folded protein (*54*). Weaker heteronuclear NOEs (0.4−0.7) are exhibited by the loops connecting the β-strands, including the H-loops, suggesting an elevated degree of mobility and a poorer degree of structure definition in these regions. The C-terminal tail (residues Gln114−His131) displays even lower NOE intensities, reaching negative values for residues Ser121 and His128. In addition, the T_2 values for the tail residues are larger than 200 ms (not shown), suggesting very high mobility of these residues.

Three-Dimensional Structure of BrucD4-4. The recombinant BrucD4-4 molecule includes the sequence of the corresponding llama VH domain (residues Asp1-Ser113), plus a tail region including the c-myc and the His$_5$ tags, or residues Gln114-His131 (Figure 3). Initially, the expected β-strand regions Val2−Ser7, Gly9−Val12, Leu18−Ser25, Ala35−Val37, Ser49−Asn52, Ser56−Tyr59, Phe67−Asp72, Thr77−Met82, Tyr91−Tyr95, and Thr107−Val111 were identified on the basis of chemical shift index (*55*) of the C^α, C^β, and H^α chemical shifts and a NOE-based contact map. These secondary structures correspond well to the β-sheet regions for the core β-barrel motif of a VH domain, i.e., residues 3−7, 11−13, 18−25, 33−40, 49−52, 56−60, 66−73, 76−82a, 88−95, 101−104, and 107−112 (*53*). After the first round of ARIA calculations (see Materials and Methods), backbone−backbone NH−NH, NH−C^αH, and
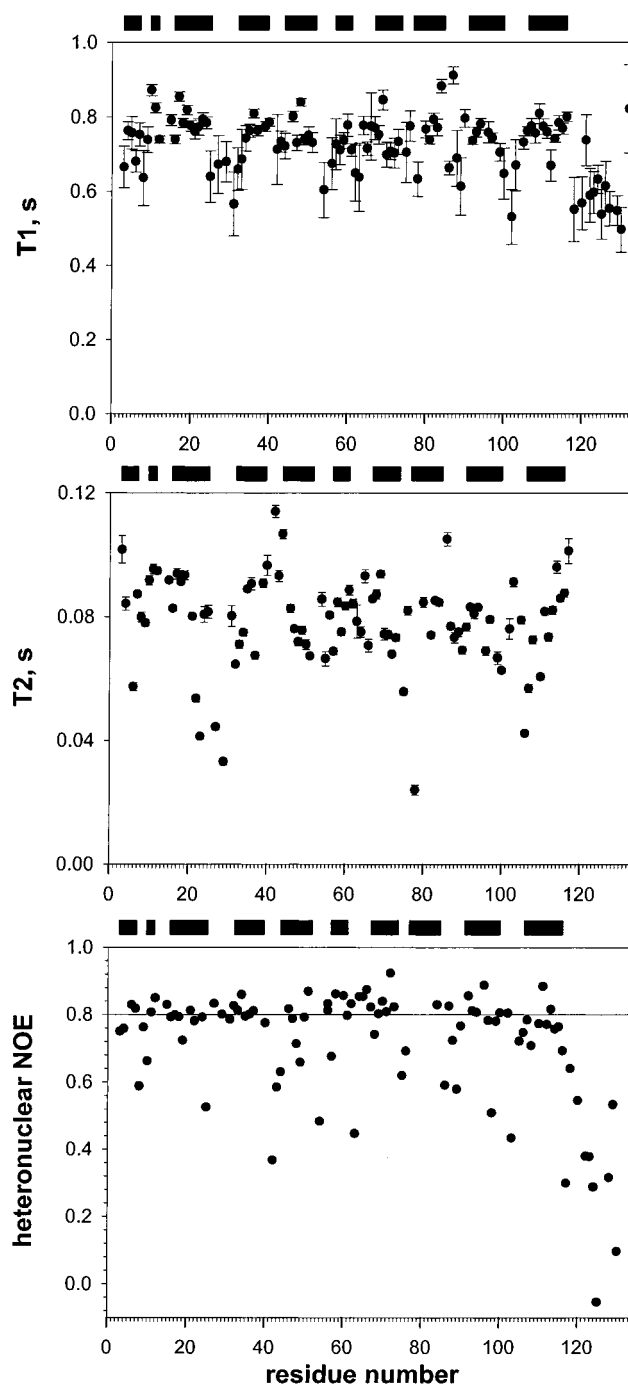
FIGURE 2: Backbone ^{15}N NMR relaxation parameters for BrucD4-4. (a) T_1 relaxation times, (b) T_2 relaxation times, and (c) steady-state {^{1}H}-^{15}N NOEs. Bars represent β-barrel amino acid residues 3−7, 10−12, 16−25, 32−40, 45−52, 56−60, 66−73, 76−82b, 87−96, and 101−112 (*34*). Other experimental conditions are the same as those in Figure 1.

