Protein Engineering, Design & Selection vol. 25 no. 1 pp. 39–46, 2012
Published online December 5, 2011 doi:10.1093/protein/gzr057

Antibody engineering reveals the important role of J segments in the production efficiency of llama single-domain antibodies in *Saccharomyces cerevisiae*

A.Gorlani[1†], D.Lutje Hulsik[2†], H.Adams[3], G.Vriend[4], P.Hermans[3] and T.Verrips[1,5]

[1]Biomolecular Imaging Group, Department of Biology, Faculty of Science, Utrecht University, Padualaan 8, 3584 CH, Utrecht, The Netherlands, [2]Unit of Virus Host Cell Interactions (UVHCI), UMI 3265, Université Joseph Fourier-EMBL-CNRS, rue Jules Horowitz 6, 38042 Grenoble, France, [3]BAC B.V., J.H. Oortweg, 2333CH, Leiden, The Netherlands and [4]CMBI, Radboud University Nijmegen Medical Centre, Geert Grooteplein 26-28, 6525 GA Nijmegen, The Netherlands

[5]To whom correspondence should be addressed.
E-mail: t.verrips@qvquality.com

Received September 21, 2011; revised November 1, 2011; accepted November 8, 2011

Edited by Hans Thogersen

Variable domains of llama heavy-chain antibodies (VHH) are becoming a potent tool for a wide range of biotechnological and medical applications. Because of structural features typical of their single-domain nature, they are relatively easy to produce in lower eukaryotes, but it is not uncommon that some molecules have poor secretion efficiency. We therefore set out to study the production yield of VHH. We computationally identified five key residues that are crucial for folding and secretion, and we validated their importance with systematic site-directed mutations. The observation that all key residues were localised in the V segment, in proximity of the J segment of VHH, led us to study the importance of J segment in secretion efficiency. Intriguingly, we found that the use of specific J segments in VHH could strongly influence the production yield. Sequence analysis and expression experiments strongly suggested that interactions with chaperones, especially with the J segment, are a crucial aspect of the production yield of VHH.

Keywords: BiP chaperone/J segment/key residues/protein production/VHH stability

Introduction

In 1993, Hamers *et al.* discovered that members of the Camelidae family possess not only conventional antibodies but also antibodies devoid of the light chain (Hamers-Casterman *et al*., 1993). Variable domains of these heavy-chain antibodies (VHH) thus have only three complementarity determining regions (CDRs) but they can nevertheless recognise their cognate antigens with binding constants equalling those of conventional immunoglobulins. Their high binding affinity and small size make them good candidates for applications ranging from biotechnology to medicine and their relatively large CDR3 may be an advantage in the recognition of epitopes in clefts or pockets unreachable by conventional antibodies (De Genst *et al*., 2006).

The importance of VHH as therapeutic agents is rapidly growing as they can be very specific and normally show little side effects (Borrebaeck and Carlsson, 2001; Wesolowski *et al*., 2009) in a wide variety of applications. The possibility to produce VHH as homo- or hetero-multimers (Conrath *et al*., 2001; Stewart *et al*., 2007), often with synergistic effects, further adds to their usefulness as therapeutic (Kruger *et al*., 2006; Saerens *et al*., 2008), anti-bacterial (Kruger *et al*., 2006) or anti-viral agents (Jahnichen *et al*., 2010; Hultberg *et al*., 2011).

We are developing VHH as microbicides to combat infectious diseases, in particular HIV (Forsman *et al*., 2008; Gorlani *et al*., 2011), in developing countries. To be successful, these microbicides should be affordable and physically stable under a wide range of conditions. We have developed continuous and fed-batch fermentation processes for VHH using *Saccharomyces cerevisiae* as production host and often have reached high production yields (van de Laar *et al*., 2007). However, neither are all VHH equally well produced nor are they all equally stable. Despite almost 20 years of study into this topic, little is known to date about the production yield and stability (together called 'production efficiency') of VHH. Production efficiency is a key factor that determines whether a promising molecule will become a marketable product.

It has been shown that VHH binding affinity is mainly determined by the CDRs while folding and secretion efficiency are mostly a function of the β-sheet scaffold (Nicaise *et al*., 2004). These characteristics allowed us to independently study and engineer the binding properties and the production efficiency (Saerens *et al*., 2005). Obviously, the production efficiency of heterologous proteins is the sum of many factors. Previously, we have studied several aspects of the production efficiency of VHH, including copy number of the heterologous gene, mRNA stability, presence of rare codons for *S.cerevisiae*, glycosylation, or charge and hydrophobicity of the protein surface (Lutje Hulsik, 2009). We here continue this systematic study with the identification of amino acids crucial for folding as folding efficiency has been shown important for production efficiency (Parekh *et al*., 1995).

Vendruscolo *et al*. (2001) studied folding using experimentally and computationally obtained Φ values (ratios of the change in stability of the transition state and native state associated with each amino acid) and found a limited

[†]The authors contributed equally to this study.

number of residues with high Φ values. These residues made many contacts in the native structure *and* in the transition state of folding. They called them 'key residues' and concluded that the presence of their interactions in the transition state is crucial for proper folding. We have used this concept to computationally identify key residues in VHH, and we have experimentally validated their importance by means of site-directed mutagenesis.

The analysis of the location of the key residues in the 3D VHH structure led to the realisation that they are all in the vicinity of the J segment. As a consequence we analysed the combinatories of variability (V), diversity (D) and joining (J) genes during the antibody maturation and found that the choice of J segment is important for VHH production yield in *S.cerevisiae*. As the endoplasmic reticulum (ER)-resident chaperone BiP/Kar2p binds specific motifs of seven amino acids consisting of large hydrophobic residues alternated to small ones (Blond-Elguindi *et al*., 1993) and such motifs are found in β-strand 9 of VH and VHH domains (Knarr *et al*., 1995), we studied the relation between folding chaperones and production efficiency of VHH.

