ISAS - INTERNATIONAL SCHOOL FOR ADVANCED STUDIES

Isas-International School for Advanced Studies

VH dimers (VHDs): a novel binding molecule

Thesis submitted for the Degree of

Doctor Philosophiae

Candidate
Jorge Sepulveda

Supervisor
Dr. Oscar Burrone

Academic Year 2000-2001

SISSA - SCUOLA INTERNAZIONALE SUPERIORE DI STUDI AVANZATI
TRIESTE
Via Beirut 2-4

TRIESTE

I have learnt that knowledge is related to modesty, honesty and the wish of helping and sharing with everybody.
I have learnt that everybody has something valuable inside. It is just a matter of understanding.
I have learnt that understanding is the most ambiguous word I have ever heard.
I hope I will learn how to understand.

Acknowledgements

During these five years I had the chance to meet a lot of people. Some left more than others did, but everybody has given me something. I hope I have replied your kindness. It is difficult to thank some friends without being unfair with others, especially considering that a thesis does not only involve science. If I begin to count the huge number of people who have helped me and have contributed to finish this period of my life I would have to add many pages to this manuscript or maybe I would have to duplicate the number of pages. …Thanks to all of you.

I thank my teacher and friend Oscar for allowing me learning from him and with him.

Special acknowledgements to my family who are always present wherever I am. To my nephews for being so nice. And especially to my parents who have made me what I am, for good or bad. Maybe, just a little less stubborn...

CONTENTS

Summary

The present thesis describes the development of a novel-binding molecule based on VH dimers. The objective was to profit from the good qualifications of that domain to bind antigens. The very large diversity of this domain in terms of sequence variations, as well as the higher number of residue contact interactions that generally forms with the antigen makes VH domains an ideal candidate to improve binding properties.

The combination of two VHs coupled in a similar way to the VL/VH would create new, non-natural antigen binding surfaces with yet undefined properties. These binders could have higher affinities than classical VL/VH, considering the number of amino acids contacts. In addition, the use of only VHs allows the possibility to form homodimers, which would provide symmetric binding surfaces that could recognisie antigen molecules with high symmetry, such as dsDNA.

The aim of the present thesis was to study whether VH domains could dimerise. We found useful to know if they could be expressed by mammalian cells, bacteria and be displayed on phages. In addition, we wanted to find out if they could bind specifically an antigen.

In the present thesis we demonstrate that VHs are efficiently expressed in mammalian cells as dimers showing a specific binding activity. These dimers are oriented similarly to VL/VH couples and are relatively stable even if they are not covalently linked. They can also be expressed in bacteria and displayed on phages keeping their binding specificities, thus, making them a biotechnological tool with high potentiality. Using the phage display technology we selected specific hetero- and homo-dimeric binders showing that also symmetrical surfaces are possible to be formed. In addition, we present data that strongly suggest the involvement of both domains in the antigen recognition. Other aspects related to the requirements of scVHD (single chain VH dimer) formation and purification such linkers and expression system were also investigated.

CHAPTER 1

Introduction

The human immune system is capable of specifically recognising and responding to an enormous number of antigens. It has evolved different mechanisms to locate foreign cells, viruses or macromolecules, to neutralise and eliminate them. Immunoglobulins play an important role in this event as soluble protein or as B-cell membrane receptor.

Immunoglobulin Structure

The basic structure of all immunoglobulin (Ig) molecules is a unit consisting of two identical light (L) polypeptide chains and two identical heavy (H) polypeptide chains linked together by non covalent interactions and stabilised by disulphide bonds. Both H and L chains are organised into domains that are defined by homology either in constant (C) or variable (V) regions on the basis of the degree of sequence variation amongst different antibody molecules. Both V domains (Fv fragment) form the antigen binding site (figure 1).

The plant protease papain cleaves the Ig molecule in the hinge region between the CH1 and CH2 domains, to give two identical Fab fragments and one Fc (crystallizable) fragment. These papain-generated fragments have been of enormous value in structure/function studies on the antibody molecule because they separated the Fab

region, which binds to the antigen, from the Fc region, which mediates effector functions. Heavy chain C domains (C region) define the classes and subclasses of mammalian Igs.

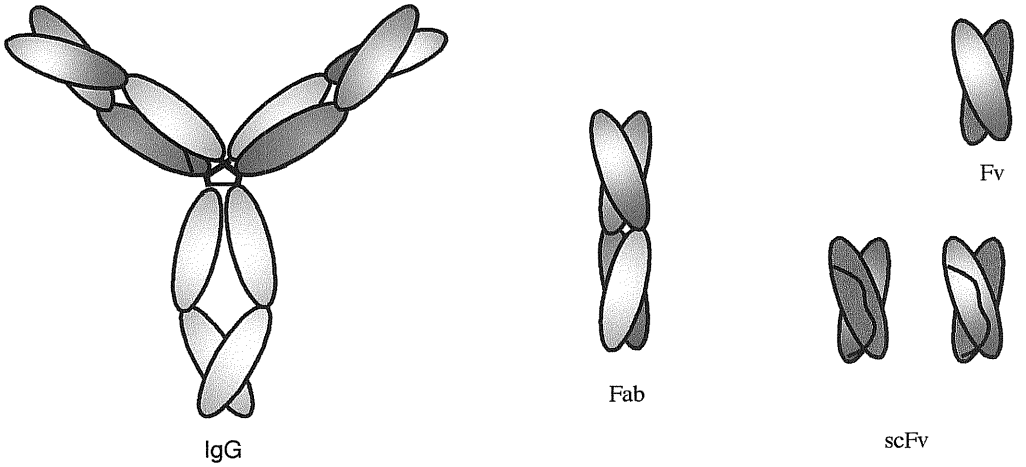


Figure 1 Structure of a complete Ig molecule and of the Fab, Fv and single chain Fv (scFv) fragments. H chain= VH-CH1-CH2-CH3. L chain= VL-CL.

Five types of H-chain exist, designated by the greek letters μ, δ, γ, ε, α, which define the corresponding immunoglobulin classes IgM, IgD, IgG, IgE and IgA. They differ in size, charge, amino acid composition and carbohydrate content. A primary difference is the presence of an extra C domain in IgM and IgE. The C region types (isotypes) mediate immunological effector functions, such as complement fixation, placental transfer, and binding to cell surface Fc receptors, that are specific to particular isotypes.

On the other hand, light chains are represented only by two isotypes, κ and λ. In humans κ light chains comprise approximately 60% of the human light chain protein, while in

mice represent approx. 95%. Functional differences have not been identified for the two isotypes of mammalian L chains and neither an H-isotype association bias.

The hinge region between the CH1 and CH2 domains, vary in length and may facilitate antigen binding by increasing H chain flexibility. It is also important in coupling the binding and effector function of the immunoglobulin, allowing formation of the complete tetrameric complex by disulphide bridges.

In addition, the Ig molecule can be expressed on the cell surface or secreted as a soluble protein with different oligomerisation patterns. On the membrane it shows a monomeric organisation although it has recently been described oligomerisation of IgM (non-covalent) using mild detergent cell lysis (1). When secreted IgG and IgE are monomers whereas IgA and IgM when expressed in conjunction with the invariant J chain protein form covalently linked dimers or pentamers, respectively. However, when no J chain protein is present, IgM is secreted as hexamer (2).

The Immunoglobulin Domain

The immunoglobulin fold has been widely identified in many organisms, including chicken, zebrafish and S. cerevisiae. It comprises a very big family of proteins known as Immunoglobulin Superfamily referring to the diverse group of genes, which contain one

or more Ig homology domains. In mammalians, apart from immunoglobulins, several molecules adopt this folding including cell surface molecules and other immune proteins.

Each domain is approximately 110 aminoacids in length with a fold (immunoglobulin fold) consisting of two stacked layers of antiparallel β sheets surrounding an internal space filled with hydrophobic aminoacid side chains with terminal exposed loops arranged in a sandwich, or β barrel structure. Each fold contains seven polypeptide β strands, four of which comprise one β pleated sheet, the other sheet consisting of the remaining three strands. The two layers are covalently linked by a disulphide bond found in most, but not all, proteins containing the immunoglobulin fold (figure 2) (3,4).

The Variable Domain

Each V domain consists of four regions of relatively conserved sequence called framework regions (FR1, FR2, FR3 and FR4). The basic secondary structure coincides with β-strands connected by loops as an immunoglobulin domain, but from which three are highly hypervariable sequences. The framework regions form the scaffold for these hypervariable loops placing them on one end of the β-barrel creating the antigen binding site, determined by the shape of the surface and so its specificity. These hypervariable regions are also referred to as complementarity determining regions (CDRs) and are located in the H-chain between 26-32 (H1), 53-55 (H2) and 96-101 (H3), and in the L-

chain between 26-33 (L1), 50-52 (L2) and 91-96 (L3), according to the Kabat numbering system (5). The length of the variable domains of H (VH) and L (VL) chains differ from the C domains by 16 aminoacids that allow to form a distinctive fold using two additional β strands and an extra loop connecting these strands (figure 2).

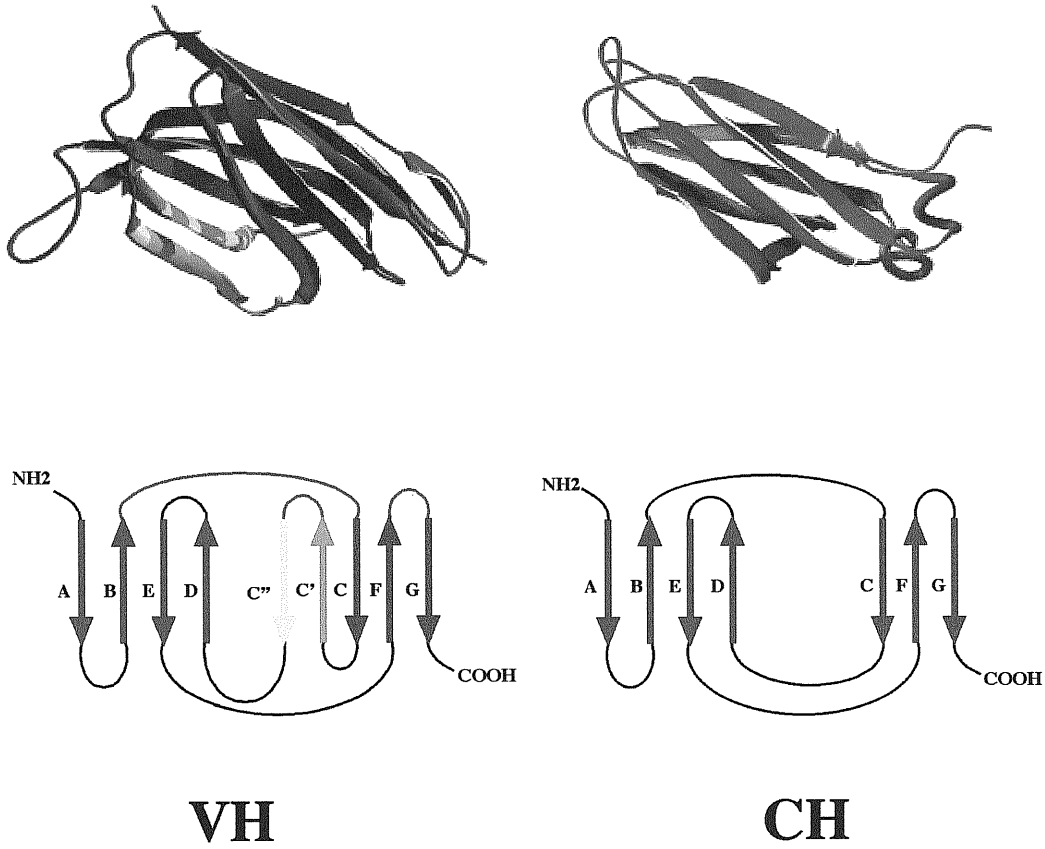


Figure 2: Immunoglobulin domain tertiary stucture. Domains VH and CH are shown in the β barrel conformation. The two additional β strands of VH are coloured with yellow and light green and the CDR loops with red.

Domain dimerisation

Immunoglobulin domains interact in such a way that an antibody actually consists of a series of dimeric modules. CH3 domain in IgG and IgA, and CH4 domain in IgM and IgE

(three-strands/four-strands sandwiches) normally associate to each other through the four-stranded face by hydrophobic surfaces in a very tight manner. In the cases of IgG, IgA and IgD, the Fc is a dimer of two domains CH2-CH3, while in IgM and IgE it consists of paired CH2-CH3-CH4 domains. In contrast, CH1 in all isotypes does not face its partner but the CL.

Similarly, V domains (four-strands/ five-strand sandwiches) form dimers using the five-strand face. The VL-VH interface consists of two closely packed β-sheets and its geometry corresponds to a nine-stranded elliptical or prism-shaped barrel, which place the six CDRs close to each other at the amino terminal end of the Fv fragment. The barrel forms the bottom and sides of the antigen-binding site. However, aminoacid residues that are part of the domain-domain interface and that appear not to be accessible to the solvent or the antigen can also contribute to antibody specificity (6). Moreover, mutations of conserved interface residues may affect antigen binding indirectly by influencing the relative position of the VL and VH domains (7). Many of the residues responsible for VH/VL interdomain contacts have been localised. Overall 12 to 21 VL and 16 to 22 VH residues participate in interchain stabilisation (3). About half of the hydrophobic core contacts are formed between FR2 of one chain (VL residues: P^{44}-hydrophilic^{45} - hydrophobic^{46} - hydrophobic^{47}, VH residues: G^{44} - L^{45} - E^{46} - W^{47} - hydrophobic^{48}) and FR4 of the other (VH/VL: W^{103}/F^{98}-$G^{104/99}$-$X^{105/100}$-$G^{106/101}$) (figure 3). Furthermore, these

regions are highly conserved in humans and murines and when modified, like in camelids, the VH becomes more soluble and does not pair with VL (8). The consequence of this is that the VHs are soluble avoiding aggregations. This natural selection has been exploited with biotechnological interest by specifically mutating residues in the VH interface ($G^{44}E$, $L^{45}R$, $W^{47}G$) allowing the monomeric expression in soluble form of defined VHs in bacteria (8, 115).

The conserved VH-VL interface suggests a very permissive VH-VL association, although, some unknown aspects provoke biased VH-VL combinations (9).

Moreover, it is accepted that the VH-VL interface is an adaptable structure, modulated by both the CDRs of the antibody and its interaction with antigen (10).

Experiments performed with oligopeptides and oligonucleotides have shown that upon antigen binding the relative orientation of the 2 V regions can take place in the complexes up to 16 degrees shifts in VH-VL association (11).

An unusual mode of domain-domain association is the VL-VL interaction found in the Bence-Jones disease which closely mimics the interface geometry of VL-VH, although, its contact area is lower by 18% (6).