C^αH−C^αH NOE connectivities were inspected, edited, and verified using the original NMR spectra. The pattern of backbone−backbone NOE connectivities also provided further evidence for the formation of β-sheet secondary structures between strands Gln5−Ser7 and Ser21−Ala24, Ser17−Ala24 and Thr77−Ser82b, Thr77−Asn82a and Phe67−Asp72, Gly10−Val12 and Val109−Ser112, Thr107−Val111 and Thr87−Tyr90, Met34−Gln39 and Ala88−Tyr95, Met34−Ala40 and Leu45−Asn52, and Thr50−Asn52 and Thr57−Tyr58. This structure pattern is in agreement with the expected organization of β-strands for a VH domain and

```
        1    5      10      15      20      25      30      35
D4.4  D V Q L Q A S G G G L V Q P G G S L R V S C A A S [G F T F S S Y]H M A

HC-V              E              A          L              R G T   D G
VH-P8    E       V E      N                 L                    A   S
FvPot    E   H   L E                        L                N I F V   S

        36   40      45      50   52   55      60      65      69
                                   a
D4.4  W(V R)Q A P G K(G L E)(W)V S(T)I[N P G D G S]T Y Y A D S V K G R F T I

HC-V       F                E R   S   A A      W D S A R      S      R
VH-P8      F                E R   I   A V S -   S   G
FvPot                                 G/V F -   S   G N  D/   A

        70   75      80   82      85      90      95      100
                          a b c
D4.4  S R D N A K N T L Y L Q M N S L K S E D T A V Y Y C A K[Y S G G A L D]

HC-V        ~ ~ ~ ~        K   V              P          T   G A G E    T W
VH-P8             S                           S A                R L K K Y  F
FvPot   T        S                            R A          I              HRVSYVLTGF

        102  105     110     115     120     125     130
D4.4  [A]W G Q G T Q V T V S S Q S E Q K L I S E E D L N H H H H H

HC-V   S
VH-P8  Y
FvPot  S                    L
```
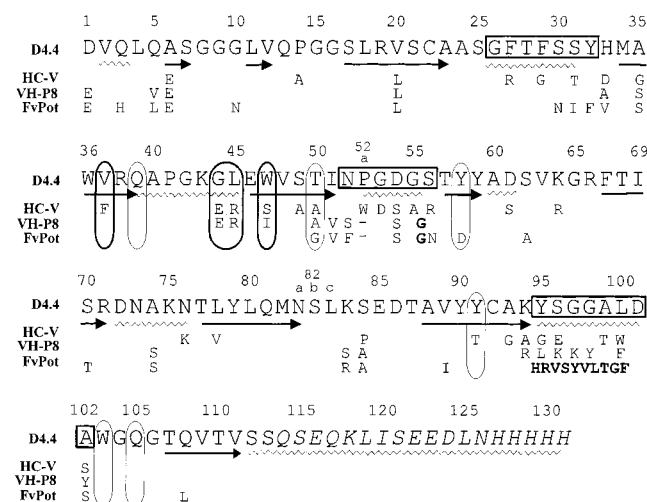
FIGURE 3: Comparison of the BrucD4-4 sequence with representative sequences (indicated underneath) from a llama $V_H H$ domain (first row, HC-V), a camelised VH domain (second row, VH-P8 human), and a VH domain from an Fv fragment (third row, Fv Pot human). Bold residues are insertions, and dashes are deletions. The hypervariable loops H1, H2, and H3 (37, 63) are boxed. VH residues interacting with the putative VL are encompassed in ovals (gray shading indicates the hallmark residues Val37, Gly44, Leu45, and Trp47) (57). C-myc and His-tag residues are shown in italics. β-Strand residues (as defined by Procheck) are indicated by arrows while poorly defined residues (local backbone RMSD > 0.3 Å) by a zigzag line.

also with the experimentally determined network of slowly exchanging amide protons.

A set of 60 structures of BrucD4-4 was calculated (Table 2). Figure 4 shows the cluster of 10 best structures superposed using the backbone atoms of the residues forming the β-barrel. The local pairwise RMSD values and the angular order parameters for those structures (Figure 5) indicate that the global three-dimensional folding is well defined by the NMR data. The less well-defined regions are within the H loops (Gly26−Ser31, Pro52a−Gly55, Tyr95−Asp101), in other loop or turn regions (Gln39−Leu45,

Table 2: Overview of the CNS/ARIA Structure Calculation Parameters of the 10 Best Structures (Figure 4) for the BrucD4-4

mean energy terms	
total CNS energy (kcal/mol)	251 ± 23
VdW energy (kcal/mol)	110 ± 8
distance constraints energy (kcal/mol)	62 ± 13
rms deviations from ideal geometry	
bond lengths (Å)	0.002 ± 0.001
bond angles (deg)	0.335 ± 0.016
improper angle (deg)	0.239 ± 0.022
distance constraints	
average rms deviation for NOE	0.030 ± 0.005
restraint violations (Å)	
total number of violations > 0.1 Å	1

Ala60−Asp61, Asp72−Asn76), and for the C-terminal residues (Ser112−His131). The high structural variability of the tail region is especially obvious from the comparison of the average pairwise RMSD values for BrucD4-4 with the tail (Asp1−His131) and without it (residues Asp1−Ser113) (Table 3). There is also a general lack of medium-range NOE contacts and secondary structures (based on chemical shift values) for this tail (residues Gln114-His131) region. In addition, no long-range NOE contacts were found between the tail and the globular part of the protein. The tail region of BrucD4-4 can therefore be considered as being conformationally flexible in solution, with no apparent interaction with the VH part of the molecule, and will be excluded from the following analysis of the solution structure.