When expressed in yeast under the lead of the invertase signal peptide, VHH are co-translationally translocated to the ER and their folding occurs as the nascent polypeptide chain enters the ER (Ng *et al*., 1996). Protein folding is usually a faster reaction than translation/translocation (Fischer and Schmid, 1990; Huard *et al*., 2006), unless interactions with chaperones are required to reach the native state, as is the case for immunoglobulins (Feige *et al*., 2010). If folding does not proceed optimally the protein will remain associated with BiP (Brodsky *et al*., 1999). Extended association with BiP or other mechanisms (Werner *et al*., 1996; Meerovitch *et al*., 1998) can label the protein as a substrate for ER-associated degradation.

The ER is a complex system in which many chaperones, cofactors and synthesised proteins interact with each other (Braakman and Bulleid, 2011). Despite similarities, the ER of lower eukaryotes like *S.cerevisiae* is less evolved than that of mammalian organisms where llama antibodies are naturally produced, in particular the difference between mammalian chaperone BiP and its yeast homologous Kar2p may be a decisive factor (Rose *et al*., 1989). Production yield differences observed in yeast as function of the J segment were not observed in human embryonic kidney cells, supporting the hypothesis that interactions between J segment and chaperones might be an important contributor to VHH production yield.

Materials and methods

Strains, media and culturing conditions

Saccharomyces cerevisiae VWK18 gal1 (CEN-PK 102-3A, MATa, leu2-2, ura3, gal1::URA3, MAL-8, MAL3, SUC2) was used for all production efficiency tests. Preparation of all VHH constructs was carried out using *Escherichia coli* strain TG1 supE hsd-5 thi$^-$ (lac-proAB) F'[traD36 proAB + lacIq lacZ-M15]. The expression tests in mammalian cells were performed in human embryonic kidney cells (HEK293E) by U-Protein Express BV, Utrecht, The Netherlands.

Selective medium for *S.cerevisiae* was yeast nitrogen base without amino acids, 6.7 g/l, glucose 20 g/l. Expression medium was YP broth (yeast extract, 10 g/l, bacto peptone, 20 g/l), glucose 20 g/l, galactose 5 g/l. LB medium, supplemented with ampicillin (100 μg/ml), was used for the selection of transformant *E.coli* colonies.

Saccharomyces cerevisiae cultures were grown overnight at 30°C, 180 r.p.m., in selective medium, then diluted 1:100, to OD_{600nm} 0.03–0.08 in expression medium and incubated at 30°C, 190 r.p.m. for 48 h. *Escherichia coli* cultures were grown in selective LB medium (Bacto Tryptone, 10 g/l, yeast extract, 5 g/l, NaCl 10 g/l) overnight at 37°C, 200 r.p.m.

Mutants construction

VHH R2 was used for the key residue experiments. All point mutations were introduced by QuickChange Site-directed Mutagenesis (Stratagene) following the manufacturer's instructions. Primers of 25–27 nucleotides in total, with the mutated codon roughly in the middle of the sequence and flanked by 8–14 wild-type nucleotides at each side were designed. Replacement of β-strand 9 was done by polymerase chain reaction (PCR). We used a forward primer that annealed to a region within the signal peptide at the 5′ of the VHH gene, containing the restriction site PstI, and custom-designed reverse primers encoding the DNA sequence specific for J4- and J7- β-strand 9 and the BstEII restriction site for compatibility with vector pUR4585. PCR products of the expected length were excised from agarose gel, digested and ligated in pUR4585.

After transformation of *E.coli* and confirmation by DNA sequencing, the purified plasmid DNA was transformed into *S.cerevisiae* using the lithium acetate method (Gietz and Schiestl, 2007). Yeast cells were plated on selective medium and at least five colonies that contained the plasmid, as confirmed by growth in selective conditions and further DNA sequencing, were screened in shake flask cultures for VHH secretion and to determine the average secretion yield.

Protein analysis

Culture supernatant (10 μl) was mixed with sample buffer [80 mM Tris/HCl pH 6.8, 33% glycerol, 6.7% sodium dodecyl sulfate (SDS), 300 mM dithiothreitol (DTT), 0.01% bromophenol blue], boiled for 10 min and loaded on 15% poly-acrylamide gels. After electrophoretic separation, proteins were blotted onto immobilon membrane (Millipore, MA, USA). Detection of the VHH protein band was performed with protein G-purified anti-VHH rabbit serum 1:5000 in milk/1% phosphate-buffered saline (PBS) for 1 h. Secondary antibody goat anti-rabbit (Li-Cor Biosciences, NE, USA) conjugated with an infrared dye IR800 (1:10 000 in milk/1% PBS for 1 h) was used for detection of the rabbit serum. Alternatively, VHH were detected with mouse anti-His monoclonal antibody (Clontech, CA, USA) 1:5000 in milk/1% PBS for 1 h. Secondary antibody donkey anti-mouse (Li-Cor) conjugated with an infrared dye IR800 (1:10 000 in milk/1% PBS for 1 h) was used for visualisation.

Protein purification

Saccharomyces cerevisiae cultures were centrifuged in a Sorvall RC-5BPlus at 6000 r.p.m. for 15 min to pellet cells. Fermentation medium was subsequently passed through 0.22 μm PES membrane filter for clarification after which we used the VivaFlow 200 (Sartorius AG, Germany) ultrafiltration device to reduce the volume and to exchange buffers.