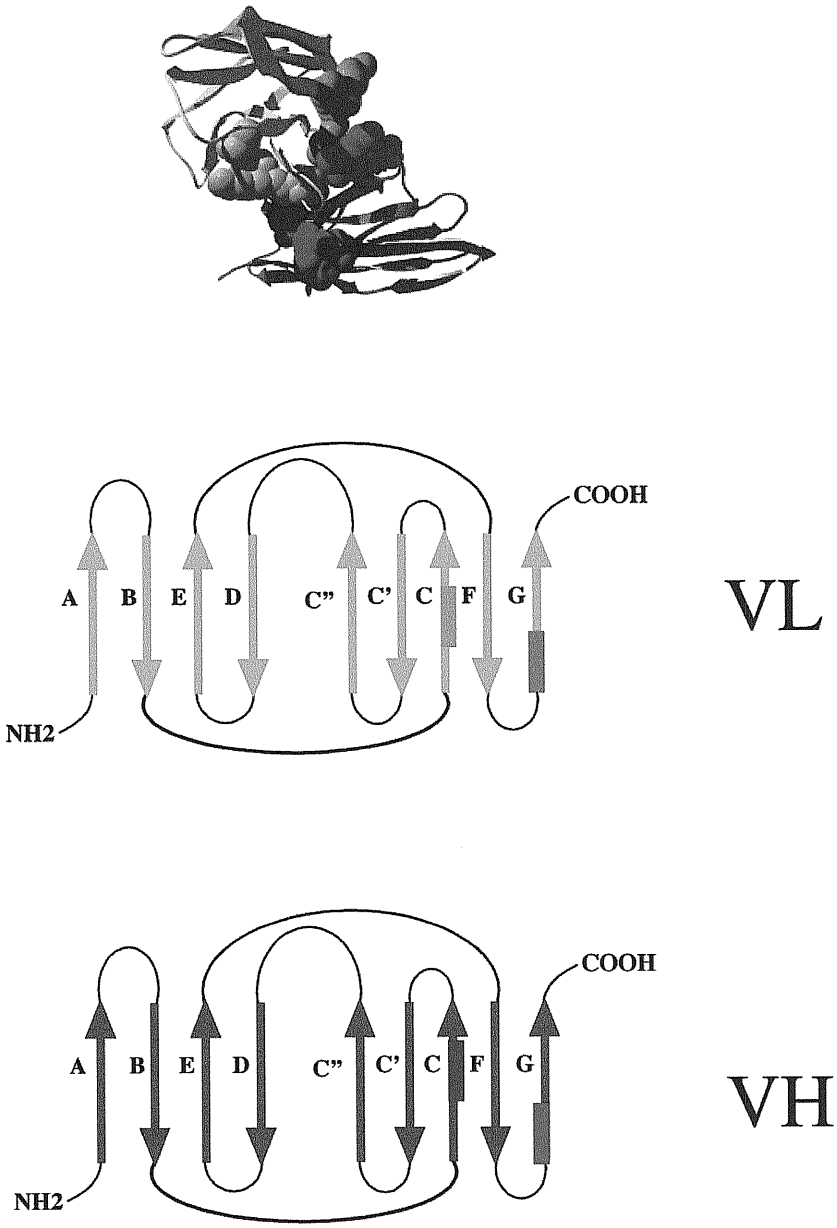


Figure 3: Schematically representation of hydrophobic b-bulges inducing VL-VH dimerisation. Light colours correspond to L-chain and dark ones to H-chain. Red spheres represent regions in FR4 and blue ones in FR2 that are located in the main β-bulges in dimerisation.

In contrast to these trans VH/VL associations that involved an area of around 1000-1700 $Å^2$ (12, 13), cis VH-CH1 and VL-CL contacts also are present. The existence of residues at the V/C domain interface, forming an extended hydrophobic patch, is preserved in all antibodies, their exact position and extend varies. A comparison of known Fab structures showed that between L-chain variable and constant domain an area of 410±90 $Å^2$ per domain is buried, while the heavy chain V/C interaction covers an area of 710±180 $Å^2$. The aminoacid conservation in these areas is not particularly high and some residues but not all are hydrophobic, predominantly aliphatic (14).

Immunoglobulin genes

The Ig molecule is encoded by three independent gene loci, namely Ig k and Ig l genes for the L chain and IgH genes for the H chain, which are located, in humans, in chromosome 2, 22 and 14, respectively. The expression of a complete immunoglobulin molecule is a consequence of the DNA recombination event that takes place independently in each lymphocyte precursor and leads to the assembly of specific DNA segments in both the H-chain and the L-chain loci.These rearrangements involve the formation of the active V regions which are encoded by a newly formed exon assembled upstream of the C region. For the H chain V regions the recombination

process involves the joining of three different gene segments: a variable (VH), a diversity (D) and a joining (JH) segment. Each VH gene segment contains two exons: the first encodes a hydrophobic leader sequence that facilitates vectorial translocation of the new synthesised molecule to the endoplasmic reticulum, while the second encodes the FR1, CDR1, FR2, CDR2 and FR3 regions. CDR3 is encoded by a D segment; short sequences derived from the VH-D and D-JH junctions, and a part of the JH (up to the W^{103}). The remaining JH segment, including the very conserved motif W^{103}-G^{104}-X^{105}-G^{106}, encodes the fourth framework region.

In humans, the IgH contains approximately 125 segments (15), classified into seven different families, of which 51 are functional (16). Approximately 25 D gene segments have been identified and just 6 functional JH segments (17, 18).

The heavy chain constant region genes are organised in clusters each encoding a different C region. In humans, the μ constant region (Cμ) is located closest to the JH segments followed by the exon clusters encoding the other H chain isotypes (19) (figure 4).

B-lymphocytes express μ H chain first during differentiation, while the other isotypes are expressed later after the process of class switching has occurred. An exception is IgD (20, 21), which is co-expressed with IgM by alternative RNA processing and termination of transcription.

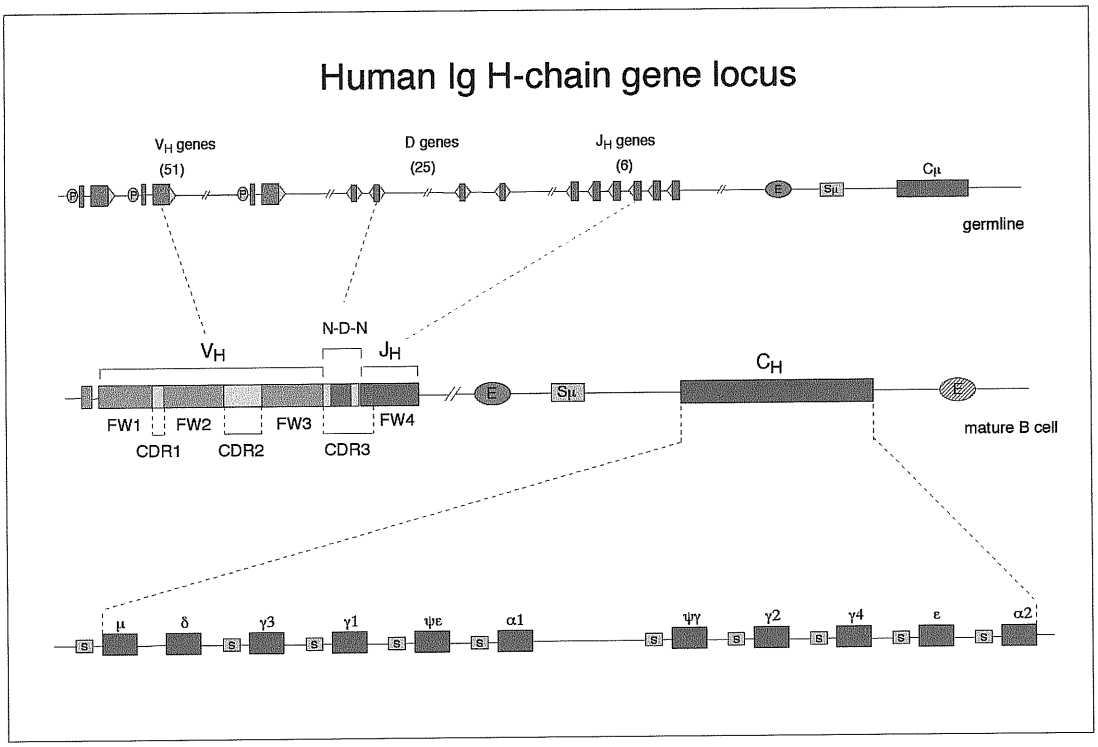


Figure 4 Human heavy chain locus organisation and rearrangement. E represents transcription enhancers and S boxes sequences involved in class-switch recombination. C regions are schematically shown as a single box.

The variable region of each L chain isotype is encoded by an upstream exon that is assembled from analogous VL and JL segments which are joined directly to each other; there are no L chain D segments. The genetic segments for the assembly of functional VL proteins derive from either of two different loci: κ (40 functional) or λ (31 functional). In humans, κ light chains comprise approximately 60% of the human light chain protein. Similar to the VH genes, each of these Vκ segments has two exons, one encoding the leader sequence and the second encoding most of the V region. The remaining of the Vκ region is encoded by one of the five functional Jκ gene segments. Downstream of the Jκ segment is located the single Cκ gene. For the λ light chain, after the Vλ regions there are seven Jλ-Cλ clusters, each composed with

one Jλ and one Cλ gene (4 functional pairs), placed in the same transcriptional orientation as the Vλ sequences (vbase directory: www.mrc-cpe.cam.ac.uk/imt-doc).

Ig genes expression

The VH-DH-JH locus undergoes V(D)J recombination during early B cell development. V(D)J recombination initiates at the pro-B cell stage and is ordered with DH to JH rearrangement preceding VH to DJH rearrangement. Additional diversity is created by the junctional imprecision of the joining events and by the deletion of nucleotides and addition in these areas of new, untemplated nucleotides between D and JH and between VH and D, called N regions. This further increases the diversity of distinct IgH chains that can be generated from the relatively modest amount of genetic information present in the germline. Generation of a μIgH chain from a productive V(D)J segment results in differentiation to the precursor-B cell stage in which most Ig light chain variable region genes are assembled, eventually producing the complete surface Ig complexes.

The V(D)J recombination involves recognition and double strand cleavage by recombinase-activating genes (RAGs) 1 and 2, which together form a recombinase (RAG 1/2), at the borders between the recombination signal sequence (RSS) and the coding DNA. The conserved heptamer and nonamer sequences of the RSSs are separeted by non-conserved spacers of 12 or 23 base pairs (forming 12-RSSs and 23-RSSs). The 12/23 rule, which is mediated at the level of RAG1/2 recognition and

cutting, specifies that the recombination process occurs only between a gene segment flanked by 12-RSS and one flanked by a 23-RSS, ensuring that only the appropriate gene segments are joined (22). The RAG 1/2 recombinase with the help of the non-specific DNA binding protein HMG1 (23) forms a synaptic complex between RSSs with different spacer lengths. The subsequent cleavage by the RAG 1/2 can be dived into two distinct steps (24). First, the RAG proteins nick the DNA at the 5' end of a signal heptamer adjacent to the coding DNA. The nick is then converted into a hairpin in a single trans-esterification step, resulting in a double strand DNA break with a blunt 5'-phosphorilated signal and a hairpin end. The joining phase is RAG 1/2 dependent, as well (25), and involves the Ku protein, the catalytic subunit of DNA-dependent protein kinase, DNA ligase IV and Xrcc4 (26). Variable number of nucleotides may be added to the 3' ends through the terminal deoxynucleotide transferase (TdT) during this process (27).

The first recombination event is a D to JH rearrangement, which invariably occurs on both H chain alleles. Precursor B cells then undergoes VH to DJH rearrangement on one or both alleles, generating potentially functional VDJ H-chain genes. When a productive rearrangement of anyone allele occurs, further changes in the second allele are inhibited (allelic exclusion) and the cell proceed to IgL rearrangement. Although it is established that this phenomenon is mediated by a feedback mechanism whereby the first productive rearrangement brings about the repression of RAG proteins which consequently prevents further rearrangement events (28). More recent evidences

indicate that additional levels of regulation exist. During lymphoid cell development, immune system loci undergo demethylation and open their chromatin structure in anticipation of the rearrangement reaction. It has been shown that demetyhlation occurs preferentially in one allele and that this modification marks the allele for rearrangement before the initiation of the recombination reaction (29). However, it is not clear yet how the allele is selected and which factors regulates the demethylation process.

Control of H chain gene transcription involves the VH promoter and multiple transcriptional enhancer elements. The Eμ is located within the JH-Cμ intron (30, 31) and a series of enhancers (collectively referred as 3'αE) lie downstream of Cα (32). This enhancer elements are active specifically in lymphoid cells, and function synergistically with the VH promoter, which is also tissue specific However, Eμ is necessary for IgH expression in precursor-B cell and dispensable for terminally differentiated B cell lines, where 3' Eα strongly activates transcription.

The assembly of L-chain genes follows H-chain gene rearrangement. This DNA recombination event joins one of the many germline Vκ regions with a particular joining (Jκ) region. If a Vκ and a Jκ gene segment are effectively joined in a cell already possessing an effective H-chain VDJ recombination, a μ-κ surface immunoglobulin B cell results. If the rearrangement of Vκ is not successful then rearrangement of the Vλ locus is initiated resulting in the expression of a μ-λ bearing B cell.

L chain gene expression appears to be regulated by mechanisms analogous to those that control H chain expression. Vκ and Vλ promoter are structurally similar to VH promoters, and are most active in lymphoid cells. Activity of the promoter of a rearranged Vκ segment depends on a tissue specific enhancer located within the Jκ-Cκ intron (33). A stronger enhancer 3'κE is located downstream to the Cκ and as the intronic enhancer can be stimulated in pre B-cells by LPS. In contrast, the human λ locus has only one enhancer located downstream Cλ (34).

Ig Assembly

Ig chain consists of a series of domains that fold independently of each other into a compact structure composed of two twisted β sheets stabilised by a single disulphide bond. In pre B cells that do not express L chains, H chains are efficiently retained in the endoplasmic reticulum (ER) due to their association with the ER chaperone, BiP (35). Initiation of L chain expression allows the formation of H2L2 molecules. BiP binds transiently to the nascent L chain in vivo when the variable domain is in an unfolded state and releases it as the domain folds. Although, BiP can bind transiently to multiple H chain domains the stable BiP binding site on unassembled heavy chains is the CH1 domain, which remains unfolded and unoxidised in the absence of light chain. The other heavy chain domains can be secreted, as partially assembled molecules, in the absence of CH1. (36). However, the role of L chain is not as a scaffold on which the CH1 can fold, but instead serves to release BiP from the heavy

chain, since this domain is perfectly capable of folding if BiP is released in vitro (37). The way that light chain interacts with CH1 is either when BiP has cycled off, thereby preventing it to rebind, or triggering BiP release. It is also plausible that VH and VL domains pair and form a semi-stable interaction that allows the CL domain to contact the CH1 domain and in some way displace BiP. Finally, CH1 is able to fold and assemble stably with the light chain, yielding a transport-competent and functional Ig molecule. The complete understanding of this process still remains unclear. Another chaperone, GRP94, has been described to act after BiP since it exclusively binds fully oxidised molecules (38). Nevertheless, no direct involvement in Ig assembly support its role in this process. GRP94 is acquiring more relevance in the field of cancer immunotherapy due to its ability to form complexes with peptides and trigger cellular immune response to intracellular antigens (39).