The general architecture of the BrucD4-4 molecule is typical of the VH domains with two pleated β-sheets formed by four and five β-strands. The structural variability in the H-loop regions for BrucD4-4 makes it impossible to assign them according to the structure classification of the canonical loops in crystallized VH and $V_H H$ domains (33). This result for a solution structure of the VH fragment is easy to understand since the ^{15}N relaxation data and hydrogen/ deuterium exchange data (see the above sections) all indicate the presence of the extensive conformational motions in the BrucD4-4 loop regions. Consistently, the lack of distance
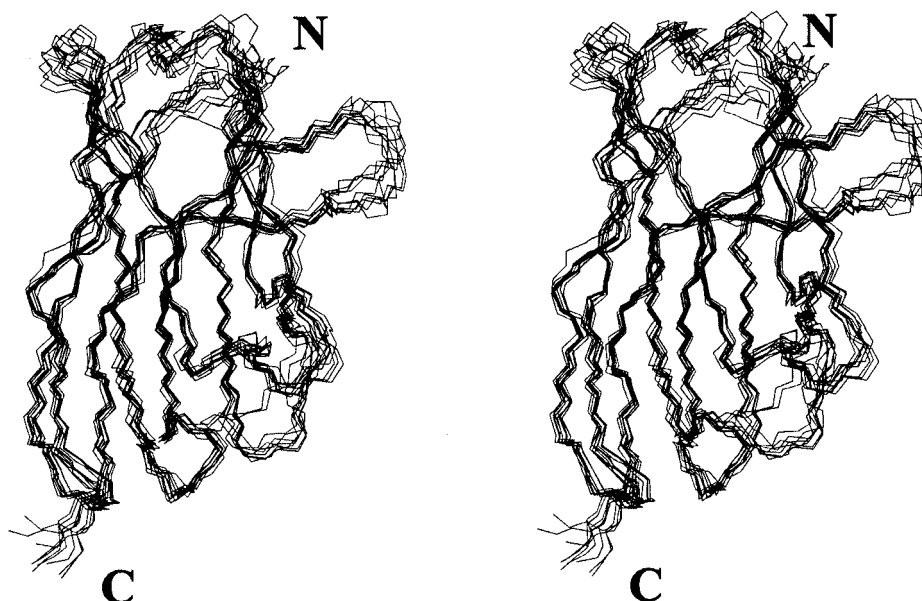

FIGURE 4: Cluster of the 10 lowest energy BrucD4-4 structures superimposed using the backbone atoms of β-barrel residues. The β-barrel residues are defined as in Figure 2. The beginning (Asp1) and end (Ser113) residues of the dAb fragment are indicated by letters N and C, respectively. The flexible C-terminal tail (residues Gln114-His131, Figure 3) is not displayed.

Table 3: Detailed Structure Comparisons between the 10 Best Structures (Figure 4) Calculated for the BrucD4-4

	all residues	V_H (1−113)	V_H, no H loops	V_H, only β-barrel
procheck analysis				
most favored	64.0%	68.8%	71.2%	79.4%
additional allowed	27.7%	24.3%	22.8%	15.2%
generously allowed	5.4%	4.4%	4.4%	3.3%
disallowed	3.0%	2.5%	1.6%	2.0%
RMSD (Å)				
all atoms	7.1 ± 2.1	1.53 ± 0.15	1.42 ± 0.15	1.21 ± 0.10
backbone	6.3 ± 2.1	1.06 ± 0.14	0.90 ± 0.17	0.67 ± 0.09

Table 4: Structural Comparison of the 10 Best BrucD4-4 Structures (Figure 4) to Previously Determined Structures of V_HH Domains and VH Fragments[a]

dAb	2−112	only β-barrel	no H loops	H1 loop	H2 loop	H3 loop
BrucD4-4	1.02 ± 0.14	0.67 ± 0.09	0.90 ± 0.17	0.76 ± 0.21	0.78 ± 0.33	0.78 ± 0.19
llama V_HH						
HC-V	2.22 ± 0.11	1.72 ± 0.08	1.55 ± 0.10	1.30 ± 0.11	0.85 ± 0.28	1.40 ± 0.14
V_HH-R2	nav	1.94 ± 0.07	1.66 ± 0.13	nav	1.67 ± 0.16	nav
camelised V_H						
VH-P8	3.47 ± 0.09	2.64 ± 0.06	2.76 ± 0.09	1.66 ± 0.21	1.84 ± 0.17	2.15 ± 0.15
VH from Fv						
Fv Pot	nav	1.69 ± 0.07	1.42 ± 0.10	1.36 ± 0.11	1.43 ± 0.18	nav
Fv 4155	nav	1.64 ± 0.06	1.51 ± 0.05	1.27 ± 0.12	nav	nav
Fv B1-8	nav	1.96 ± 0.06	1.87 ± 0.07	1.28 ± 0.14	0.95 ± 0.29	nav