Samples were affinity-purified using an Akta Xpress machine (GE Healthcare Europe, Diegem, Belgium) equipped with HisTrap FastFlow 1 ml column (GE Healthcare). Binding buffer was 20 mM Tris-HCl pH 7.4, 300 mM NaCl, 20 mM imidazole. Elution buffer was 20 mM Tris-HCl pH 7.4, 500 mM NaCl, 500 mM imidazole. Immediately after elution, VHH were desalted to PBS using 2×5 ml HiTrap desalting columns, (GE Healthcare). Purified proteins were stored at -20°C.

Circular dichroism spectra and melting temperature (T_m) measurement

Jasco J-810 spectropolarimeter (Jasco Inc., Great Dunmow, UK) was used to obtain circular dichroism (CD) measurements and melting temperature (T_m) values of VHH. Purified VHH were dialysed against MilliQ water prior loading the machine. After identifying 205 nm wavelength as optimal for following state changes in VHH, a temperature gradient 20–90°C (unfolding) and back to 20°C (refolding) was applied to the samples, with a slope 2°C/min. Temperature-gradient experiments were repeated three times with fresh protein preparation from the same purification batch.

Bioinformatic tools and analysis

WHAT IF (Vriend, 1990) was used for sequence and structure analyses and homology modelling. YASARA (www.yasara.org) was used for visual inspection of 3D structures. Llama VHH structures were collected from the Protein Data Bank (Berman *et al*., 2000).

The following structures were used: 1G9E (Renisio *et al*., 2002), 1HCV (Spinelli *et al*., 1996), 1I3U (Spinelli *et al*., 2001), 1I3V (Spinelli *et al*., 2001), 1QD0 (Spinelli *et al*., 2000), 1SJX (Dolk *et al*., 2005), 1U0Q (Cambillau, unpubl.), 2BSE (Spinelli *et al*., 2006), 2XA3 (Hinz *et al*., 2010) and 3EZJ (Korotkov *et al*., 2009). A total of 1152 unique llama VHH sequences, obtained from proprietary and public databases, were used in this study (http://swift.cmbi.ru.nl/mcsis/systems/ABVDDB/).

Determination of important residues for *in vitro* folding was done with WHAT IF's implementation of the criteria of Vendruscolo *et al*. For each residue the Cα distance to each other Cα was measured, excluding the direct neighbouring residues. Only residues that made five or more contacts in all structures were taken into account. The standard WHAT IF options for mutant prediction and analyses were used to analyse the structure of mutant VHH and to analyse their gain or loss of contacts.

To determine the most likely germline genes of V and J segments, an alignment using WHAT IF's implementation of all-against-all combined DNA-protein alignments were done. The DNA sequences were translated in the correct reading frame, resulting in the following sequence format: cagQgtgVcagQ, i.e. first the codon, then the corresponding amino acid. Each sequence was stored using a special format, called a DNA−protein sequence, which dictates that the triplet and the corresponding amino acid remain associated and in-phase in subsequent alignment procedures.

Results

Determination of key residues for folding and secretion in vivo

Vendruscolo *et al*. 2001 identified and described, in the enzyme acylphosphatase, the interactions of a small number of amino acids that form a native-like contact network in the transition state, which was found to be essential to efficiently obtain the native fold of the protein. They called these residues 'key residues' and observed that they have more interactions with other amino acids, compared with non-key residues.

We analysed 10 3D structures of VHH available from the Protein Data Bank and we picked out between 26 and 39 amino acids in each structure that make an above-average number of contacts, like key residues, and thus are likely to form the points of nucleation for the folding of the protein. An alignment of these 10 VHH sequences revealed that 8 of the putative key residues were conserved: L20, W36, R38, V63, M82, L82c, D86 and Y90 (Kabat numbering. IMGT numbering: L21, W41, R43, V71, M91, L94, L98, Y102) (Fig. 1). These same residue positions are also highly conserved in the non-redundant database of 1152 VHH sequences from immunised and non-immunised llama libraries that we utilised in this study. Moreover, the interactions take place with amino acids belonging to the VHH framework, and not to the CDRs.

Using VHH R2 (Spinelli *et al*., 2000) as a model protein, we performed an alanine scan on the hydrophobic residues and made conservative mutations on the three hydrophilic ones. Additionally, we introduced mutations that were suggested by the alignment of the 1152 VHH sequences. We constructed L20A, W36A, V63A, M82A, L82cA, D86A, Y90A and L20V, L20I, R38K, D86E, Y90F. The conservative mutations were meant to maintain properties of the wild-type amino acid, as opposed to the introduction of alanine. Table I shows that the introduction of alanine at positions 20, 36, 38, 86 and 90 abolished secretion, whereas introduction of alanine at positions 63, 82 and 82c reduced the secretion to 15–70% of wild-type VHH R2. Conservative mutations had in some cases mild effects on secretion, like L20I (40% of wild type), R38K (30%) and Y90F (25%), while in other cases strong effects were observed, like L20V (0%) and D86E (5%). On the base of the effect on secretion we concluded that only five of the eight putative key residues are crucial for secretion. Moreover, we observed a lower T_m compared with VHH R2 wild type in all mutants for which enough material could be obtained. Even though a trend could be observed, that linked the reduction in stability and the lower secretion yield, an actual correlation between the two variables did not exist (P value >0.05).

Analysis of VHH segments and their association with secretion in S.cerevisiae

Figure 4A and B show the location in the VHH structure of the eight residues tested. It struck us that seven out of eight putative key residues are located in the vicinity of β-strand 9, the strand encoded by the J segment. The five crucial and two non-crucial key residues form a hydrophobic groove where the J segment docks upon folding. This prompted us to study the relation between V and J segments and the production yield.