In addition, it is generally believed that complete immunoglobulin molecules, H2L2, are generated by the assembly of two HL intermediates. However, at least part of the IgE molecules assemble along a H+L->HL+H->H2L+L->H2L2 pathway supported by the presence of H2L intermediates in the cellular extracts of both ε-s1 and ε-s2, two alternatively spliced isoforms of IgE, transfectomas. Moreover, the ε-s2, which contain a C-terminal cysteine, was found to be covalently linked, suggesting a high rate of oxidation of the Cys^{554}, which allows formation of a disulphide bridge between the two H chains, even before the association with the second L chain. On the other hand, no covalently linked ε-s1, isoform without the C-terminal cysteine, H2L were

detected, indicating that the oxidation of the two cysteines from CH2 domain occurs more slowly and may require association into complete H2L2 molecules (40).

Membrane versus secreted Immunoglobulin

The differentiation of the mature B cell to plasma cell is accompanied by the loss of surface immunoglobulin expression as a consequence of conversion from the synthesis of membrane to secretory immunoglobulin molecules. The secreted form of a given immunoglobulin molecule has the same VDJ segments of the membrane form, but differ at the carboxy terminal amino acids. The two forms are encoded by mRNA species that arise by alternative splicing of transcripts derived from a single heavy chain gene. It has been proposed that differential transcription, termination and polyadenilation of the primary RNA transcript determines which of the two mRNA is expressed. Each immunoglobulin heavy chain gene contains at least two polyadenilation sites. Usage of the promoter proximal polyadenilation site, located immediately 3' to the stop codon of the last CH domain results in a mRNA encoding a secretory form of heavy chain. If however, transcriptional termination and polyadenilation occur at a site located several kilobases to the 3' of the first polyadenilation site; a mRNA containing sequence from two additional exons (M1 and M2) is produced. In most cases exon M1 encodes the EMPD and TMD while M2 codes for the cytoplasmic C-terminal domain. In the case of human ε H-chain gene, alternative splicing in the 3' part o the ε locus involving exons M1 and M2, generates

a number o different ε-mRNAs coding for two functional secretory (named ε s1 and ε s2)(41,40) and two functional membrane isoforms (named ε-short and ε-long) (42, 43, 44, 45). RNA processing events remove the first polyadenilation site from the membrane mRNA by splicing from a conserved donor site in the terminal CH exon (either the CH3 or CH4 exon depending on the heavy chain isotype) to the M1 exon. Thus, in the case of immunoglobulin heavy chain gene expression, transcription and termination at alternative end sites determines the expression of either the membrane or secreted protein. Furthermore, no function for EMPD has been reported, although, recent studies evidence a stabilisation role of membrane bound immunoglobulins due to cysteine residues that are involved in inter-chain bridges (46).

B cell development

Progression of B cell along the B-cell differentiation pathway can be monitored by the expression of developmentally regulated genes. Even though several surface markers exist, also the rearrangement statuses of the immunoglobulin genes are useful developmental markers (figure 5).

RAG1 and RAG2 are co-ordinately expressed during B cell development, as sequential IgH D to JH and VH to DJH occur. The D to JH rearrangement is initiated in pro-B cells that also express the Igα and Igβ component of the antigen receptor, as well as the surrogate light chains (ψLC). Completion of a successful VH to DJH

rearrangement marks the transition to pre B cells when it pairs with ψLC, allowing the surface expression of the pre-B cell receptor (μVHψLC, Igα and Igβ).

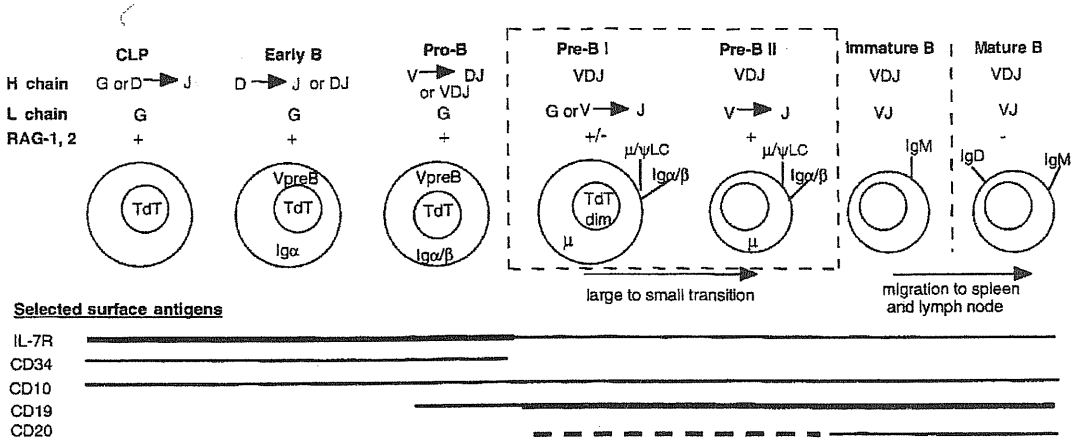


Figure 5 B-cell development and selected surface markers. TdT= terminal deoxytransferase. Igα/β correspond to CD79a and b accessory molecules. G: locus germline configuration.

Selection is achieved by the pre B cell proliferation with complete pre-BCR over the non-proliferating cells expressing non-fitting μH. The pre-BCR also signals for allelic exclusion at the IgH locus by preventing rearrangements in the second IgH allele. Subsequent rearrangement of a kappa or lambda chain V gene permits cell surface expression of a conventional membrane IgM receptor on the immature B cell which exists in the bone marrow and migrates to the periphery. Mature B cells in spleens and secondary lymphoid organs are competent to respond to cognate antigen and signalling from T cells. They circulate through the blood and lymph nodes until they contact specific antigens, typically in the T cell area in the spleen or lymph nodes.

Some B-cells are activated in situ, proliferate to form foci, and then differentiate to the antibody-secreting cell-state. Other activated B-cells migrate to the follicular region, where they initiate the formation of germinal centres.

The maturation of the immune response

In T-cell dependent immune responses (TD) (responses to protein antigens), the affinity of a given antibody for its cognate antigen increases from the relatively low affinity of the primary response to the high affinity antibodies predominant in a memory response. This process, referred to as affinity maturation, occurs in specialised microenvironments called germinal centres (GC) (47) (figure 2). Germinal centres arise from a limited number of B cells activated by antigen and migrating to primary follicles where they interact with follicular dendritic cells (FDC). The FDC carry on their surface antigen, forming complexes with antibodies and components of the complement system, which are critical to the selection and maturation of high affinity antibody. After proliferative expansion discrete dark and light zone are identified. The dark zone contains rapidly cycling centroblast, whereas the light zone harbours resting centrocytes derived from the centroblasts, FDC that sequester the antigen, and antigen specific T cells and macrophages (figure 6).

Within GC, three processes are essential to the maturation: somatic hypermutation, selection of high affinity variants and class switching. The process of somatic hypermutation that occurs in proliferating centroblasts further diversify the Ig repertoire created by V(D)J recombination by introduction of nucleotide changes into the variable regions of the immunglobulin V genes.

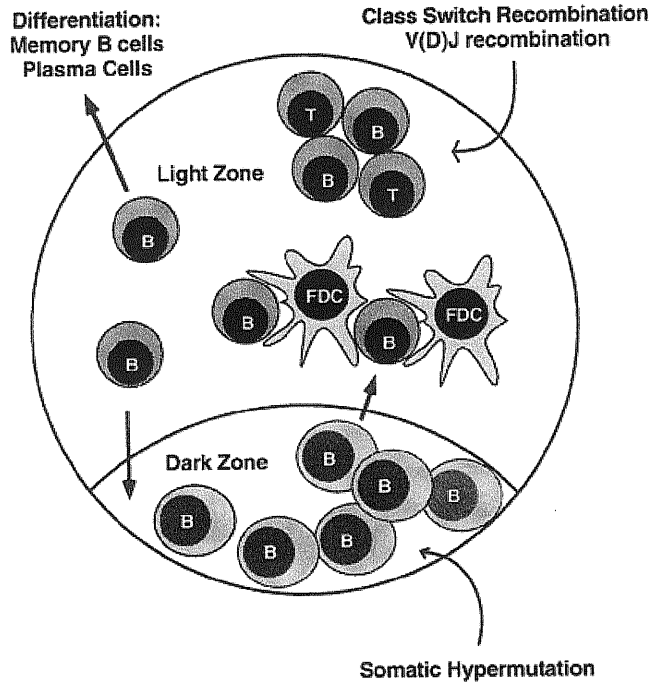


Figure 6: Schematic representation of the germinal centre

The hypermutation mechanism involves a mutational rate of 1 per 1000 base pairs per generation, implying that for each cell division almost one mutation takes place in either the H- or L- chain V region of an individual cell. Although the mutations can be scattered over the rearranged V genes, this process is not random, in that intrinsic hot

spots can be discerned (48). Replacement mutations are often over represented at the level of CDRs and underrepresented in the framework regions, whereas the opposite is true for silent (S) mutations. This process creates an enormous increase in antibody diversity in an age dependent manner (49). Although most of these mutations will either not affect or lower the affinity with which the antibody binds its ligand, some will increase it. Since proliferating GC cells are programmed to die unless rescued by signals involving antigen and antigen specific T-cells, only cells expressing high affinity antibody are positively selected in the light zone of the GC. Centrocytes selected by antigen may re-enter the dark zone and undergo further clonal expansion and somatic hypermutation or may exit the germinal centre to differentiate into memory cells or plasma cells. Memory B-cells are long lived and more easily triggered than virgin B-cells upon re-encounter with the same antigen, contributing to render secondary immune responses highly efficient.

A second aspect involved in the maturation of the immune response is class switching to Ig isotypes different from the IgM and IgD, characteristic of primary responses. Class switching occurs within the light zone of GC through a recombination event that produces a change in the class of antibody synthesised, hence a change in the effector function of the antibody produced, maintaining unaltered the specificity of binding. The change in antibody class is effected by a deletional DNA recombination event called switch recombination, which occurs between tandemly located sequences called switch regions (S), which are located upstream of each of the CH genes (see

figure 4). B cells that have undergone a class-switch recombination event usually have deleted all of the CH gene segments between the Sμ region and the S region of the CH gene that is expressed, since extra chromosomal circles generated by S-S mediated deletions and containing intervening DNA have been isolated (50). The induction of a switching process is dependent on the action of a specialised set of B-cell stimulants, like LPS and CD40L expressed on the surface of T-cells.

The targeting of the C region that will be expressed as a result of switching is largely determined by cytokines. IL-4 determines that switch will be to the εC region and to the γ4, in human, or γ1, in mouse, while IFN-γ determines switch to γ2a and TGF-β to α. A variety of evidences also indicate that once a B-cell has switched to a particular CH gene, it can, upon appropriate stimulation, undergo an additional switch to a downstream CH genes (51). Sequential switching appears to be a physiologically relevant process, but its precise role has yet to be determined.

It was recently discovered that, in addition to somatic hypermutation and class switch recombination, germinal centre B cell may also activate a mechanism, called receptor revision, which involves reactivation of the V(D)J recombination machinery. It appears that RAG genes in germinal centre B cells are actively suppressed when they bind avidly to an antigen but maintained or re-activated when the binding is weak. This may allow some low-affinity receptor to escape apoptotic death by the

acquisition a higher affinity receptor through replacement of L-chain (rarely the H-chain) (52).

In addition, some experiments performed with activation-induced cytidine deaminase (AID) deletion and overexpression suggest that this enzyme participate in both hypermutation and class switching depending in part on a common molecular mechanism. Overexpression of AID increase the class switching from IgM to IgA without cytokine stimulation, and AID deficiency not only blocks class switching in the immunised AID deficient mice, hence, abrogate hypermutation (53).

CDRs

The CDRs structurally form loops that vary in length as well as in sequence. Analysis of the relationship between the sequence and the three dimensional structure of the antibody combining sites revealed that, except for VH CDR3, the other loops have a restricted number of main chain conformations or canonical structures (54, 55). The canonical structure formed in a particular loop is determined by its size and the presence of certain residues at key sites in the loop and in framework regions. The conformation of VH CDR3 shows some regularities, from which rules relating sequence to conformation can be stated (56, 57), but the longer the CDR3 the more imprecise the conformation can be predicted.

Although all loops participate in antigen recognition the diversity of the antibody is mainly given by the VH and in particular the aromatic residues, specially the tyrosine, play a crucial role (58). In addition, another study analysing the buried surface between CDRs and antigen evidenced that L1, L2, L3, H1, H2 and H3 were responsible for the antigen surface contact in a 9%, 4%, 21%, 10%, 23% and 29%, respectively, whereas only a 4% corresponded to the framework (13).

CDR3 diversity

The H3 region is much more variable in length and sequence than the other loops (5, 59) because it is formed by the rearrangement of the V-D-J genes during the immunoglobulin heavy chain assembly. The presence of multiple V, D and J gene segments, the recombination and the inaccuracy of this process contribute to the generation of diversity of an antibody repertoire.

The extraordinary diversity of H3 may arrive to a number higher than 10^{14} different peptides, according to the calculation done by Sanz I, taking in account several mechanisms involved in this process (60). He analysed more than 500 H3-sequences and explained their diversity including alternative mechanism like inverted D segments, D-D fusion, gene conversion and the utilisation of DIR genes, longer sequences interspersed amongst the functional D segments that are flanked by 12bp and 23 bp spacer RSS (recombination signal sequence). In contrast, recent evidences

obtained by sequencing the D locus and by comparison of the complete sequence of this with a database of rearranged sequences, disagree for the use of alternative mechanism in the repertoire of natural antibodies (17). However, several analyses, including the latter's, agreed with D segments that can be used in all three reading frames. Moreover, the studies coincide in determining a biased utilisation of D segments as well as in the reading frame (17, 60, 61). Eventhough, the mechanism and degree of diversity are still unclear; its consequence might be enough for large antigen recognition (62,63).

The CDR3 encompasses the 3' end of VH, all of D, and the 5' end of JH. It often contains N nucleotides (figure 7), which are randomly inserted at both the VH-D and D-JH junctions by the enzyme terminal deoxytransferase (TdT) at the moment of recombination (64).

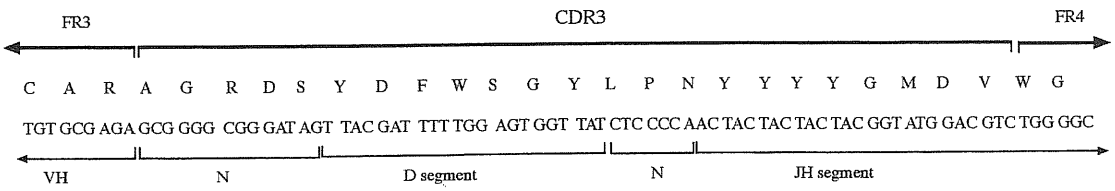


Figure 7 Example of a VH CDR3 and the localisation of the N nucleotides.