[a] Average pairwise RMSD values for BrucD4-4 are included for comparison. All RMSD values are backbone comparisons, and values are in Å. "nav" indicates that comparisons are not available because of differing number of residues.
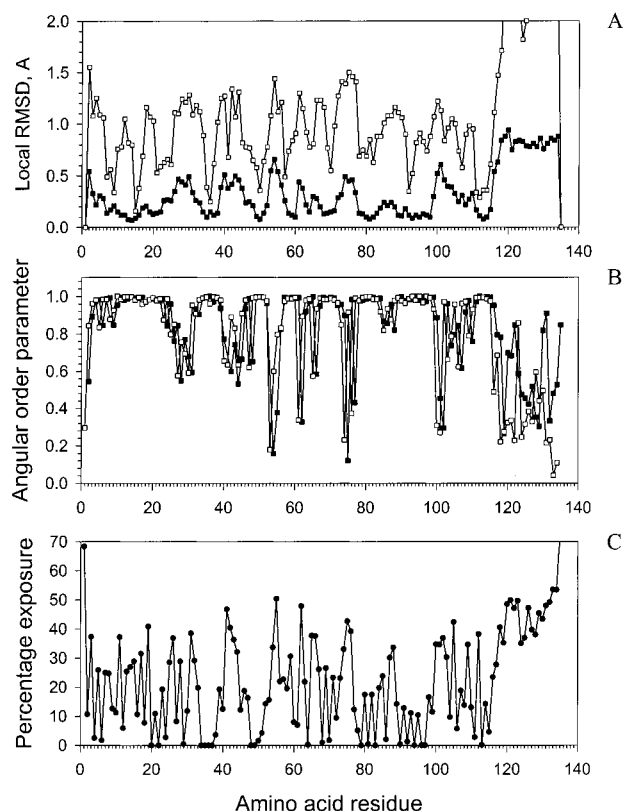

FIGURE 5: Structural parameters of the 10 best BrucD4-4 structures. (A) Local average pairwise RMSD for backbone (■) and heavy atoms (□) based on tripeptide fragments with the residue in question as the central amino acid. (B) Angular order parameters for BrucD4-4 as reported by MOLMOL for φ (■) and ψ (□) angles (*50*). (C) Average solvent exposure of BrucD4-4 residues as reported by MOLMOL, with solvent radius 1.4 and precision 3.

restraints for the H loops in BrucD4-4 is likely a result of inherent conformational flexibility. It is possible that these flexible H loops can only adopt well-defined conformations

upon binding to target molecules, a subject that will require more detailed studies of the dynamic properties of VH and V_HH domains and structures of dAb-antigen complexes.

In Table 4, the BrucD4-4 structure is compared with the previously determined structures of other V_HH and VH domains. We used for the comparison two llama V_HH proteins, free HC-V (*26*) (PDB entry 1HCV) and V_HH-R2 in complex with a hapten (*31*) (PDB entry 1QD0). In addition, we compared the structure of BrucD4-4 with a camelised human VH domain VH-P8 (*34*) (PDB entry 1VHP), and VH domains from intact Fv fragments, the human Fv Pot (*53*) (PDB entry 1IGM), Fv4155 (*56*) (PDB entry 1BFV), and Fv B1-8 from a mouse (Simon et al., PDB entry 1A6U). The BrucD4-4 structure strongly resembles those of the human Fv Pot and Fv4155, closely followed by the llama's V_HHs HC-V and V_HH-R2. Although the structure of the H loops is not sufficiently well-defined to refer them to the canonical structures found in crystallized VH and V_HH domains (*33*), it is possible to say on the basis of RMSD values that the H1 loop of BrucD4-4 is very similar to the corresponding loops of the HC-V llama V_HH and the VH domains from intact Fv, while the H2 loop matches the ones found in the V_HH domain HC-V and the VH domain from Fv B1−8. The H3 loops differ in length and can only be compared in some cases: the H3 loop from the V_HH domain HC-V is again very similar. Interestingly, the BrucD4-4 structure appears to differ most from the solution structure of the camelised human VH domain VH-P8 determined by use of NMR spectroscopy (*34*).

One of the most pronounced differences between the three-dimensional structures of the VH-P8 and VH domains in VH−VL pairs is the significant distortion of the former VL side in VH-P8 at the residues Val37 and Arg38 (*34*). Specifically for the corresponding region in BrucD4-4, unambiguously assigned NOE connectivities were identified prior to full structural analysis between the H^δ protons of
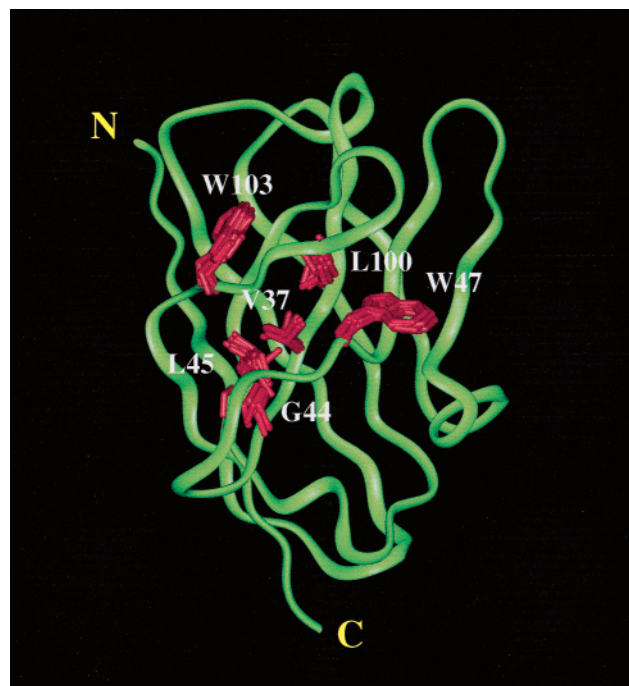
FIGURE 6: View of the side chains of residues Val37, Gly44, Leu45, Trp47, Leu100, and Trp103 forming a hydrophobic cluster in the 10 best BrucD4-4 structures. The structures are superimposed using the backbone atoms of the β-barrel amino acid residues as defined in Figure 2.