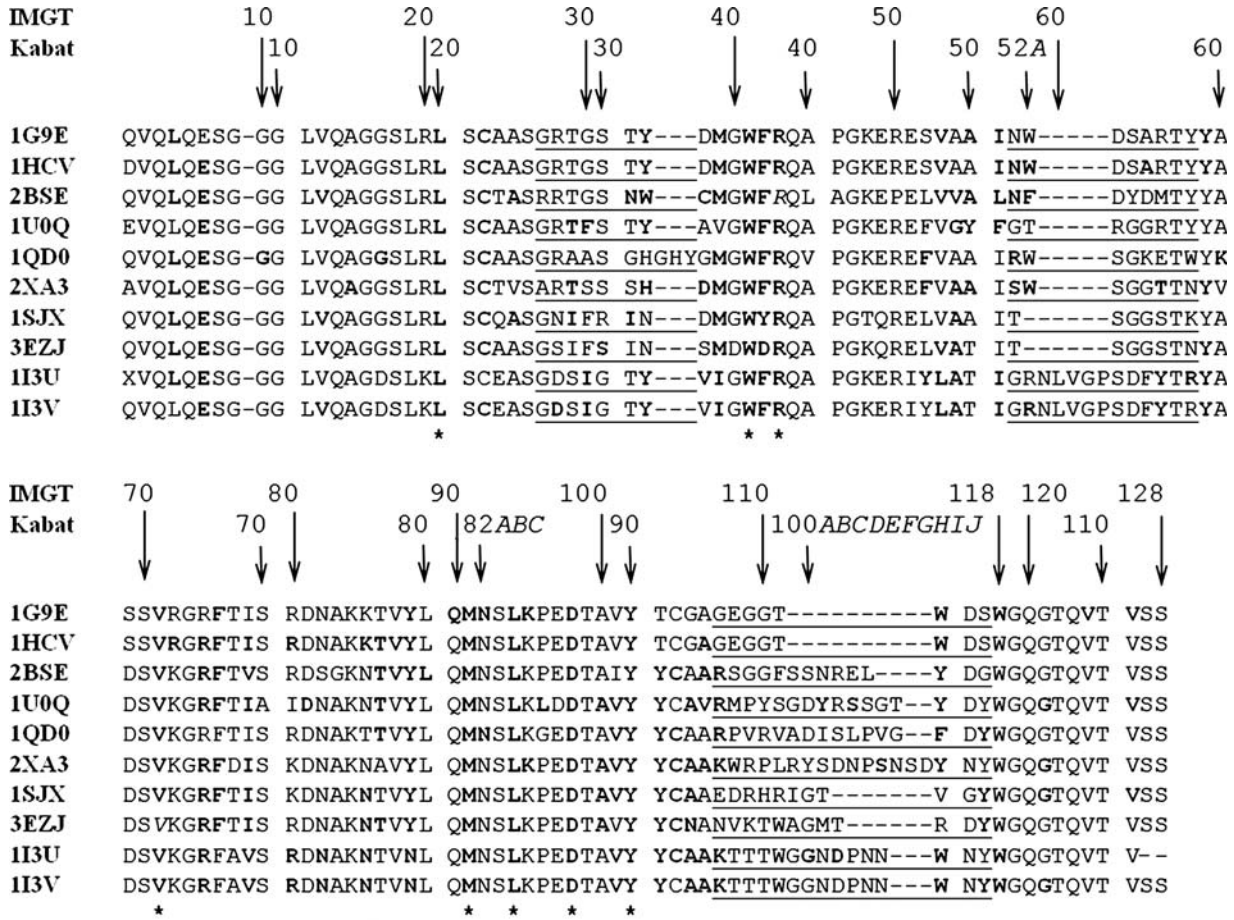

Fig. 1. Sequence alignment of the 10 VHH structures used in this study (both IMGT and Kabat numbering are indicated. Above the sequences, the apostrophe mark indicates the IMGT numbers; the dot indicates the Kabat numbers). CDR residues are underlined. Residues that make at least five contacts are in bold. Residues that make at least five contacts and are conserved in the alignment are indicated with a star. R38 2BSE and V63 3EZJ make four contacts but they were included in the selection for potential key residues.

Table I. Effect of key residue mutations on secretion and on thermal stability

Amino acid[a] (Kabat)	Number of interactions	Mutation introduced	Residual secretion (WT = 100%)[b]	ΔT_m (T_m mutant—T_m wild type)[b]
L20	7	L20A	0%	n.d.[c]
		L20V	0%	n.d.
		L20I	40% (SD 5%[d])	−14.2
W36	8	W36A	0%	n.d.
R38	5	R38K	30% (SD 4%[d])	−42.8[e]
D86	5	D86A	0%	n.d.
		D86E	5% (SD 1%[d])	−7.3
Y90	5	Y90A	0%	n.d.
		Y90F	25% (SD 3%[d])	−1.3
V63	6	V63A	15% (SD 3%[d])	−5.5
M82	5	M82A	20% (SD 3%[d])	−5.9
L82c	5	L82cA	70% (SD 8%[d])	−6.6

[a]The indicated amino acids were replaced by alanine or by amino acids similar to the wild type (conservative mutation).
[b]Secretion in *S.cerevisiae* and melting temperature (T_m) were measured and compared with the wild-type VHH.
[c]Not determined because not enough material could be obtained.
[d]SD = standard deviation.
[e]Aggregates of the mutant VHH R38K started appearing at 30°C, thus the measurement was stopped.

We expressed in *S.cerevisiae* 71 VHH that were part of the 1152-sequence database. These 71 VHH were selected from different libraries, recognise different antigens and the crucial key residues were present in all of them (except for VHH 7, 8 and 57, that carried the mild L20I substitution that was showed non-crucial for production yield). We measured the amount of VHH present in the culture medium and defined five secretion classes, from low to high yield (Fig. 2). Despite that it is commonly believed that VHH secrete well, we observed a 100-fold production yield difference from class 1 to 5 and almost half of the VHH measured had a secretion level <10 mg/l.