The introduction of junctional nucleotides is developmentally regulated; N insertions are found in 68% of fetal B cells, 86% of neonatal B cells, and in 91% to 100% of mature adult B cells. Another source of diversity in CDR3 appears to be random

deletion by nucleases of the terminal nucleotides of rearranging VH, D and JH genes (65,66). In addition to N nucleotides at junctional sites in which no nucleotide deletions occur, another repairing process may take place, introducing sequences known as "P" nucleotides. These sequences, usually only one or two bases long, are complementary copies of the last nucleotides of the coding region they adjoin. They may be the consequence of repair of asymmetrical breaks in the hairpin ends of V genes undergoing recombination. All of these variables contribute to the CDR3 H chain "fingerprint" before somatic mutations, that further distinguish and individualise V genes mainly by point mutations, which only rarely involved deletions and/or insertions. The resulting CDR3 sequence is unique to each rearrangement, and that therefore identifies individual B cells or clonal B-cell expansions.

Interestingly, analysis of large numbers of rearranged genes indicates that few germline segments dominate the mature repertoire and that the sequence diversity encoded by these segments is focused on residues at the centre of the antigen binding site. With somatic hypermutation, diversity spreads to regions at the periphery of the binding site that are highly conserved in the primary repertoire (67, 68, 69). This complementarity has probably been selected by evolution as an efficient strategy for searching sequence space (70, 71).

Phage display libraries

The advent of monoclonal antibody technology (72) showed that immunoglobulins have a wide range of biological activities and specificities, useful in treatment of disease, diagnosis and research.

Since the variable domains of the immunoglobulins interact directly with the antigen, technique aimed at the production and selection of immunoglobulin fragments, rather than the complete molecule have been developed. However, only after the derivation of single-chain molecules (scFv) (73,74) and after showing that functional antibody fragments could be produced and secreted by *Escherichia Coli* (75, 76) this technology became important.

A scFv is a single protein, which contains one VL joined to one VH by a flexible linker. Thus, with the complete antigen-binding site of an antibody and the same monomeric binding affinity as the parental monoclonal antibody (73, 74). All these features and the small size, make a scFv a molecule with high potential in the medical and biotechnological field.

Phage display technology has been developed as a means of making antibodies in vitro (77, 78, 79, 80, 81). In general, the affinity of the antibodies isolated is proportional to the initial diversity size of the library used for selection (80).

Correctly folded antibody fragments (Fv, scFv or Fab) are expressed by routing the nascent antibody chain(s) to the periplasm of the bacterium (75, 76), where the intradomain disulphide bridge between the two beta-sheets is formed, and the VH and VL will pair. Periplasmic expression in the bacterium resembles the natural production route in the endoplasmic reticulum (ER) of the lymphocytes.

Antibody displayed on filamentous fd phages is accomplished by fusing the coding sequence of the antibody variable (V) regions to the amino terminus of the phage protein p3. When antibody V-genes are cloned into phage display vectors, based on either phage or phagemid, functional antibody fragments are expressed on the surface of infective particles while the encoding genes reside within the phage particles. In the case of phage vectors, the gene encoding the displayed protein is included in the phage genome and as a result all phages will display the fusion protein, and will contain only the recombinant phage genome. Furthermore, each copy of p3 will diplay a recombinant protein. In contrast, in the case of phagemid vectors, the recombinant fusion protein is encoded on a plasmid (phagemid) which also contains the packaging signal. This phagemid makes large amounts of the recombinant display protein, but is unable to make a phage unless the bacteria carrying the phagemid also contains a helper phage, which supplies all the other proteins required to make a functional phage. Helper phages are essentially normal fd phages with a disabled packaging signal and carrying a different antibiotic resistant gene. The disabled packaging signal does not prevent the helper phage from making phages when alone

in a bacterium, but in the presence of a phagemid, which contains an optimal packaging signal, the phagemid will be packaged in preference to the helper phage. As a result, phagemid preparations will produce phenotypically and genotypically coupled phages, where the p3 can be whether wild type or recombinant.

Two different designs for the display of antigen binding sites have been successfully demonstrated: single chain Fv fragments (82) and heterodimeric Fab fragments (83) fused to the amino-terminus of p3, a protein of the phage that allows the bacterial infection through the F' pilus. Heterodimeric Fab fragments can be assembled on the surface of phages by linking one chain to the phage coat protein (i.e., VH-CH1) and secreting the other into the bacterial periplasm (i.e., VL-CL), where the two chains associate.

The linkage between antibody genotype and phenotype allows the enrichment of antigen specific phage antibodies, using immobilised or labelled antigen. Phages that display a relevant antibody will be retained on a surface coated with antigen while non-adherent phages will be washed away. Bound phages can be recovered from the surface, reinfected into bacteria and regrown for further enrichment and eventually for binding analysis.

Phage antibody libraries

With phage display, antibodies can be made completely in vitro, by-passing the immune system and the immunisation procedure, allowing in vitro "tailoring" of the affinity and specificity of the antibody. The concepts behind phage antibody library selections are identical to those used for any of the other biological molecular diversity techniques; the creation of diversity, followed by a series of recursive cycles of selection on antigen (each of which involving binding, washing, and elution) followed by amplification.

There are two main routes to antibodies libraries via phage display. The first is to select antibodies from libraries prepared from immune donors e.g. immunised animal or, in some instances, human immune B-cells. Immune libraries have two main advantages. First, they are highly biased towards V-genes that encode antibodies against the immunogen (if IgG specific primers are used), which means that relatively small (10^5 clones) libraries can be successfully screened. And second, many of the genes will encode both affinity matured antibodies (increasing the number of high affinity antibodies in the library) and highly specific antibodies.

Disadvantages of this strategy, however, include the time required to immunise animals, the unpredictability of immune response to the antigen of interest and the lack of immune response to some antigens (self-antigens or toxic molecules). In addition it involves the construction of the library itself. Moreover the scrambling of

heavy and light chains which occurs during library construction makes it almost impossible to isolate "original" VH-VL combinations with high specificity and affinity. Most of the VH-VL pairs formed are created de novo, although the success obtained with this method indicates this may not be very important. In particular, most of the specificities are provided by VH domains, several VLs could prove to be good partners for Ag recognition (58). Furthermore, in selection from very large naïve phage antibody libraries (80), different antigens always selected different VH genes, whereas some VL genes were found in scFvs binding to different antigens, suggesting a surrogate role for VL.

The second route to antibody libraries contains naïve or semi-synthetic libraries. In both cases, the ultimate goal is to assemble libraries of very large diversity size that high affinity antibodies of any specificity can be selected (79, 80, 81, 84).

The primary immune response involves a large array of IgM antibodies that recognise a variety of antigens. The murine naïve repertoire has been estimated to contain $<5x10^8$ different B-lymphocytes while the human repertoire may be 100 to 1000 times' bigger (85). This array of antibodies may be cloned as a naïve repertoire of rearranged genes, by harvesting the V-genes from the IgM mRNA of B-cells of unimmunised human donors, isolated from peripheral blood lymphocytes (PBLs), bone marrow, spleen cells, or from animal sources. The functional VH genes can be amplified either from IgM mRNA (84) or total mRNA (80). However, in an early

paper on phage antibody libraries two libraries were made in which VHs were derived either from IgM or igG (using respectively μ or γ primers). Interestingly only the IgM derived library yielded binders to the antigens used for selection (84).

Provided that the repertoire is highly diverse, the greatest advance of this approach is: that one single library can be used for all antigens; high affinity human antibodies can be isolated; antibodies to self, non-immunogenic or toxic substances can be generated; and antibody generation does not take much time (2 weeks).

The first of such "single pot" repertoires was made from the PBLs of two healthy human volunteers and contained $3x10^7$ clones. From this pool antibodies to over 25 different antigens were isolated. The antigens were either foreign, such as bovine serum albumin, and hapten, or self, such as thyroglobulin, CD4 etc. Recently, many larger scFv repertoires have been made by "brute force" cloning (80) yielding $1.4x10^{10}$ independent antibody clones. From this library, antibodies with affinities typical for a secondary immune response were isolated (average Ka of around 10^8 M^{-1}. Thus, neither immunisation nor affinity maturation are needed for the generation of high affinity (human) antibodies.

However there are also some disadvantages in particular the largely unknown and uncontrollable contents of these naïve libraries, and the lower affinity of the antibodies where smaller-sized repertoire are used.

Library performance is demonstrably improved by increasing size and diversity. However, other factors including expression levels, folding, and toxicity to E. coli may all reduce the functional repertoire size. The advantage of synthetic V-gene repertoires is that they can be designed and tuned to optimise library function. To construct a synthetic antibody library, V-genes are assembled by introducing a pre-determined level of randomisation in the CDRs (or possibly also bordering FR regions) into ideally germline V-gene segments (86). Several synthetic libraries have been designed with the heavy chain CDR3 sequence encoded by oligonucleotide primers encoding a stretch of randomised amino acids residues (79, 87, 88). In fact, from structural studies, it has become apparent that five of the six CDR regions (all but the CDR3 of VH) have limited structural variation, and frequently follow a certain canonical fold. Thus the CDR3 of VH is the most diverse loop, in composition length and structure, and this is the region that should be partially or completely randomised using oligonucleotide-directed mutagenesis or PCR-based methods.

Antibodies with affinities comparable to those obtained using traditional hybridoma technology can be selected from large naïve antibody libraries, and the affinity of these can be further increased, to levels unattainable in the immune system, by using the selected antibodies as the basis for subsequent libraries and selection.

Although a number of different phage(mid) antibody libraries have been published, the number which are 'naïve' is relatively small.

In general, the affinity of the antibodies selected is proportional to the size of the library, with Kds ranging from $10^{-6/7}$ for the smaller libraries (84, 87) to 10^{-9} for the larger ones (79, 80, 81). Antibodies selected from immunised libraries tend to have higher affinities for the antigen used for immunisation from an equivalent library size.

Recently, a new approach to easily increase diversity has been introduced. The method involves the creation of a primary phage scFv library in a phagemid containing two lox sites, which are unable to recombine with each other. Infection of Cre recombinase expressing bacteria by such a primary library at a high multiplicity of infection, results in the introduction of many different phagemids into the same bacteria cell where different plasmid molecules take place, generating a large number of new VH/VL combinations. On the basis of the observed recombination, Sblattero et al calculated a potential diversity between 10^{11} and 10^{12}, limited only by the volume of culture medium used (89). This secondary library created by in vivo recombination was validated by selection against a large number of different protein antigens achieving with good affinities.

CHAPTER 2

Results

Introduction

Antigen-binding sites are formed by the surface derived from the parallel association of VL and VH regions. The variable domains resemble each other in both secondary and tertiary structure having each two layers of β-sheets supporting six hypervariable loops, three of the heavy (H1, H2 and H3) and three of the light chain domains (L1, L2 and L3) that constitute the antigen binding loops. VL and VH domains associate noncovalently to form a β-barrel structure, which place the six CDRs close to each other at the amino terminal end of the VL-VH dimer, also known as the Fv region. In general, all loops participate in antigen recognition. However, it has been shown that in many cases the buried surface of the antigen is mainly occupied by VH sequences, specially loop H3 (13, 58).

It is also known that L and H chains differ in their secretory properties when expressed in mammalian cells. When transfected alone in lymphoid cells L chains dimers are secreted in a high level, even though, the VL-VL interaction surface is 18% lower than a normal VH-VL. On the other hand, in the absence of L chains, H chains can not be expressed (35, 36, 90, 91) due to the association with BiP and GRP94 within the endoplasmic reticulum (38, 92). However, a mouse myeloma mutant can secrete CH1-truncated heavy chains without L chain (36, 93). Moreover, isolated VHs tend to form dimers when

renatured under physiological conditions (94). These results suggest that VHs may interact spontaneously in the absence of CH1.

Taking into consideration that VL and VH are structurally similar, that VLs can form dimers, and that there is no reported evidence of VH/VH associations, one aim of this project was to verify if VH domains could associate with each other, maintaining a proper folding. Secondly, it was of interest to know if the VH domain could keep its property to bind an antigen or to form new antigen recognition sites. Thus, we wanted to know if novel biotechnological applications could be obtained from the manipulation and induction of association of single VH domains.

VH dimer simulation

The first approach followed to see if VHs could interact with each other was to model a VH-VH dimer by computer simulation. For this purpose, we used the Swiss-PDB viewer (http://www.expasy.ch/spdbv/text/getmac.htm) and the following immunoglobulin structures (PDB entry): 1ADQ (Homo sapiens), 1VFB (Mus musculus), and 1IGM (Homo sapiens).

In order to do the models and since both VL and VH tertiary structures are very similar we assumed that VH could be aligned over the VL backbone without significant changes in the β-barrel structures. Furthermore, we kept the same relative parallel orientations,

thus neutralising the hydrophobic interface surfaces of both domains. We considered these models could put in evidence obvious restrictions between the surfaces of the interacting domains that could be taken in consideration in other experimental approaches.

We initiated by replacing the VL domain from three different VL-VH couples by the same VH of the pair used to form homodimers. Figures 8 shows different views of homodimeric models showing no steric hindrances between the interacting surfaces. However, in two cases a loop produced a clash (differently coloured in figure 8a and c). Interestingly, these loops corresponded to H3 which, as previously described, can be very variable in length (3-30 aa) as well as aminoacid composition. The short H3 of figure 8b (8 aminoacids long) does not produce a clash between the structures suggesting that the length of H3 could be a restricting factor in permitting VH/VH association. In addition, we did not find differences in the behaviour of mouse and human VHs when doing the modelling.