Arg38 and the H^γ protons of Val48, between the $H^{\beta,\delta}$ protons of Arg38 and the H^η proton of Tyr90, and between the $H^{\beta,\delta}$ protons of Arg38 and the H^β protons of Ala88. These contacts suggested that the side chain of Arg38 is buried in the interior of BrucD4-4, similar to that in the VH domain from Fv Pot (53). Upon formation of the VH-like fold in the first rounds of calculations, a number of ambiguous assignments were resolved by the ARIA structure-based iterative approach, revealing additional numerous contacts between the side chain of Arg38 and the side chains of residues Val48, Ala88, and Tyr90, thus confirming the interior location of Arg38. Consistently, we observed slowly exchanging amide hydrogen for Glu46, which was assigned to the hydrogen bond between residues Glu46 and Arg38 in a typical β-structure. On the putative VH−VL interface, the side chains of residues Val37, Leu45, and Trp47 in BrucD4-4 form a hydrophobic cluster, enlarged by Leu100 and Trp103 (Figure 6). The presence of the cluster is directly evidenced by the NOE contacts between the $H^{\beta,\gamma}$ protons of Val37 and the $H^{\gamma,\delta}$ protons of Leu45, between the H^γ protons of Val37 and the $H^{\beta,\epsilon1}$ protons of Trp47, between the H^δ protons of Leu45 and the $H^{\epsilon1}$ proton of Trp103, between the H^δ protons of Leu100 and the H^β protons of Trp103, and between the H^δ protons of Leu100 and the $H^{\beta,\epsilon1}$ protons of Trp47. The positions of the side chains are rather well defined and very similar to those of the corresponding residues in Fv Pot, Fv4155, and HC-V.

DISCUSSION

In this work, we present the solution structure of a llama dAb BrucD4-4, a soluble member of a llama dAb library subpopulation with amino acid sequences typical of conventional llama antibody VH domains. Four members of this subpopulation, BrucE3-1, BrucB8, C44, and 1B1E10, were directly identified by using CDR3-specific oligonucleotide probes as originating from the conventional H_2L_2 antibodies (36). Some of the expressed dAbs from this subpopulation, including BrucD4-4 and 1B1E10, displayed NMR spectra with resonance line widths characteristic of well-folded proteins with no marked aggregation. The direct identification of the origin of the BrucD4-4 sequence was not possible due to the nonspecific interaction of the CDR3-specific oligonucleotide probe with *E. coli* DNA (unpublished observations). However, BrucD4-4 has four amino acids, Val37, Gly44, Leu45, and Trp47, that are considered to be a hallmark distinguishing the camelid VH from the V_HH imprinted at the germline level (38−41), and BrueD4-4 is highly homologous to llama VH fragments BrucE3-1, BrucB8, C44, and 1B1E10 (36). Nevertheless, it is an important task to establish the solution structure of a soluble VGLW-containing llama dAb and compare it with those of other VH domains and dAb fragments.

NMR analysis showed that BrucD4-4 is essentially monomeric in solution and has practically the same three-dimensional fold as VH domains from F_v fragments and camelid V_HHs with two pleated β-sheets formed by four and five β-strands. The formation of classic and stable core of a β-barrel-structure along with flexible loops is supported by the results of structure calculations, amide hydrogen−deuterium exchange experiments, and ^{15}N NMR relaxation data. The H loops do not appear to form uniquely defined structures in solution and therefore cannot be attributed to any canonical forms. NMR resonance broadening caused by conformational exchange processes on a microsecond−millisecond time scale in the H loops was reported previously for a llama V_HH fragment (19) and designed camelised human VH fragments (25). The eight-residue H3 loop folds over the side chain of Val37 similarly to that in V_HHs; however, this interaction might be transient due to the H3 conformational mobility.