In the antibody maturation process, different V, D and J genes are assembled to generate the whole VH domain; therefore, we investigated whether certain of V, D and J combinations correlate with the secretion level of VHH in *S.cerevisiae*. We matched the 71 sequences with the 23 germline genes that encode for their V segment and linked the genes to the secretion class of the VHH. The same was done with the seven germline genes encoding for the J segment. The D segment consists of essentially CDR3 and it is mainly involved in antigen recognition therefore it was not considered. In the large sequence-database 23 V and 6 J

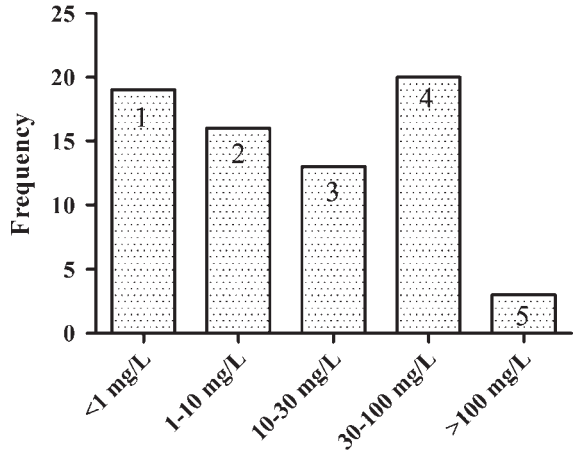

Fig. 2. Ranking of the production yield of 71 VHH. Five classes were defined, from <1 to >100 mg/l of VHH secreted in culture medium. Almost half of the VHH measured had a secretion level <10 mg/l.

Table II. Relation between V and J germline genes usage and secretion class of VHH

Germline[a]	Secretion class[b]		
	1/2	3	4/5
Vc–Ve	2	0	0
Vg–Vj	7	0	2
Vk–Vn	4	2	8
Vo–Vq	3	2	2
Vr–Vt	12	5	11
Vv–Vw	5	3	0
J4	18	8	17
J4 or J6	8	4	6
J7	9	1	0

[a]Four V and four J genes were not found in our dataset.
[b]No significant correlation exists between V genes and secretion yield. J7 gene correlates with low secretion yield in *S.cerevisiae* ($P = 0.01$).

Table III. Characterisation of VHH—J segment chimeric constructs

Protein	J segment used	Secretion class	T_m (°C)	Refolding ability
VHH3 WT	J4	5	74.5	+
VHH3 mut	J7	1	70.1	−
VHH70 WT	J7	1	70.8	−
VHH70 mut	J4	5	76.7	+
VHH A12 WT	J4	5	74.2	+
VHH A12 mut	J7	1	n.d.	n.d.
VHH 3F3 WT	J7	1	n.d.	n.d.
VHH 3F3 mut	J4	4	73.9	+

genes were found, whereas 18 V and 3 J genes were represented in the 71 VHH that we expressed.

No relation was found between V germlines and secretion class, whereas it appeared that J genes 4 and 6 were the only genes represented in the two highest secretion classes, while J7 was almost exclusively found in the two lowest secretion classes (Table II).

J segment (β-strand 9) swap between class 1 and class 5 VHH

To study the role of the J segments further, we analysed VHH3 and VHH70. These two VHH are in the highest and lowest secretion class, respectively, and they bind the same antigen. These VHH use the same V gene but use a J4 and J7 gene, respectively. Investigation of their modelled 3D structure showed that the inter β-strand interactions and the conformations of both β-strands 9 are very similar. We exchanged the J segments between the two VHH and measured the production yield. Mutant VHH70 (VHH70-J4) had a secretion yield comparable to VHH3, while mutant VHH3 (VHH3-J7) was almost not secreted (Table III). In this experiment the J gene 7 seemed the bottleneck for high secretion.

VHH3 and VHH70 shared the same V gene. We therefore investigated whether the effect of the J gene would be independent from the V gene used. VHH A12 (Forsman *et al.*, 2008) and 3F3 (Strokappe *et al.*, accepted) fitted the requirements because they originated from different V genes, and their J genes were J4 and J7, respectively. A12 was naturally well-secreted, whereas 3F3 was not. Swapping the J segments had on these molecules the same effect observed for VHH3 and VHH70 (Table III), thus strengthening the suggestion that the J segment is an important factor in the production efficiency, independently from the V segment used.

Thermal stability of VHH3, VHH70 and J-segment mutants.

Figure 3 shows the CD spectra and thermal denaturation of VHH3, VHH70 and their J-segment mutants. The spectra at 25°C show the native state of each protein; it appeared that replacing the J segment had an effect on the optical activity of both VHH3 and VHH70, with significant variations occurring at different wavelengths: at 215 and 235 nm in VHH3, and at 200 and 215 nm in VHH70. Measurement of the spectra at 90°C showed that all four proteins were unfolded. Thermal denaturation occurred for VHH3 at a higher temperature, 74.5°C (SD 1.1), than VHH70, 70.8°C (SD 0.8). However, introduction of segment J4 in VHH70 increased its thermal stability by almost 6 degrees, to 76.7°C (SD 1.4), and consistently, introduction of segment J7 in VHH3 had the opposite effect, lowering the unfolding temperature to 70.1°C (SD 1.1). The unfolding of VHH3 wild type and VHH70-J4 was reversible, whereas VHH3-J7 and VHH70 wild type would not refold when the temperature decreased. It seems that the presence of J4 relates to both a higher T_m and to reversibility of folding. These data indicate an important role played by the J segment.