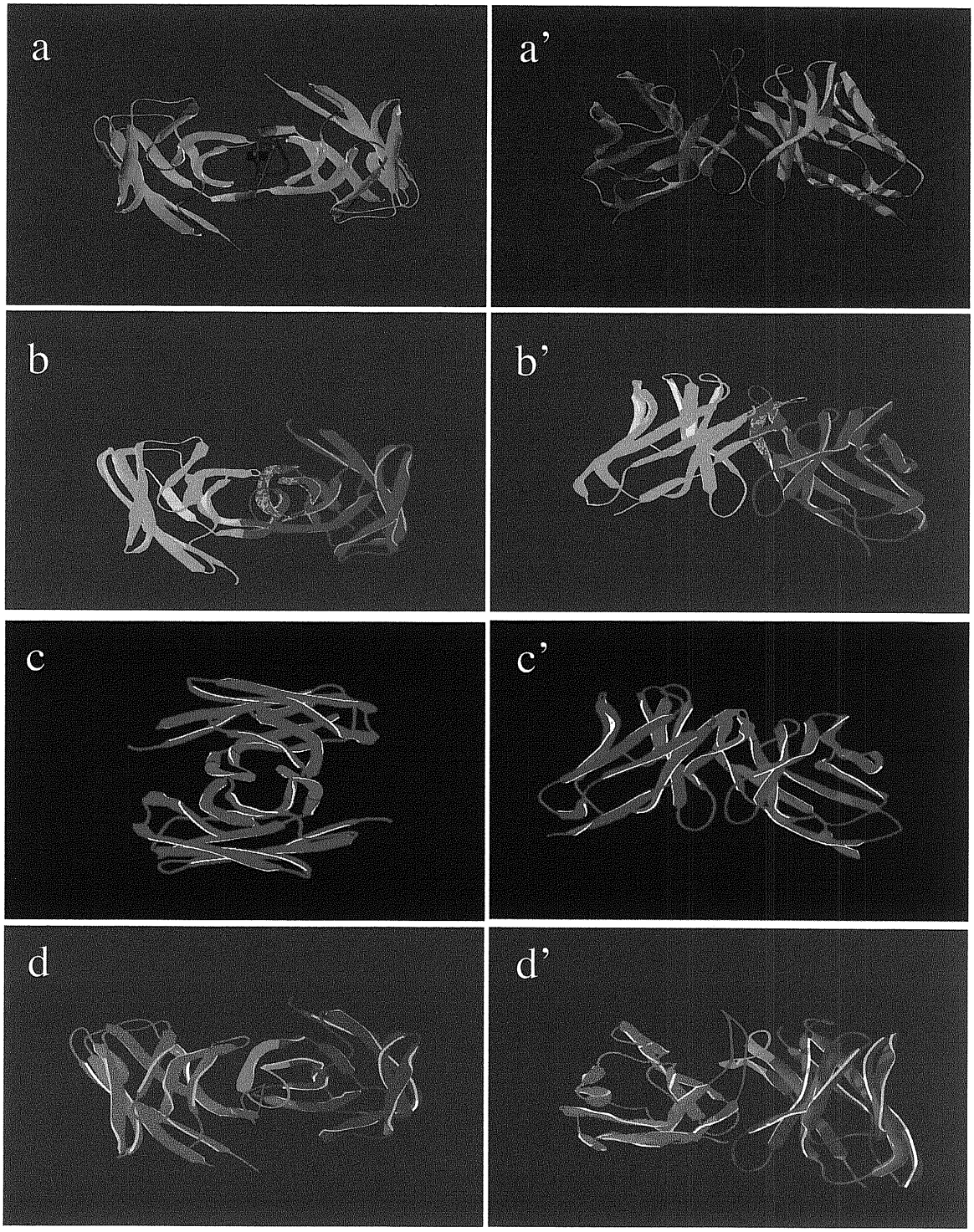


Figure 8 Models of VH-VH association. Top (a, b, c, d) and front (a', b', c' and d') view of two (a, b) human homodimers (1ADQ, 1IGM) and of a mouse (c) homodimer (1FVB). d represents a human heterodimer (combination 1ADQ and 1IGM).

To address the question whether dimerisation was only allowed for homodimers we also simulated a heterodimer using the same criteria as for homodimers. In this case, instead of replacing the VL with the same VH from the template, a different VH was used. The 1IGM structure was used as template and its VL replaced with the VH^{1ADQ}. As can be observed in figure 8d, similarly to the homodimers, no evident clashes, with the exception of the H3 loops, were present in the model. Both homodimer and heterodimer simulations suggested that H3 might create problems for VH/VH associations. Hence, a small H3 could avoid the potential clashes and allow the dimerisation. To test this hypothesis we decided to look for a natural VH with a small H3 to further investigate the possibility of VH/VH associations.

H3 Fingerprinting

In order to search for a VH with a small H3, several VHs were obtained by RT-PCR from peripheral blood lymphocytes RNA. The VHs were cloned and the length of their H3s analysed. Although H3 is a very variable sequence, it is located between FR3 and FR4, which are relatively conserved in most VH family genes. The cloned VH were further amplified using external primers followed by an additional extension cycle performed with a 5'γ^{32}P end-labelled oligonucleotide that annealed to the conserved FR3. The products were loaded in a sequencing gel, and the sizes of H3 calculated using a ladder of bands from a sequence. As shown in figure 9 different pattern of H3 lengths were

observed. The smallest H3 (orange triangle) corresponded to 3 aminoacids. The aminoacid sequence of this VH (figure 8c) corresponded to family VH1 (hv12635) and confirmed the 3 aa length (Pro-Gly-Gly) of H3. This VH was selected for further studies.

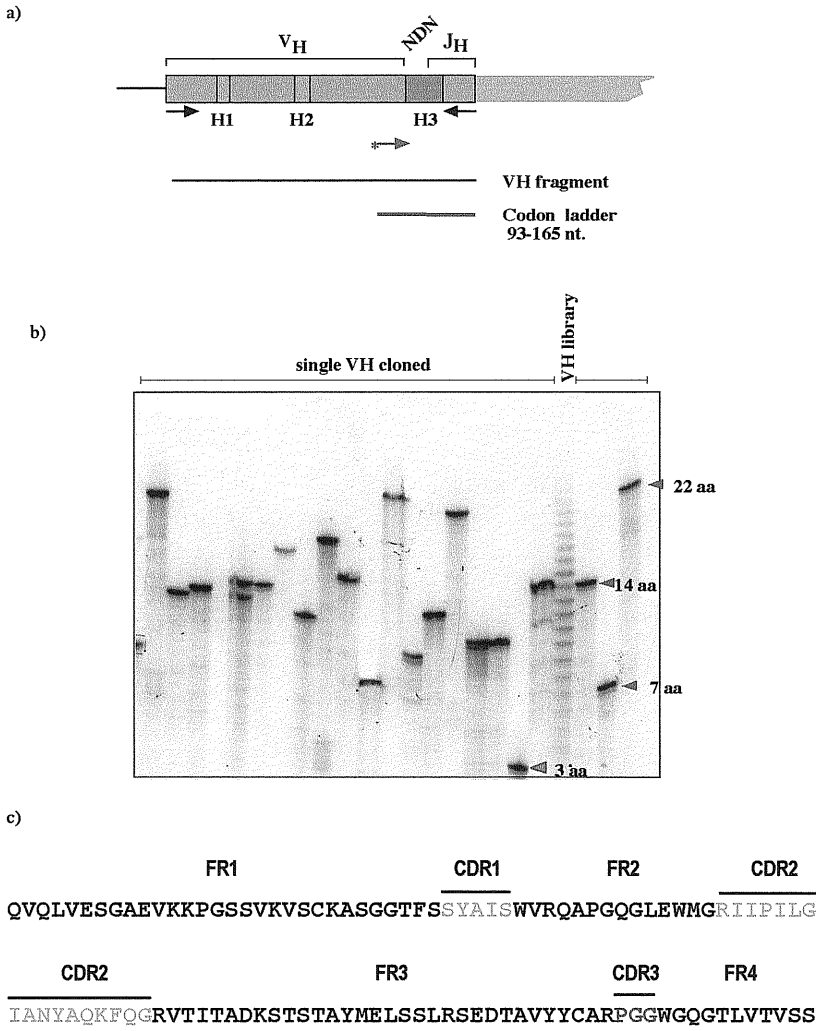


Figure 9 CDR3 fingerprint analysis of VHs. a) Schematic representation of a VH region. The position of the PCR primers (black arrows), the ^{32}P-hV2 (red arrow with asterix) extension primer, and the expected fragment product size (red line) are indicated. b) Autoradiography of labelled fragments of single VH clones. Triangles show different H3 lengths. c) Sequence of VH with H3 of 3 aa (PGG).

VH dimer simulation with a short H3

With the porpoise of knowing whether the selected VH containing the 3 aminoacids H3 (VH^{PGG}) fulfilled the requirements for VH dimerisation, a model was produced using the Swiss-Model package (95, 96, 97). We first chose five templates, with the highest sequence homology scores, using the BLASTP2 program, with known 3D structures from the ExNRL-3D database. The following templates were selected (PDB entry): 1AD9H, 1AXSB, 1DHAH, 1GC1H and 1MLBB. It is known that H3 modelling is difficult to perform in particular when long loops are involved (4). We assumed that the modelling of the VH^{PGG} H3 should be of enough confidence since it was extremely short. To further visualise if H3 clashes could be prevented with this model, the resulted VH^{PGG} structure was aligned over both VH^{1ADQ} in the homodimer scaffold where H3s clashed to each other (figure 10).

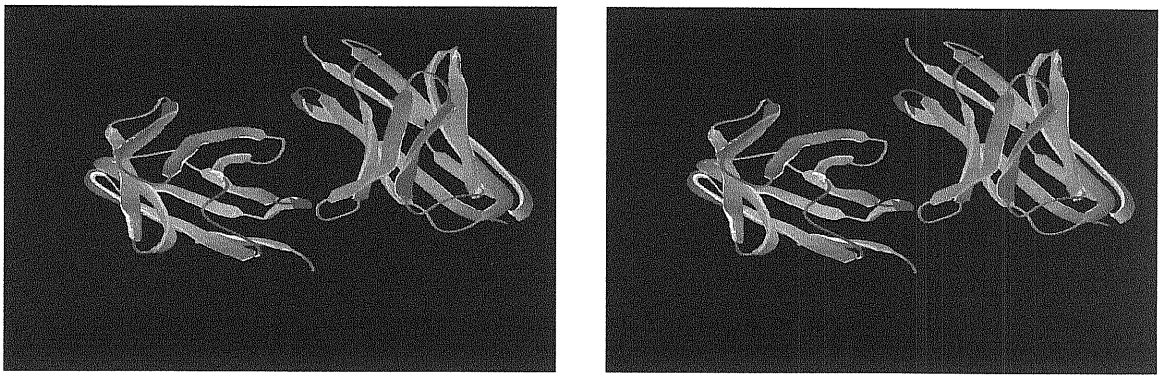


Figure 10 Model of VH-VH association with a short H3 (VH^{PGG}). VH monomer has a CDR3 length of 3 aminoacids composed of PGG (proline-glycine-glycine) (red coloured loop). Two different views are shown.

First criteria for correct folding and dimerisation

Secretion from mammalian cells can be consider as a criteria for proper folding, Indeed, quality control mechanism operate within cells to generate that only rightly folded and assembled molecules are produced and expressed (98). In addition, mammalian cells can not secrete H chains in the absence of L-chains in part because of H chain association to BiP and also because of the hydrophobic surfaces that create solubility problems (99, 100). Therefore, in order to know whether VH^{PGG} folds and assembles correctly, we tested its capacity to be secreted from mammalian cells. We first decided to express the selected VH^{PGG} fused to an Ig dimerising domain, like in the SIP (small immune protein) system (101), which was developed in our lab.

The SIP system consists of a scFv coupled to an Ig dimerising domain [CH3 for γ or α, CH4 for ε]. The protein product is well secreted when transfected into mammalian cells (101). One advantage of this system is that the dimerising domain can also be used as a tag, and therefore, the protein easily purified and followed by ELISA and/or western blotting. When replacing the γCH3 domain with the dimerising domain of the ε-H chain constant region, εCH4-s2, the protein is expressed as a dimer covalently linked by a disulphide bridge. εCH4-s2 is derived from alternative splicing of the 3' part of the ε locus, that contains a C-terminal cysteine involved in an inter H-chain disulphide bridge (45, unpublish results). The VH-ε CH4 was clearly identified as a dimer of approx. 80

kDa when analyzed by western blotting performed under non-reducing conditions (figure 11).

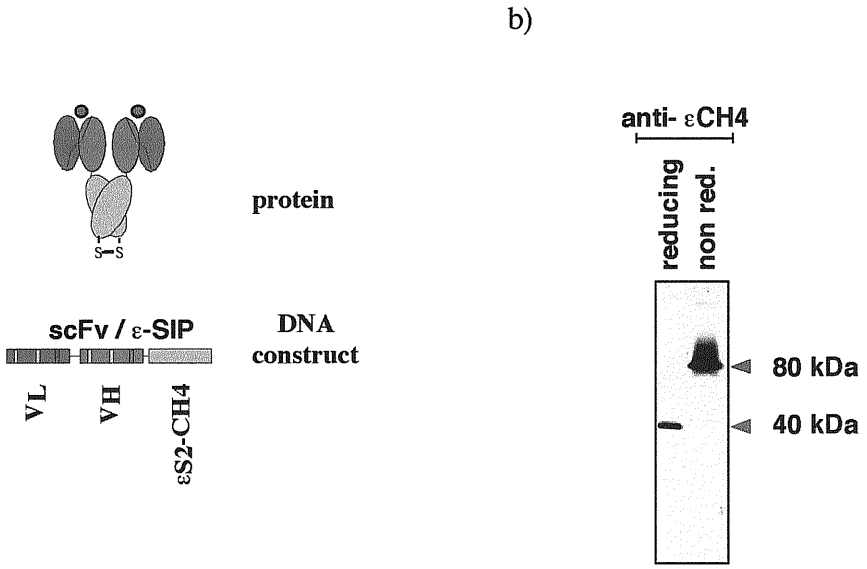


Figure 11 a) Schematic representation of a SIP construct containing a mouse VL^{6C6} and VH^{6C6} (6C6-SIP) and the corresponding protein. b) Western blot analysis of the 6C6-SIP, using a mouse anti-εCH4 antibody and an anti-mouse Ig-HRP-conjugated antibody, developed by chemioluminiscence. Reducing and non-reducing patterns of the protein are shown.

We decided to use the εCH4-s2 dimerising domain to exploit its ability to stabilise the dimer even if weak interactions between VHs occurred. The construct consisted of VH^{PGG} directly fused to εCH4-s2 and was used to transfect sp2/0 cells. Supernatants of positive clones (detected by ELISA) were analysed by Western blotting, in reducing and non-reducing conditions, using an anti εCH4 antibody. If VH/VH interaction occurs it can be expected that VH^{PGG}-ε CH4 should be produced as a dimeric protein of aprox. 50 kDa. that under reducing conditions, should appeared as monomer of about 25 kDa. The

results shown in figure12a are in agreement with the expected sizes. Practically all the material secreted corresponded to the dimeric protein.