BrucD4-4 exposes on its surface a number of hydrophobic residues typical of VH domains from conventional H_2L_2 antibodies with intimately associated VH/VL domains. In VH−VL pairs, the framework VH residues Val37, Gly44, Leu45, Trp47, Tyr58, Tyr91, and Trp103 create a large hydrophobic surface for interacting with the VL domain (57). Exposure of these residues into the solvent normally leads to aggregation and stickiness of the isolated mouse and human VH molecules (13, 14). Aggregation of the isolated VH molecules in turn reduces the solubility and the proton transverse relaxation times to the values making NMR structural analysis impossible (14, 25). In the camelised human VH molecule VH-P8, some of these residues were replaced by more hydrophilic amino acids (14), which somewhat improved the solubility of VH-P8, and the determination of the three-dimensional structure by NMR was made possible in the presence of the detergent CHAPS (34). However, the VH-P8 molecule displayed backbone deformations at the former VL interface with the side chain of Val37 buried in the interior of VH-P8 and the side chain of Arg38 exerted into the solvent. This deformation caused a kink in the polypeptide chain and the breakage of a hydrogen bond between the NH of residue 46 and the carbonyl group of residue 38 (34). Contrary to VH-P8,
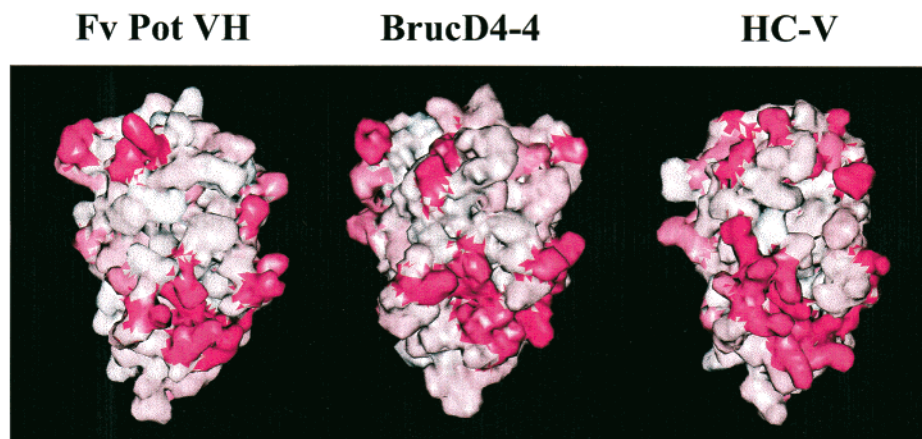
FIGURE 7: Surface plot of the VL interface of VH fragments in the lowest-energy structure of BrucD4-4 (middle), a VH fragment from an Fv Pot (left), and a llama V$_H$H fragment HC-V (right). Amino acid residues are colored according to the InsightII software definitions based on a Engelman-Steitz scale (*64*) from hydrophobic (white) to hydrophilic (red).

BrucD4-4 did not display any signs of such structure deformation.

An interesting aspect of this work is related to the BrucD4-4 solubility allowing NMR structural analysis, although BrucD4-4 has a number of residues typical of conventional VH fragments. A plot of the surface corresponding to the VL interface in the VH domain from Fv Pot, BrucD4-4, and the llama V$_H$H fragment HC-V (Figure 7) shows that the BrucD4-4 surface can be situated between the other two with respect to hydrophobicity. This observation may explain an enhanced solubility of BrucD4-4 sufficient for the solution structure determination.

The foldability and reasonable stability of anti-idiotypic llama VH fragments indicate that despite the absence of the VL and CH1 partners, they still contain a sufficient number of stabilizing interactions to avoid misfolding and/or aggregation. Interestingly, recent modeling and experimental studies demonstrated that a limited subset of structural contacts scattered over the protein sequence may have a greater effect on the conformational stability and foldability of proteins than others (*58*−*62*). Mirny and Shakhnovich proposed to look at universally conserved positions in protein folds as the evolutionary signals for protein stability and foldability (*61*). Some of these general principles on protein stability have already found practical applications in the engineering of antibody single-chain Fv fragments (*2*). Similar strategies have been suggested for the design of libraries of water-soluble VH fragments (*36*). Five amino acids at positions 6, 74, 83, 84, and 108 (residues Ala6, Ala74, Lys83, Ser84, and Gln108 for BrucD4-4) were identified as having conservative substitutions between the set of water-soluble llama VHs and human VH3 family of sequences and are proposed to represent foldability/stability determinants for water-soluble llama VH fragments (*36*). Indeed, a library of human VH fragments incorporating the corresponding amino acid substitutions has been constructed and shown to have significantly improved solution properties (unpublished results).

The set of dAb molecules reported here also represents a good system for studying the correlations between conformational motions, thermal stability, and solution behavior of members of engineered protein libraries. Single-domain antibodies 1B1E10, BrucB3, BrucC6, BrucD4-4, and BrucC7.2 are similar in their primary and, presumably, tertiary

structures and have sufficient solubility for rigorous NMR relaxation studies. They span a significant range of unfolding transition temperatures of 57−72 °C and display markedly different tendency toward aggregation upon denaturation. NMR relaxation studies, which are underway in our laboratory, will define the locations of the conformational motions in dAb molecules and may assist in engineering new stable VH-based protein libraries.