Expression of VHH3 and VHH70 in HEK293 cells

Although protein stability and production yield often correlate in *S.cerevisiae* and other organisms (Kwon *et al.*, 1996; Kowalski *et al.*, 1998), we observed that this relation does not hold very well when we mutate key residues that are involved in the proper association of β-strand 9 with the rest of the VHH framework. Since several differences exist between the ER of yeast and mammalian cells, we investigated if aspects of the cell folding machinery could explain the observed effect. We tested whether distinct J segments had the same impact on the production of VHH3 and VHH70 in HEK293 cells, as they did in yeast cells. Expression experiments showed that the secretion difference between the two proteins was negligible. This supported our hypothesis that in *S.cerevisiae* the low production yield of VHH using J7-derived β-strand 9 is caused by suboptimal interactions between Kar2p and the folding VHH. In fact the

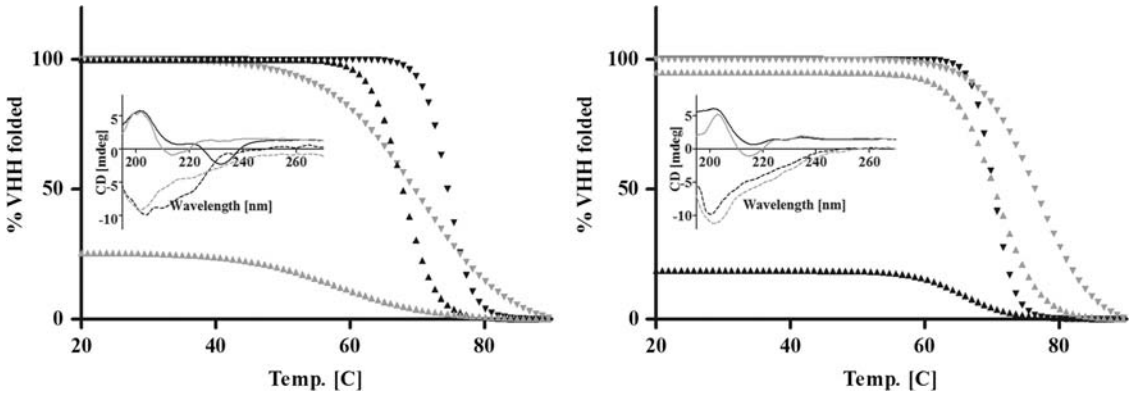

Fig. 3. Thermal denaturation and refolding of VHH. Per cent of folded VHH vs. temperature was studied by measuring their circular dichroism. Left: VHH3 wild type is represented by black triangles, VHH3 mutant by grey triangles. Heating cycles are shown with triangles pointing down; cooling cycles are shown with triangles pointing up. VHH3 mutant shows lower T_m than wild type and inability to refold. The spectra depicted in the small image show that both VHH are folded at 25°C (VHH3 wild-type black solid line, VHH3 mutant grey solid line) and are unfolded at 90°C (VHH3 wild-type black dashed line, VHH3 mutant grey dashed line). Right: The same symbols and colours apply to VHH70 wild type and mutant. VHH70 wild type has lower T_m than mutant and is unable to refold. The spectra in the smaller image show that both VHH are folded at 25°C (VHH70 wild-type black solid line, VHH70 mutant grey solid line) and are unfolded at 90°C (VH70 wild-type black dashed line, VHH70 mutant grey dashed line).