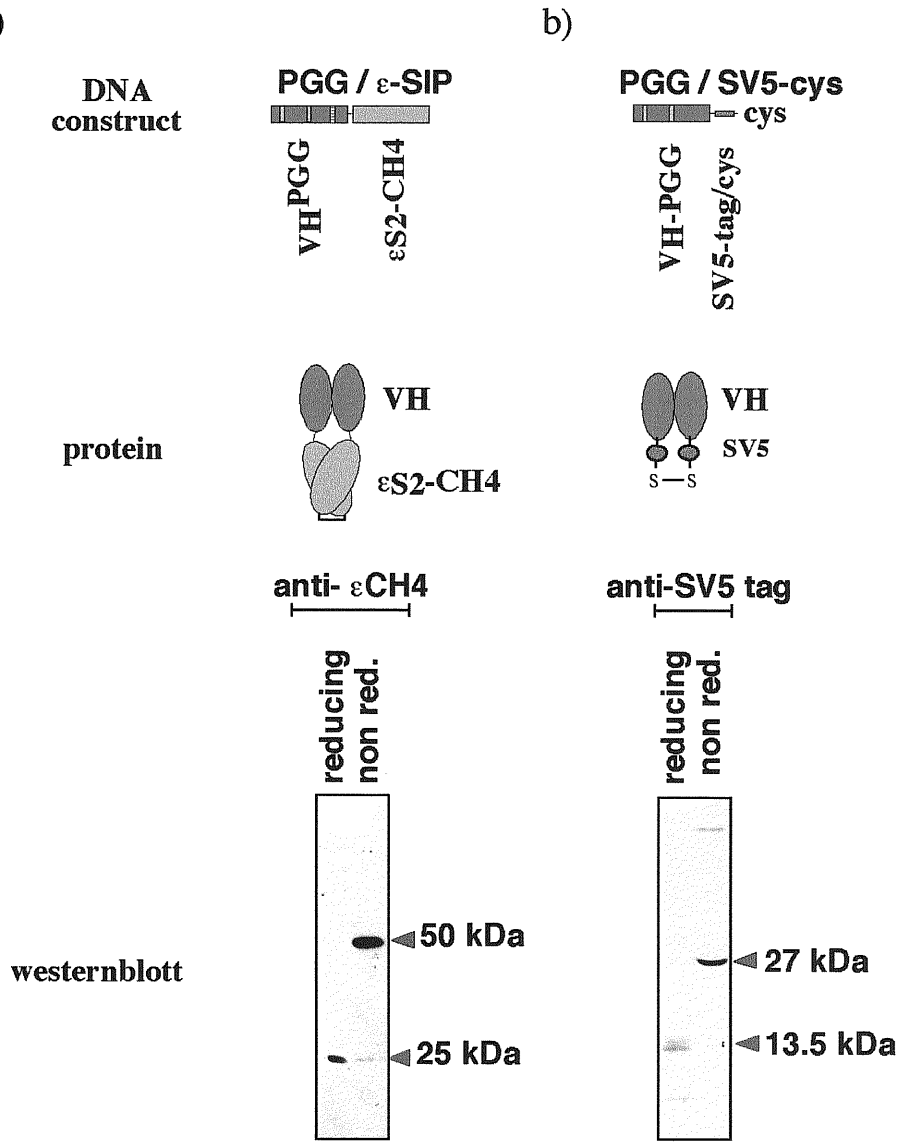


Figure 12 Schematic representation of constructs and proteins of VH^{PGG} in two different formats and their corresponding western blot analysis. a) PGG/ε-SIP western blot analysis was performed using a mouse anti-ε CH4 antibody and an anti-Ig mouse-HRP conjugated antibody. b) PGG/SV5-cys protein was identified with a monoclonal anti-SV5 antibody and an anti-mouse Ig-HRP conjugated antibody. Both proteins were visualised by chemioluminiscence. Molecular weights of reducing and non-reducing samples are indicated.

This result suggested that the domains were correctly folded, and more importantly, that the assembly allowed by the cells prevented the exposed hydrophobic areas to remain impaired. This was a good indication that the VH associations could take place.

However, the presence of the ε CH4 dimerising domain could have biased the formation of the VH/VH homodimer. We therefore expressed the single VH using the 14 aa long SV5 (simian virus 5) peptide tag (which does not dimerise) fused to the carboxy end. We also decided to place a C-terminal cysteine after the SV5 tag sequence to test the orientation of the domains and to facilitate visualisation of dimers. A interchain disulphide bridge was expected if the C-terminal parts were close to each other due to the parallel orientation of the two interacting VH domains (figure 12). The analysis of the expression of this construct was done by westernblotting using a monoclonal antibody against the SV5 tag peptide. Interestingly, under non-reducing condition the VH was exclusively found as a dimer (27 kDa) supporting the hypothesis that C-terminal portions were oriented in a similar manner as in VL-VH pairs (figure 12). In addition, the fact that a single band was observed indicates that no association with other proteins containing reactive cysteines occurred.

In order to confirm these results we planned a second strategy deleting the C-terminal cysteines and assaying the oligomerisation-state by in vitro crosslinking. The crosslinking-reagent used (DSP: dithiobis succinimidylpropionate) has the pecularity of

being divalent with an internal disulphide bridge. The dimeric protein would then migrate with a mobility of 27 kDa whereas monomer would migrate as 13.5 kDa under non-reducing conditions (figure 13). This strategy eliminates the formation of covalent disulphide bridges in vivo between VHs and looks directly at the oligomerisation-state of the already secreted material, thus, avoiding a possible bias due to the presence of the cysteines. In addition it also allows to detect higher degrees of oligomerisation.

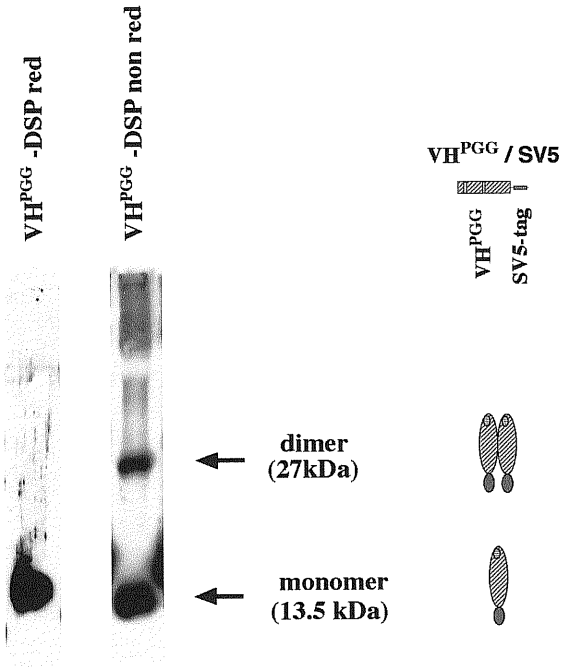
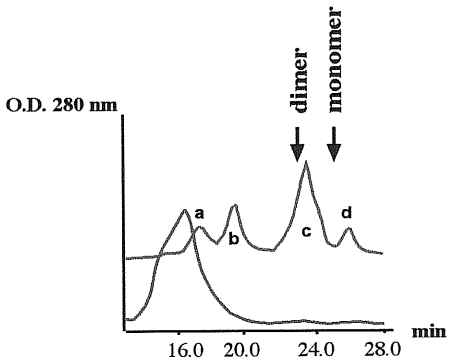


Figure 13 Westernblotting of crosslinked mammalian cells supernatant transfected with VH^{PGG}-SV5. Dialysed surnatants were crosslinked with DSP reagent and loaded into a 15% PAGE-SDS gel in reducing (red) and non-reducing (non-red) conditions. Arrows indicate monomers (VH^{PGG}) and dimers (VHD^{PGG}).

As shown in figure 13 the crosslinking assay indicates that the dimeric form of VH is the main oligomerisation-state. In contrast to the previous results a large amount of monomer appeared in this analysis. We have observed variability in the relative range of dimer/monomer depending in part on the concentration of protein and reagent. We also extended this analysis to other VHs and found that in all cases the main products obtained were dimers and monomers. It is important to mention that we never found higher order oligomers as main product.

We also attempted to analyse the VH dimer (called also VHD) by gel filtration. We have many difficulties in obtaining reproducible results; in most cases the purification yield ended with the formation of large aggregates. This was particularly true when the material was concentrated due to the purification procedure. In figure 14 it can be seen that only a small fraction of the protein constitute dimers while the rest are aggregates with molecular weights higher then 67 kDa. This may be due to the intrinsic solubility of VHs that somehow expose the hydrophobic patches located at the bottom part of the β-barrel, opposite to the CDRs normally located at the top, which are always neutralised by CH1 in the complete Ig (14, 102). This aggregation might have been helped by the lack of a linker between the interacting domains as suggested by Glockshuber et al (103). Nevertheless the result shows that the non-aggregated material consist mainly of homodimers.

a)



b)

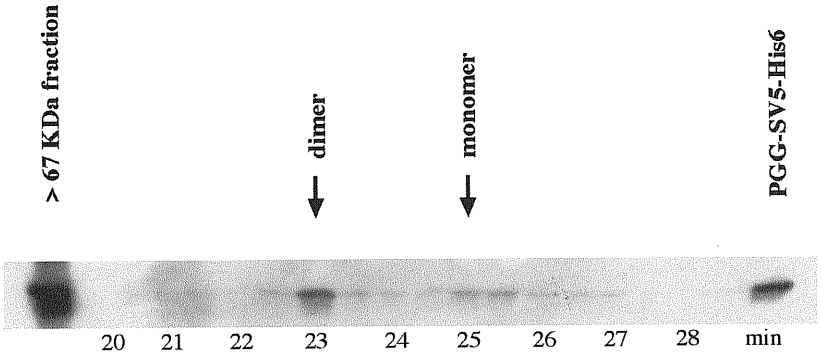


Figure 14 a) Gel filtration Superdex 75/30 profile of VHD^{PGG}/SV5-His6 produced by sp2/0 cells and purified on Ni^{++} column. Molecular weight markers: a=67 kDa, b=42 kDa, c=25 kDa and d=13.7 kDa. Column was run in PBS. The blue line corresponds to the sample and the red one to the molecular weight markers. b) Western blot analysis of samples corresponding to the region of the dimmer (27 kDa) and monomeric (13.5 kDa) forms. Direct mammalian cell surnatant producing VHD^{PGG} /SV5-His6 is shown in the last lane and aggregation products (out of the column exclusion range) in the first one.

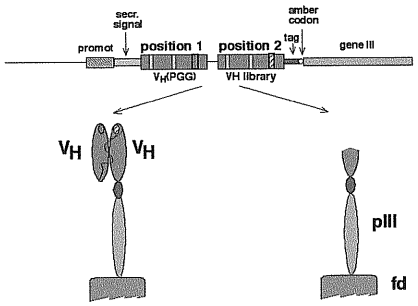
Our results allowed us to conclude that VH^{PGG} interacts to itself mostly as a dimer (VHD^{PGG}) with a similar orientation to the VL-VH scaffold, most likely through the same hydrophobic surfaces localised in the interface interacting with VL.

VHD phage display

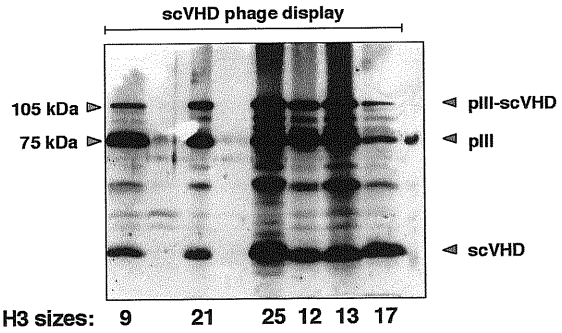
As mentioned before fd phages can display functional fragments of immunoglobulin, mainly scFvs, when fused to the pIII protein (77, 78, 79, 80, 81, 82, 83). With the aim of investigating the possibility that VHD could be used as specific binders we decided to use phage display of scVHD to select for functional molecules. We constructed a library of scVHD, in which position 1 has VH^{PGG} and position 2 contained a library of human VH derived from PBL. This relative small size library (diversity 10^5) was first analysed for its ability to display the fusion protein. This was done by Western blot with the monoclonal antibody against the SV5 tag as well as with an anti-pIII serum. The expected size for the pIII- scVHD fusion is of around 105 kDa while for pIII alone is around 75 kDa. As shown in figure 15, we found that 80% of randomly analysed VHD-pIII fusions were properly displayed on the fd phage. However, the fact that 20% of the clones did not displayed also indicated that several sequences are either toxic for the bacteria or that they need an appropriate partner to be correctly folded and assembled.

One particular mouse derived VH (VH^{6C6}) that was previously shown in our lab to be unable to form homo VHD was used with VH^{PGG} to evaluate the possibility of being rescued for display. The results of figure 16 show that VH^{6C6} can not be displayed with VH^{PGG} producing a complete degradation of the VHD part of the fusion protein. While anti-SV5 does not detect any protein, anti-pIII shows out the degraded pIII moiety.

a)



b)



c)

```
PGG       QVQLVESGAEVKKPGSSVKVSCKASGGTFSSYAI--SWVRQAPGQGLEWMGRIIPI--LGIANYAQKFQG

CLONE8    QVQLVQSGAEVKKPGASVKVSCKASGYTFTSYGI--SWVRQAPGQGLEWMGWISAY--NGNTNYAQKLQG
CLONE21   QVQLVQSGAEVKKPGASVKVSCKASGYTFTSYGI--SWVRQAPGQGLEWMGWISAY--NGNTNYAQKLQG
CLONE3    QVQLVQSGAEVKKPGASVKVSCKASGYTFTSYDI--NWVRQATGQGLEWMGWMNPN--SGNTGYAQKFQG
CLONE10   QVQLVQSGAEVKKPGASVKVSCKASGYTFTSYAM--HWVRQAPGQRLEWMGWINAG--NGNTKYSQKFQG
CLONE23   EVQLVESGGGVVRPGGSLRLSCAASGFTFDDYGM--SWVRQAPGKGLEWVSGINWN--GGSTGYADSVKG
CLONE11   EVQLVESGGGLVQPGRSLRLSCAASGFTFDDYAM--HWVRQAPGKGLEWVSGISWN--SGSIGYADSVKG
CLONE13   QVQLVESGGGLVKPGGSLRLSCAASGFTFSDYYM--SWIRQAPGKGLEWVSYISSS--GSTIYYADSVKG
CLONE9    EVQLLESGGGLVQPGGSLRLSCAASGFTFSSYAM--SWVRQAPGKGLEWVSAISGS--GGSTYYADSVKG
CLONE24   EVQLVESGGGLVQPGRSLRLSCTASGFTFGDYAM--SWFRQAPGKGLEWVGFIRSKAYGGTTEYTASVKG
CLONE7    QVQLQQWGAGLLKPSETLSLTCAVYGGSFSGYYW--SWIRQPPGKGLEWIGEINHS---GSTNYNPSLKS
CLONE12   QVTLKESGPVLVKPTETLTLTCTVSGFSLSNARMGVSWIRQPPGKALEWLAHIFSN---DEKSYSTSLKS

PGG       RVTITADKSTSTAYMELSSLRSEDTAVYYCARP-G--------------G-------WGQGTLVTVSS--

CLONE8    RVTMTTDTSTSTAYMELRSLRSDDTAVYYCARVDG-IAEND--------A----FDIWGQGTTVTVSS--
CLONE21   RVTMTTDTSTSTAYMELRSLRSDDTAVYYCARCVG-VVPATPAPYYYYYG----MDVWGQGTTVTVSS--
CLONE3    RVTMTRNTSISTAYMELSSLRSEDTAVYYCARG------------IAVAG----TDYWGQGTLVTVSS--
CLONE10   RVTITRDTSASTAYMELSSLRSEDTAVYYCVRGNW-KYDGY---NYKYYG----MDVWGQGTTVTVSS--
CLONE23   RFTISRDNAKNSLYLQMNSLRAEDTALYYCARD----MDGSSSPLNYYYG----MDVWGQGTTVTVSS--
CLONE11   RFTISRDNAKNSLYLQMNSLRAEDTALYYCAKG----VVPAPRGGLYYYG----MDVWGQGTTVTVSS--
CLONE13   RFTISRDNAKNSLYLQMNSLRAEDTAVYYCARGK---LVPGATWAYYYDG----MDVWGQGRTVTVSS--
CLONE9    RFTISRDNSKNTLYLQMNSLRAEDTAVYYCAKS--------LSQQQTFNA----FDIWGQGTMVTVSS--
CLONE24   RFTISRDGSKSIAYLQMNSLKTEDTAVYYCTRDGGERFLEWLSINPDCYG----MDVWGQGTTVTVSS--
CLONE7    RVTISVDTSKNQFSLKLSSVTAADTAVYYCAR---GGVRGVLRSRKAPYG----MDVWGQGTTVTVSS--
CLONE12   RLTISKDTSKSQVVLTMTNMDPVDTATYYCARIISDNVWGSFRYNEDRNDRNDAFDIWGQGTTVTVSS--
```

Figure 15 Analysis of hetero-scVHD$^{PGG/library}$ [VH^{PGG}-VH library] displayed on M13 phages. a) Schematic design of the construct and the scVHD-pIII fusion protein displayed on M13 phage. Two different infective phages can arise after induction with helper phage: display of complete scVHD-pIII or pIII with degraded scVHD. b) Western blot analysis of individual isolated phage particles reacted with anti-SV5 Ab. H3 lengths of the different VHs in position 2 are shown c) Aligned sequences of different VHs compared to VH^{PGG} (top).