REFERENCES

1. Muyldermans, S., Cambillau, C., and Wyns, L. (2001) *Trends Biochem. Sci. 26*, 230−235.
2. Worn, A., and Pluckthun, A. (2001) *J. Mol. Biol. 305*, 989−1010.
3. Skerra, A. (2000) *J. Mol. Recognit. 13*, 167−187.
4. Winter, G., and Milstein, C. (1991) *Nature 349*, 293−299.
5. Skerra, A., and Pluckthun, A. (1988) *Science 240*, 1038−1041.
6. Glockshuber, R., Malia, M., Pfitzinger, I., and Pluckthun, A. (1990) *Biochemistry 29*, 1362−1367.
7. Ueda, H., Tsumoto, K., Kubota, K., Suzuki, E., Nagamune, T., Nishimura, H., Schueler, P. A., Winter, G., Kumagai, I., and Mohoney, W. C. (1996) *Nat. Biotechnol. 14*, 1714−1718.
8. Huston, J. S., Levinson, D., Mudgett-Hunter, M., Tai, M. S., Novotny, J., Margolies, M. N., Ridge, R. J., Bruccoleri, R. E., Haber, E., Crea, R., and. (1988) *Proc. Natl. Acad. Sci. U.S.A. 85*, 5879−5883.
9. Bird, R. E., Hardman, K. D., Jacobson, J. W., Johnson, S., Kaufman, B. M., Lee, S. M., Lee, T., Pope, S. H., Riordan, G. S., and Whitlow, M. (1988) *Science 242*, 423−426.
10. Huston, J. S., McCartney, J., Tai, M. S., Mottola-Hartshorn, C., Jin, D., Warren, F., Keck, P., and Oppermann, H. (1993) *Int. Rev. Immunol. 10*, 195−217.
11. Reiter, Y., and Pastan, I. (1998) *Trends Biotechnol. 16*, 513−520.
12. Hudson, P. J. (1999) *Curr. Opin. Immunol. 11*, 548−557.
13. Ward, E. S., Gussow, D., Griffiths, A. D., Jones, P. T., and Winter, G. (1989) *Nature 341*, 544−546.
14. Davies, J., and Riechmann, L. (1994) *FEBS Lett. 339*, 285−290.
15. Kortt, A. A., Guthrie, R. E., Hinds, M. G., Power, B. E., Ivancic, N., Caldwell, J. B., Gruen, L. C., Norton, R. S., and Hudson, P. J. (1995) *J. Protein Chem. 14*, 167−178.

16. Hamers-Casterman, C., Atarhouch, T., Muyldermans, S., Robinson, G., Hamers, C., Songa, E. B., Bendahman, N., and Hamers, R. (1993) *Nature 363*, 446−448.

17. Ghahroudi M. A., Desmyter, A., Wyns, L., Hamers, R., and Muyldermans, S. (1997) *FEBS Lett. 414*, 521−526.

18. van der Linden, R. H., Frenken, L. G., de Geus, B., Harmsen, 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.

19. Perez, J. M., Renisio, J. G., Prompers, J. J., van Platerink, C. J., Cambillau, C., Darbon, H., and Frenken, L. G. (2001) *Biochemistry 40*, 74−83.

20. Dumoulin, M., Conrath, K., Van Meirhaeghe, A., Meersman, F., Heremans, K., Frenken, L. G., Muyldermans, S., Wyns, L., and Matagne, A. (2002) *Protein Sci. 11*, 500−515.

21. Ewert, S., Cambillau, C., Conrath, K., and Pluckthun, A. (2002) *Biochemistry 41*, 3628−3636.

22. Lauwereys, M., Arbabi, G. M., Desmyter, A., Kinne, J., Holzer, W., De Genst, E., Wyns, L., and Muyldermans, S. (1998) *EMBO J. 17*, 3512−3520.

23. Riechmann, L., and Muyldermans, S. (1999) *J. Immunol. Methods 231*, 25−38.

24. Davies, J., and Riechmann, L. (1995) *Biotechnology 13*, 475−479.

25. Tanha, J., Xu, P., Chen, Z., Ni, F., Kaplan, H., Narang, S. A., and MacKenzie, C. R. (2001) *J. Biol. Chem. 276*, 24774−24780.

26. Spinelli, 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. Desmyter, A., Transue, T. R., Ghahroudi, M. A., Thi, M. H., Poortmans, F., Hamers, R., Muyldermans, S., and Wyns, L. (1996) *Nat. Struct. Biol. 3*, 803−811.

28. 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.

29. Voordijk, S., Hansson, T., Hilvert, D., and van Gunsteren, W. F. (2000) *J. Mol. Biol. 300*, 963−973.

30. Decanniere, K., Desmyter, A., Lauwereys, M., Ghahroudi, M. A., Muyldermans, S., and Wyns, L. (1999) *Structure. Fold. Des. 7*, 361−370.

31. Spinelli, S., Frenken, L. G., Hermans, P., Verrips, T., Brown, K., Tegoni, M., and Cambillau, C. (2000) *Biochemistry 39*, 1217−1222.