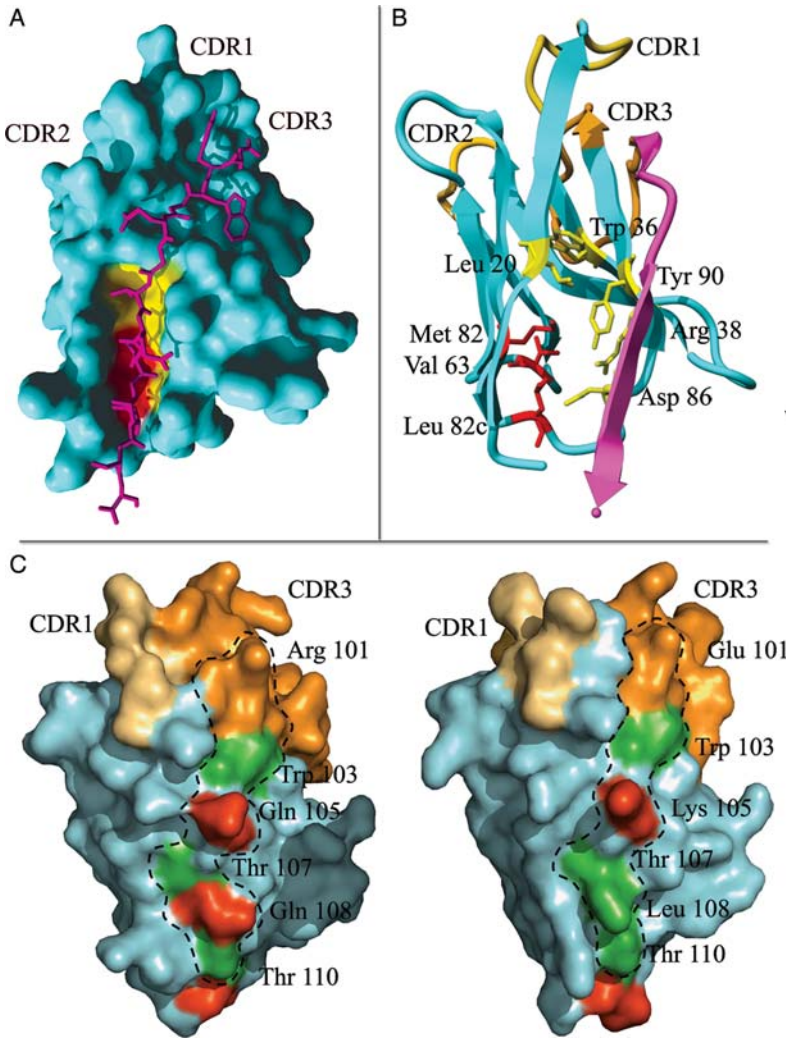

Fig. 4. (**A**) The surface of VHH R2 with β-strand 9 left out is shown in cyan. After surface calculation, β-strand 9 was put back as a purple stick model. The surface made up by the five key residues that essentially abolish production yield if mutated to alanine is shown in bright yellow, whereas the surface of residues M82 and L82c is in red. V63 does not contribute to this surface. (**B**) 3D ribbon representation of VHH R2. The scaffold is shown in cyan and the three CDR loops are in shades of orange. Crucial key residues L20, W36, R38, D86 and Y90 are shown in yellow, residues V63, M82, L82c are in red. β-Strand 9 is represented in purple. (**C**) Molecular surface of VHH3 (left) and VHH70 (right). The scaffold is shown in cyan and the three CDR loops are in shades of orange. β-Strand 9 residues are shown in green (hydrophobic side-chain) or red (hydrophilic side-chain). BiP-recognised motifs according to Blond-Elguindi *et al*. are circled with a dashed line. Note that β-strand 9 derived from J4 (left) and J7 (right) display different hydrophobic patterns.

3D models of VHH3 and VHH70 revealed that the side chains of the solvent-accessible residues in β-strand 9 form different hydrophobic motifs, as highlighted in the dashed area recognised by BiP/Kar2p (Fig. 4C). We recorded CD spectra of VHH produced in mammalian cells and we compared them with spectra of yeast-produced VHH that we previously acquired. We observed that both the unfolding and refolding transitions occurred in the same manner and at the same temperature, confirming that the native conformation of VHH is not dependent on the host used for their production.

Conclusions

The growing importance of llama heavy-chain antibody fragments, VHH, in health care and as research and imaging tools has raised the question how to produce them efficiently. We use *S.cerevisiae* as production host because of its outstanding record in biotechnology and because it is a 'generally recognised as safe' microorganism (Buckholz and Gleeson 1991; Gellissen *et al*., 1992; Idiris *et al*., 2010).

We previously analysed several host- and process-related factors (Lutje Hulsik, 2009) and here we looked at key residues and specific V or J genes as determinants of secretion. We found that amino acids L20, W36, R38, D86 and Y90 behave like Vendruscolo's key residues and we determined that they are very important for the proper folding of VHH *in vivo*. However, key residues are not more important than other residues for the stability of the structure, as demonstrated by the comparable loss in T_m of key residue mutants and the other destabilising mutants V63, M82 and L82c.

The observation that all five crucial key residues are located close to the side of the β-sandwich where β-strand 9 is meant to fold brought our attention to its role in folding. The large difference in secretion efficiency recorded among the 71 VHH tested, despite all of them having the five key residues in place, suggested an important role in folding and production yield for the β-strand 9 that is encoded by the J gene. The correlation of J gene 7 with inefficient secretion indicated that independently of the V segment, amino acids on the J segment can still have a strong effect on the folding. Once we identified J segments specific for high and for low production yield we were able to clone these J segments in other VHH (based on unrelated V and D segments) and transfer the beneficial or detrimental effects.

The involvement of chaperones in determining the folding efficiency of the J segment was confirmed by the parallel expression of a J4- and J7-related VHH in yeast and mammalian systems. Whereas *S.cerevisiae* has a rather primitive chaperone system, mammalian cells have excellent chaperones, and in fact the large secretion difference reported in yeast cells completely disappeared in mammalian cells. The equivalent thermal stability of yeast-produced and mammalian-produced VHH is the proof that the folding pathway, rather than the final conformation, causes the production yield differences.

The above observations led us to speculate on a possible two-step folding mechanism for VHH. As chaperones seem to play a role in keeping the C-terminal domains unfolded (J segment and possibly CDR3), this might suggest that the N-terminal V segment folds first, provided that the key residues are in place, and forms the cavity where the J segment subsequently docks.

Our work points out a new aspect of the production efficiency of VHH, which is the importance of folding kinetics, rather than the stability of the native structure. We anticipate that interesting results might be obtained by cloning mammalian chaperones such as BiP into the yeast used for VHH expression.

Acknowledgements

A.G. would like to acknowledge Jacques Doux and Daniele Porcelli for assistance with the CD experiments. G.V. thanks NBIC for support.

Funding

This work was partially funded by the European Union program Europrise.