As control a scFv containing the same VH^{PGG} with an irrelevant VL shows proper display of the Fv-pIII fusions. When sequences from the positively displayed VHs were analysed, we found members of the VH1, VH2, VH3 and VH4 families, containing J regions of JH3, JH4 and JH6 segments. Interestingly, the H3 of these VHs had length ranging from 9 to 25 aminoacids (figure 15). These results clearly indicate that contrary to our expectation the length of H3 appears not to be a strong restriction for the ability of different VHs to form VHD with VH^{PGG}.

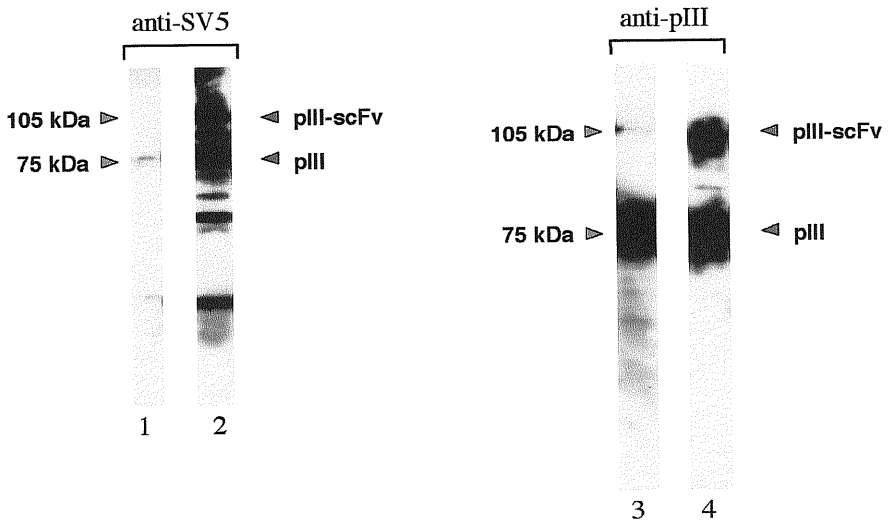


Figure 16 Westernblott analysis of hetero-$scVHD^{PGG/6C6}$ [VH^{PGG}-VH^{6C6}] compared to a scFv [VL-VH] used as positive control and developed through the SV5 tag (lanes 1 and 3) and pIII protein (lanes 2 and 4).

Selection of VHD binders

Several hints suggested that VHs dimerise like VL-VH scaffolds; however, the functionality was a main topic to address in order to find further usage of this novel type of binders.

To easily achieve some functional VHD we applied phage display technology, using two different vectors, pDAN3 and pDAN5. The difference between them is that the latter allows "in vivo" recombination of V regions between different single chain molecules within the same bacteria cell. This represents a technical advantage to increase the library size (89). The pDAN5 vector was developed from pDAN3 and it contains two different recombination sites (loxP511 and loxP) that permit reciprocal recombination of sequences flanked by these sites in cells that express the recombinase Cre. Both vectors were developed by Sblattero and Bradbury and kindly provided for our VHD studies.

Our first attempt to obtain VHD binders was performed by constructing a small size library on vector pDAN3 (pDAN5 was not available at that time) (figure17).

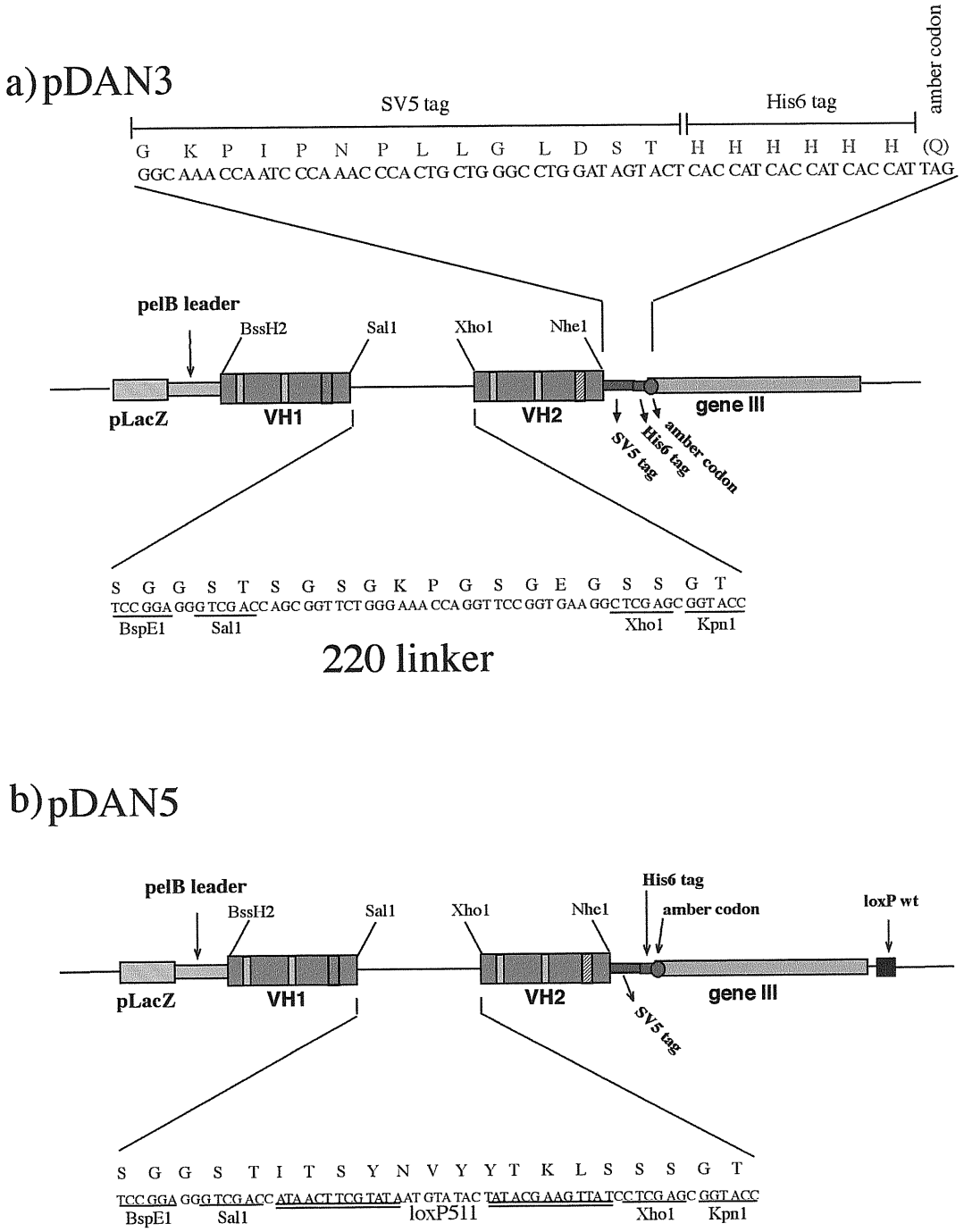


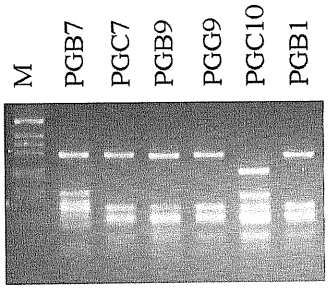
Figure 17 Schematic representation of the relevant region of the phagemid vectors pDAN3 and pDAN5. The sequence of the two linkers is shown. The SV5/His6 tag is common to both vectors.

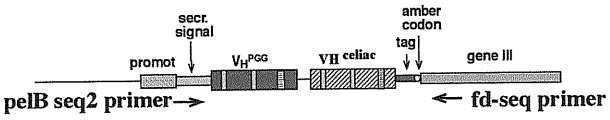
Using VH^{PGG} as a fixed partner in position 1 of the scVHD, we cloned in position 2 VHs derived from IgA of celiac disease patients that can be considered as natural immunised donors (this collection of VHs was gently provided by Sblattero, Bradbury and Marzari). The total library (celiac) size was of 10^5 recombinant colonies and was used to test the presence of VHD binders on immobilised wheat extract (mainly gliadin).

After two rounds of selection 96 individual clones were analysed by ELISA against a total wheat extract. Fifteen percent of the clones showed binding activity to the extract (figure 18b).

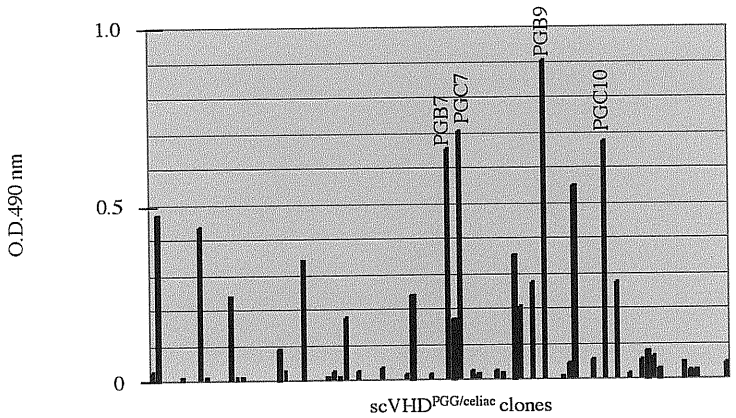
BstNI-fingerprinting analysis of the positive selected clones revealed the presence of different VHs (figure 18a). One particular sequence VH^{H16} (clone PGC7) was highly represented (65%) and corresponded to a VH of the VH4-DP79 family with the JH3b and D3-22 segments, containing a 16 aa long H3 (figure 18c, red sequence). A sequence similar to VH^{H16} was found in clone PGB9 carrying only three mutations indicated in blue in figure 18. Two different sequences were also found, one with a 6 aa long H3 (clone PGB7) and the second with another 16 aa long H3 (PGC10). The VH family, D and JH segments to which they belong are indicated between parenthesis. Curiously, the clone with the shortest H3 was badly expressed (figure 20), although it was positive in ELISA (figure 18b).

a)





b)



c)

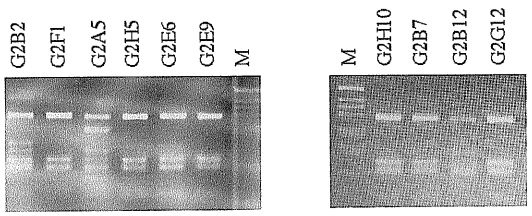
```
PGB9   QVQLQESGPGLVKPSQTLSLTCTVSGGW-ISSGGYYWSWIRQPPGKGLEWIGSIYYSG--STYYNPSLKSRVTISVDTSKNQFSLKLSSMTAADTAVYY
PGB7   QVQLQQSGPGLVKPSQILSLTCAISGDS-VSSNSAIWNWIRQSPSRGLEWLGRTYYRSKWYYDYALSVKNRITLNPDTSKNQFSLHLNSVTPEDTAVYY
PGC7   QVQLQESGPGLVKPSQTLSLTCTVSGGS-ISSGGYYWSWIRQPPGKGLERIGSIYYSG--STYYNPSLKSRVTISVDTSKNQFSLKLSSMTAADTAVYY
PGC10  QVQLVQSGAEVKKPGESLKISCKGSGYS---FTSYWIGWVRQMPGKGLEWMGIIYPGDS-DTRYSPSFQGQVTISADKSISTACLQWGSLKASDTAMYY

PGB9   CAG---IHPSGAYYDSSDAFDIWGQGTTVTVSS (VH4-DP79/D3-22/JH3b/H3 lenght=16)
PGB7   CAT-------GAWL------QQWGQGTLVTVSS (VH6-DP74/D7-27/JH5b/H3 lenght=6)
PGC7   CAR---IHPSGAYYDSSDAFDIWGQGTTVTVSS (VH4-DP79/D3-22/JH3b/H3 lenght=16)
PGC10  CARHGGGAAAGGYY---YGMDVWGQGTTVTVSS (VH5-DP73/D6-13/JH6b/H3 lenght=16)
```

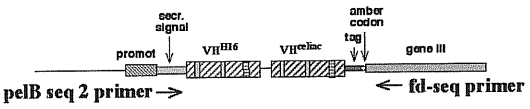
Figure 18 a) Fingerprinting of scVHD [VH^{PGG}-VH^{celiac}] selected clones. The scVHDs were amplified with pelB seq 2 and fd seq primers, further digested with BstN1 restriction enzyme and analysed in a 2% agarose gel. At the bottom a schematic representation of the amplified sample is shown. b) ELISA test of selected scVHD [VH^{PGG}-VH ^{celiac}] (celiac stands for celiac disease library) c) Alignment of position two VH sequences of selected clones. VH, D and JH genes used and length of H3s are indicated in parenthesis.

In order to investigate if the VH^{PGG} (3aa long H3) could be replaced by VHs with longer H3s, the VH^{PGG} was replaced with the selected VH^{H16}. The resulting library contained VH^{H16} in position 1 and the same variability as the previous library in position 2 (figure19). This library was used in a single round of selection. By ELISA, 65% of the single clones showed anti-extract activity (figure 19b). Within these positive clones a large number were homodimers in which the same VH^{H16} or with only few mutations were represented in position 2 in 70% of the binders. The other VHs selected contained H3 of different lengths ranging from 10 to 16 residues (figure 19c). Remarkably, all positives clones contain a central GXY motif in H3 (green letters in figure 19c). However, if clone G2B12 is not considered a central motif XHXSGXY and $F^{100}D^{101}$ can be identified. It is also interesting the fact that clone G2B12 contained the same VH as clone PGC10, likely indicating that this sequence was the consequence of a selection process. Moreover, also in the negative clones a bias to a particular VH was found (clones NC8, NC9 and NC10).