32. Spinelli, S., Tegoni, M., Frenken, L., van Vliet, C., and Cambillau, C. (2001) *J. Mol. Biol. 311*, 123−129.

33. Decanniere, K., Muyldermans, S., and Wyns, L. (2000) *J. Mol. Biol. 300*, 83−91.

34. Riechmann, L. (1996) *J. Mol. Biol. 259*, 957−969.

35. Davies, J., and Riechmann, L. (1996) *Protein Eng 9*, 531−537.

36. Tanha, J., Dubuc, G., Hirama, T., Narang, S., and MacKenzie, R. (2002) *J. Immunol. Methods 263*, 97−109.

37. Kabat, E., Wu, T. T., Perry, H. M., Gottesman, K. S., and Foeller, C. (1991) *Sequence of Proteins of Immunological Interest*, Publication 91-3242, US Public Health Service, NIH, Bethesda, MD..

38. Muyldermans, S., Atarhouch, T., Saldanha, J., Barbosa, J. A., and Hamers, R. (1994) *Protein Eng. 7*, 1129−1135.

39. Vu, K. B., Ghahroudi, M. A., Wyns, L., and Muyldermans, S. (1997) *Mol. Immunol. 34*, 1121−1131.

40. Nguyen, V. K., Muyldermans, S., and Hamers, R. (1998) *J. Mol. Biol. 275*, 413−418.

41. Nguyen, V. K., Hamers, R., Wyns, L., and Muyldermans, S. (2000) *EMBO J. 19*, 921−930.

42. Anand, N. N., Dubuc, G., Phipps, J., MacKenzie, C. R., Sadowska, J., Young, N. M., Bundle, D. R., and Narang, S. A. (1991) *Gene 100*, 39−44.

43. MacKenzie, C. R., Sharma, V., Brummell, D., Bilous, D., Dubuc, G., Sadowska, J., Young, N. M., Bundle, D. R., and Narang, S. A. (1994) *Biotechnology 12*, 390−395.

44. Sattler, M., Schleucher, J., and Griesinger, C. (1999) *Prog. Nucl. Magn. Res. Spectrosc. 34*, 93−158.

45. Delaglio, F., Grzesiek, S., Vuister, G. W., Zhu, G., Pfeifer, J., and Bax, A. (1995) *J. Biomol. NMR 6*, 277−293.

46. Johnson, B. A., and Blevins, R. A. (1994) *J. Biomol. NMR 4*, 603−614.

47. Farrow, N. A., Muhandiram, R., Singer, A. U., Pascal, S. M., Kay, C. M., Gish, G., Shoelson, S. E., Pawson, T., Forman-Kay, J. D., and Kay, L. E. (1994) *Biochemistry 33*, 5984−6003.

48. Linge, J. P., O'Donoghue, S. I., and Nilges, M. (2001) *Methods Enzymol. 339*, 71−90.

49. Brunger, A. T., Adams, P. D., Clore, G. M., DeLano, W. L., Gros, P., Grosse-Kunstleve, R. W., Jiang, J. S., Kuszewski, J., Nilges, M., Pannu, N. S., Read, R. J., Rice, L. M., Simonson, T., and Warren, G. L. (1998) *Acta Crystallogr. Crystallogr., Sect. D 54*, Part 5, 905−921.

50. Koradi, R., Billeter, M., and Wuthrich, K. (1996) *J. Mol. Graphics 14*, 51−32.

51. Jones, J. A., Wilkins, D. K., Smith, L. J., and Dobson, C. M. (1997) *J. Biomol. NMR 10*, 199−203.

52. Wilkins, D. K., Grimshaw, S. B., Receveur, V., Dobson, C. M., Jones, J. A., and Smith, L. J. (1999) *Biochemistry 38*, 16424−16431.

53. Fan, Z. C., Shan, L., Guddat, L. W., He, X. M., Gray, W. R., Raison, R. L., and Edmundson, A. B. (1992) *J. Mol. Biol. 228*, 188−207.

54. Grasberger, B. L., Gronenborn, A. M., and Clore, G. M. (1993) *J. Mol. Biol. 230*, 364−372.

55. Wishart, D. S., and Sykes, B. D. (1994) *J. Biomol. NMR 4*, 171−180.

56. Trinh, C. H., Hemmington, S. D., Verhoeyen, M. E., and Phillips, S. E. (1997) *Structure 5*, 937−948.

57. Padlan, E. A. (1994) *Mol. Immunol. 31*, 169−217.

58. Fersht, A. R. (1995) *Proc. Natl. Acad. Sci. U.S.A. 92*, 10869−10873.

59. Itzhaki, L. S., Otzen, D. E., and Fersht, A. R. (1995) *J. Mol. Biol. 254*, 260−288.

60. Shakhnovich, E., Abkevich, V., and Ptitsyn, O. (1996) *Nature 379*, 96−98.

61. Mirny, L. A., and Shakhnovich, E. I. (1999) *J. Mol. Biol. 291*, 177−196.

62. Tiana, G., Broglia, R. A., and Shakhnovich, E. I. (2000) *Proteins 39*, 244−251.

63. Chothia, C., Lesk, A. M., Gherardi, E., Tomlinson, I. M., Walter, G., Marks, J. D., Llewelyn, M. B., and Winter, G. (1992) *J. Mol. Biol. 227*, 799−817.

64. Engelman, D. M., and Steitz, T. A. (1981) *Cell 23*, 411−422.