References

Berman,H.M., Westbrook,J., Feng,Z., Gilliland,G., Bhat,T.N., Weissig,H., Shindyalov,I.N. and Bourne,P.E. (2000) *Nucleic Acids Res.*, **28**, 235–242.
Blond-Elguindi,S., Cwirla,S.E., Dower,W.J., Lipshutz,R.J., Sprang,S.R., Sambrook,J.F. and Gething,M.J. (1993) *Cell*, **75**, 717–728.
Borrebaeck,C.A. and Carlsson,R. (2001) *Curr. Opin. Pharmacol.*, **1**, 404–408.
Braakman,I. and Bulleid,N.J. (2011) *Annu. Rev. Biochem.*, **80**, 71–99.
Brodsky,J.L., Werner,E.D., Dubas,M.E., Goeckeler,J.L., Kruse,K.B. and McCracken,A.A. (1999) *J.Biol.Chem.*, **274**, 3453–3460.
Buckholz,R.G. and Gleeson,M.A. (1991) *Biotechnology (NY)*, **9**, 1067–1072.
Conrath,K.E., Lauwereys,M., Wyns,L. and Muyldermans,S. (2001) *J.Biol.Chem.*, **276**, 7346–7350.
De Genst,E., Silence,K., Decanniere,K., Conrath,K., Loris,R., Kinne,J., Muyldermans,S. and Wyns,L. (2006) *Proc. Natl Acad. Sci. USA*, **103**, 4586–4591.
Dolk,E., van der Vaart,M., Lutje Hulsik,D., Vriend,G., de Haard,H., Spinelli,S., Cambillau,C., Frenken,L. and Verrips,T. (2005) *Appl. Environ. Microbiol.*, **71**, 442–450.
Feige,M.J., Hendershot,L.M. and Buchner,J. (2010) *Trends Biochem. Sci.*, **35**, 189–198.
Fischer,G. and Schmid,F.X. (1990) *Biochemistry*, **29**, 2205–2212.
Forsman,A., Beirnaert,E., Aasa-Chapman,M.M., *et al*. (2008) *J. Virol.*, **82**, 12069–12081.
Gellissen,G., Melber,K., Janowicz,Z.A., Dahlems,U.M., Weydemann,U., Piontek,M., Strasser,A.W. and Hollenberg,C.P. (1992) *Antonie Van Leeuwenhoek*, **62**, 79–93.
Gietz,R.D. and Schiestl,R.H. (2007) *Nat. Protoc.*, **2**, 31–34.
Gorlani,A., Brouwers,J., Mc Conville,C., *et al*. (2011) *AIDS Res. Hum. Retroviruses*, doi: 10.1089/aid.2011.0133.
Hamers-Casterman,C., Atarhouch,T., Muyldermans,S., Robinson,G., Hamers,C., Songa,E.B., Bendahman,N. and Hamers,R. (1993) *Nature*, **363**, 446–448.
Hinz,A., Lutje Hulsik,D., Forsman,A., *et al*. (2010) *PLoS One*, **5**, e10482.
Huard,F.P., Deane,C.M. and Wood,G.R. (2006) *Bioinformatics*, **22**, e203–e210.
Hultberg,A., Temperton,A.J., Rosseels,V., *et al*. (2011) *PLoS ONE*, **6**, e17665.
Idiris,A., Tohda,H., Kumagai,H. and Takegawa,K. (2010) *Appl. Microbiol. Biotechnol.*, **86**, 403–417.
Jahnichen,S., Blanchetot,C., Maussang,D., *et al*. (2010) *Proc. Natl Acad. Sci.USA*, **107**, 20565–20570.
Knarr,G., Gething,M.J., Modrow,S. and Buchner,J. (1995) *J. Biol. Chem.*, **270**, 27589–27594.
Korotkov,K.V., Pardon,E., Steyaert,J. and Hol,W.G. (2009) *Structure*, **17**, 255–265.
Kowalski,J.M., Parekh,R.N. and Wittrup,K.D. (1998) *Biochemistry*, **37**, 1264–1273.
Kruger,C., Hultberg,A., Marcotte,H., Hermans,P., Bezemer,S., Frenken,L.G. and Hammarstrom,L. (2006) *Appl. Microbiol. Biotechnol.*, **72**, 732–737.
Kwon,W.S., Da Silva,N.A. and Kellis,J.T., Jr. (1996) *Protein Eng.*, **9**, 1197–1202.

Lutje Hulsik,D. (2009) Improving properties of llama heavy-chain antibodies using in silico analysis. *Ph.D. Thesis* Utrecht University, Faculty of Science, Department of Biology. pp. 80–103.
Meerovitch,K., Wing,S. and Goltzman,D. (1998) *J. Biol. Chem.*, **273**, 21025–21030.
Ng,D.T., Brown,J.D. and Walter,P. (1996) *J. Cell Biol.*, **134**, 269–278.
Nicaise,M., Valerio-Lepiniec,M., Minard,P. and Desmadril,M. (2004) *Protein Sci.*, **13**, 1882–1891.
Parekh,R., Forrester,K. and Wittrup,D. (1995) *Protein Expr. Purif.*, **6**, 537–545.
Renisio,J.G., Perez,J., Czisch,M., Guenneugues,M., Bornet,O., Frenken,L., Cambillau,C. and Darbon,H. (2002) *Proteins*, **47**, 546–555.
Rose,M.D., Misra,L.M. and Vogel,J.P. (1989) *Cell*, **57**, 1211–1221.
Saerens,D., Ghassabeh,G.H. and Muyldermans,S. (2008) *Curr. Opin. Pharmacol.*, **8**, 600–608.
Saerens,D., Pellis,M., Loris,R., Pardon,E., Dumoulin,M., Matagne,A., Wyns,L., Muyldermans,S. and Conrath,K. (2005) *J. Mol. Biol.*, **352**, 597–607.
Spinelli,S., Desmyter,A., Verrips,C.T., de Haard,H.J., Moineau,S. and Cambillau,C. (2006) *Nat. Struct. Mol. Biol.*, **13**, 85–89.
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.
Spinelli,S., Frenken,L.G., Hermans,P., Verrips,T., Brown,K., Tegoni,M. and Cambillau,C. (2000) *Biochemistry*, **39**, 1217–22.
Spinelli,S., Tegoni,M., Frenken,L., van Vliet,C. and Cambillau,C. (2001) *J. Mol. Biol.*, **311**, 123–129.
Stewart,C.S., MacKenzie,C.R. and Hall,J.C. (2007) *Toxicon*, **49**, 699–709.
van de Laar,T., Visser,C., Holster,M., Lopez,C.G., Kreuning,D., Sierkstra,L., Lindner,N. and Verrips,T. (2007) *Biotechnol. Bioeng.*, **96**, 483–494.
Vendruscolo,M., Paci,E., Dobson,C.M. and Karplus,M. (2001) *Nature*, **409**, 641–645.
Vriend,G. (1990) *J. Mol. Graph.*, **8**, 52–56, 29.
Werner,E.D., Brodsky,J.L. and McCracken,A.A. (1996) *Proc. Natl Acad. Sci. USA*, **93**, 13797–13801.
Wesolowski,J., Alzogaray,V., Reyelt,J., *et al*. (2009) *Med. Microbiol. Immunol.*, **198**, 157–174.