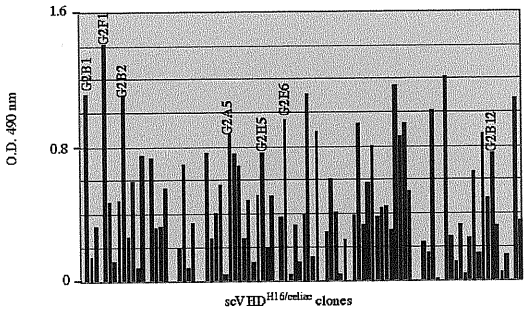
These results indicated that functional VHD could be obtained containing VHs with both H3 of average lengths (i.e. VH^{H16} paired with VHs containing H3 of 14aa, 15aa and 16aa long). A phage display analysis was carried out on several different positives and negatives clones. As shown in figure 21 the level of display between them varies.

a)





b)



c)

```
NB8     EVQLVETGGGLVKPGGSLRLSCAASGFTGLTFSDYYMNWIRQAPGKGLEWVSYITESGR-NTYYADSVKGRFTISRDNAKNSLYLQMNSLRADDTAVYY
NB9     EVQLLETGGGLVQPGGSLRLSCVASGFT---FSTYPMHWVRRAPRKGLEYVSAILGDGS-HPQYADSVKGRFTISRDNSQNTLYLHMGSLRPEDMGVYY
NC8     QVQLQQSGPGLVKPSQTLSLTCAISGDS-VSSNSAAWNWIRQSPSRGLEWLGRTYYRSKWYTDYALSVKSRIAVNPDTSKNQFSLQLNSVTPEDTALYY
NC9     QVQLQQTGPGLVKPSQTLSLTCAISGDS-VSSNSAAWNWIRQSPSRGLEWLGRTYYRSKWISDYALSVKSRITINPDTSKNQFSLQLNSVTPEDTALYY
NC10    QVQLQQTGPGLVKPSQTLSLTCAISGDS-VSSNSAAWNWIRQSPSRGLEWLGRTYYRSKWISDYALSVKSRITINPDTSKNQFSLQLNSVTPEDTALYY

G2A5    QVQLQESGAGLLQPSETLSLTCTVSGGS-VSSGSYYWSWIRQPPGKGLEWIGYIYYSG--STNYNPSLKSRVTISVDTSNNQFSLKLRSVTAADTAVYY
G2B1    QVQLQESGPGLVKPSQTLSLTCTVSGGW-ISSGGYYWSWIRQPPGKGLEWIGSIYYSG--STYYNPSLKSRVTISVDTSKNQFSLKLSSMTAADTAVYY
G2H5    QVQLQESGPGLVKPSQTLSLTCTVSGGW-ISSGGYYWSWIRQPPGKGLEWIGSIYYSG--STYYNPSLKSRVTISVDTSKNQFSLKLSSMTAADTAVYY
G2B2    QVQLVQSGAEVKKPGASVKVSCKASGYT---FTSYYMHWVRQAPGQGLEWMGIINPSGG-STSYAQKFQGRVTMTRDTSTSTVYMELSSLRSEDTAVYY
G2E6    QVQLQESGPGLVKPSQTLSLTCTVSGGW-ISSGGYYWSWIRQPPGKGLEWIGSIYYSG--STYYNPSLKSRVTISVDTSKNQFSLKLSSMTAADTAVYY
G2F1    QVQLQESGPGLVKPSQTLSLTCTVSGGS-ISSGGYYWSWIRQPPGKGLEWIGSIYYSG--STYYNPSLKSRVTISVDTSKNQFSLKLSSMTAADTAVYY
G2B12   QVELVQPGAEVKKPGESLKISCKGSGYS---FTSYWIGWVRQMPGKGLEWMGIIYPGDS-DTRYSPSFQGQVTISADKSISTACLQWGSLKASDTAMYY

VH^H16  QVQLQESGPGLVKPSQTLSLTCTVSGGS-ISSGGYYWSWIRQPPGKGLERIGSIYYSG--STYYNPSLKSRVTISVDTSKNQFSLKLSSMTAADTAVYY

NB8     CARVG-RHPHNYA------IDAWGQGTTVTVSS  (VH3-DP35/D1-07 or D4-04/JH6b/H3 lenght=12)
NB9     CARD---YDTGWS------LDYWGQGTLVTVSS  (VH3-DP61/D1-26/JH4b/H3 lenght=10)
NC8     CVR------DGGLG---LDAYDIWGHGTTVTVSS  (VH6-DP74/D1/JH3b/H3 lenght=11)
NC9     CVR------DGGLG---LDAYDIWGHGTTVTVSS  (VH6-DP74/D1/JH3b/H3 lenght=11)
NC10    CVR------DGGLG---LDAYDIWGHGTTVTVSS  (VH6-DP74/D1/JH3b/H3 lenght=11)

G2A5    CARAP-YHASGSYP----YFDSWGQGTLVTVSS  (VH4-DP66/D3-10/JH4b/H3 lenght=14)
G2B1    CAG---IHPSGAYYDSSDAFDIWGQGTTVTVSS  (VH4-DP79/D3-22/JH3b/H3 lenght=16)
G2H5    CAG---IHPSGAYYDSSDAFDIWGQGTMVTVSS  (VH4-DP79/D3-22/JH3b/H3 lenght=16)
G2B2    CASTP-THFSGSYP---APFDYWGQGTLVTVSS  (VH7-VHGL1.2/D3-10/JH4b/H3 lenght=15)
G2E6    CAR---IHPSGAYYDSSDAFDIWGQGTTVTVSS  (VH4-DP79/D3-22/JH3b/H3 lenght=16)
G2F1    CAR---IHPSGAYYDSSDAFDIWGQGTMVTVSS  (VH4-DP79/D3-22/JH3b/H3 lenght=16)
G2B12   CARHGGGAAAGGYY---YGMDVWGQGTTVTVSS  (VH5-DP73/D6-13/JH6b/H3 lenght=16)

VH^H16  CAR---IHPSGAYYDSSDAFDIWGQGTTVTVSS  (VH4-DP79/D3-22/JH3b/H3 lenght=16)
```

Figure 19 a) Fingerprinting of scVHD$^{H16/celiac}$ [VH^{H16}-VH^{celiac}] clones (celiac= celiac disease library). Amplified scVHDs were digested with BstN1 restriction enzyme. At the bottom a schematic representation of the amplified sample is shown. b) ELISA test of individual scVHD$^{H16/celiac}$ clones. c) Alignment of sequences from different clones (position 2). VH, D and JH genes used and length of H3s are indicated in parenthesis.

Most of the clones that showed binding activity were well displayed, although some of them (such as clone B2) showed lower ability to be displayed. The ELISA value for clone G2B2 however was comparable to the others, suggesting a higher affinity and/or efficient folding.

Interestingly, negative clones also displayed well, thus indicating that in those cases the VH^{H16}/VH combination was not permissive for antigen binding and suggesting that both VH segments contribute to the specificity of the binders. Practically in all cases the recombinant VHD was recovered from supernatant of bacteria cultures as a 30 kDa protein (figure 20b).

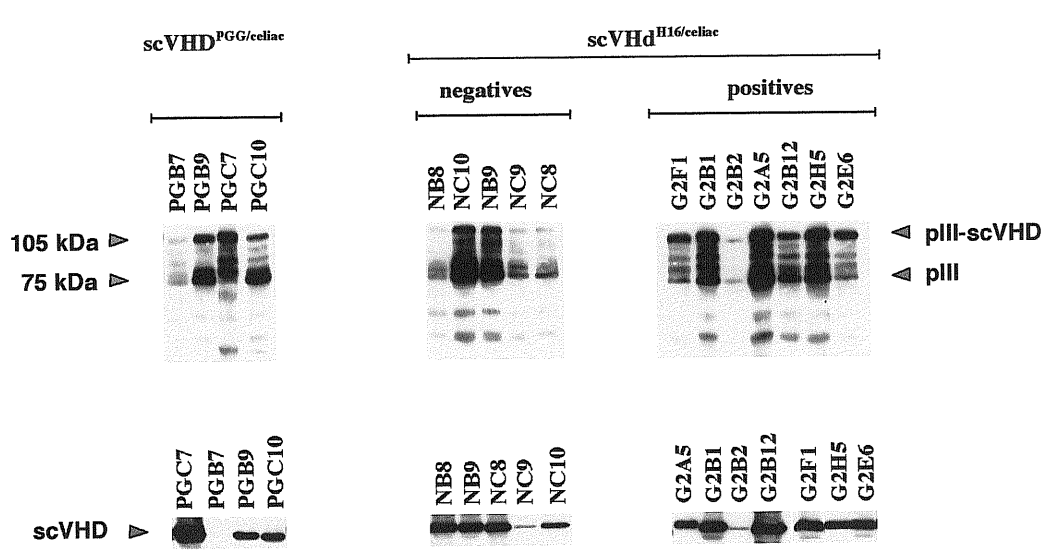


Figure 20 Western blot of phage displayed scVHD and soluble scVHD detected with anti-SV5 tag mAb. a) Phage-display analysis of positive and negative clones was performed on selected $scVHD^{PGG/celiac}$ and $VH^{H16/celiac}$. b) Soluble scVHD was analysed on supernatants of DH5αF' bacteria culture induced with IPTG.

VHD library on pDAN5

In an attempt to further develop the use of VHD we decided to construct a fully naïve scVHD library with random sequences in both position 1 and position 2. For this reason we used vector pDAN5. As previously mentioned, this vector has the enormous advantage of allowing the formation of very large diversity libraries through in vivo recombination in Cre^+ bacteria. The usefulness of such vector had been well demonstrated by Sblattero et al (89), and therefore we decided to construct our VH library in pDAN5. This was done by independent RT/PCR of peripheral blood lymphocytes RNA using primers appropriate for cloning in position 1 and 2. After fingerprinting analysis with BstN1, the quality of the library appeared acceptable with >90% of clones containing a scVHD. After sequencing analysis a bias on the presence of VH and JH families with preferential use of VH4 and VH6 with JH4 and JH6. Table 1 shows the relative frequency of VH and JH families represented in the library.

VH family	Position 1 (% of presence)	Position 2 (% of presence)	JH family	Position 1 (% of presence)	Position 2 (% of presence)
VH1	5	14	JH1	3	3
VH2	17	8	JH2	15	0
VH3	17	13	JH3	15	13
VH4	39	19	JH4	23	31
VH5	5	8	JH5	9	12
VH6	17	38	JH6	35	41

TABLE 1 Representation of the different VH and JH families derived from the recombined library. Ninety clones were analysed.

The diversity of the original primary library (before recombination) was 10^7. This number was greatly increased ($>10^4$, see ref. 89) after recombination. Analysis of our VHDs by superinfecting with several different phages revealed efficient mixing of VHs from position 2.

This recombined library was used to select binders of different antigens: lysozyme, rotavirus NSP5, GST, GST-HIV/TAT, human IgG and thioredoxin. All these antigens were used in several different conditions. Unfortunately, the results were always disappointing.

In many cases we did not obtain any binder. In others, strong binders were selected that upon further analysis demonstrated to be truncated fragments of only one VH without a complete V region sequence.

These results heavily contrasted with the selections performed on the wheat extract. The main difference between the two libraries were:

a) The vector: pDAN3 in the original, pDAN5 in the present

b) The variability of VH in position 1: VH^{PGG} in the original, VH library in the present.

The difference in the composition of position 1 does not seem to be relevant since it was also shown that either VH^{PGG} or VH^{H16} were able to sustain pairing with other VH that

showed binding activity. Therefore, the main candidate to explain the different behaviour of both libraries was the vector which differed only in the sequence of the linker that in pDAN5 contains the loxP511 recombination site (figure 17).

We then decided to check the ability of constructs based on the pDAN5 linker sequence to produce acceptable levels of display. Figure 21 shows such analysis of randomly picked clones from the pDAN5 library. It is clear that none of them were able to display the VHD-pIII fusion protein. As a control a similar construct on pDAN3 is shown on the left.

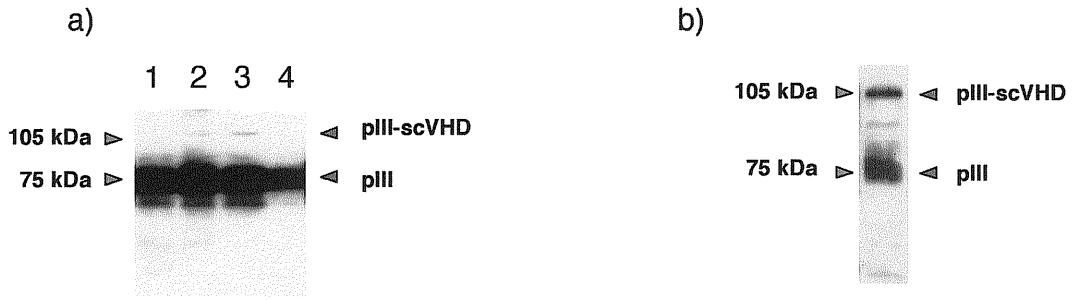


Figure 21 a) Phage display of scVHD-pIII using pDAN5 vector. Phage preparation was loaded into a 10% PAGE-SDS gel and the western blotting performed using an anti-SV5 tag mAb. b) Positive control of a displayed scVHD-pIII using phagemid pDAN3.

These results strongly suggested that the linker was responsible for the failure to select binders. This was further demonstrated by replacement of the linker. Figure 22 shows that the display of 3 different clones expressed with either linker loxP511 (pDAN5) or linker 220 (pDAN3). While the pDAN5 linker did not sustain any display of the fusion protein, all three VHD were efficiently displayed with the pDAN3 linker.

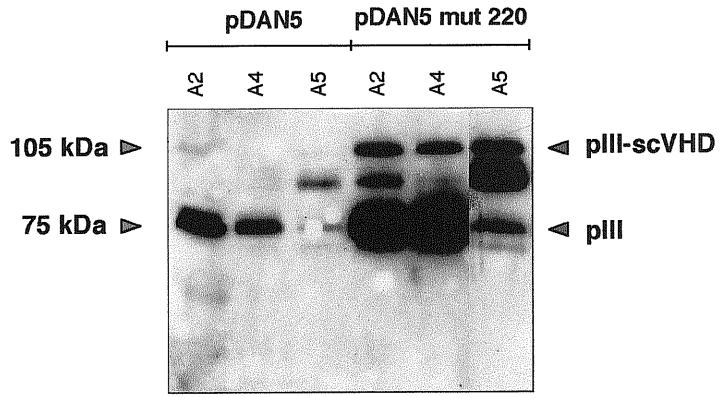


Figure 22 Phage display of scVHD-pIII using different linkers. Three different clones of the pDAN5 library were tested for display with the original (pDAN5) and the modified (pDAN5 mut 220). Western blot of phage display protein was visualised using the anti-SV5 mAb.

This result prompted us to reconstruct a small fully naïve library in pDAN3. Not surprisingly, analysis of different VHD (as judged by the BstN1 fingerprinting) showed that most of them were able to fully display VHD-pIII protein. In addition, all clones with capacity to display also produced significant amounts of soluble VHD protein (figure23b bottom).

All together these results demonstrated that a suitable linker (such as 220 linker) is required for the functional display of VHD. In addition, they also showed that the recombination strategy in the library developed with the vector pDAN5 is completely deleterious to use with VHD.

From this experience we concluded on the need of libraries based on a suitable vector, such as pDAN3, for proper selection of VHD binders.

It is noteworthy to mention that while this work was in progress, alternative strategies were been investigated in the lab (by Hulin Jin) that exclusively make use of homo-VHD. Binders against several different antigens selected by Hulin Jin demonstrated that VHD are fully functional. Some of these binders were used in this thesis for the studies presented in the next section.

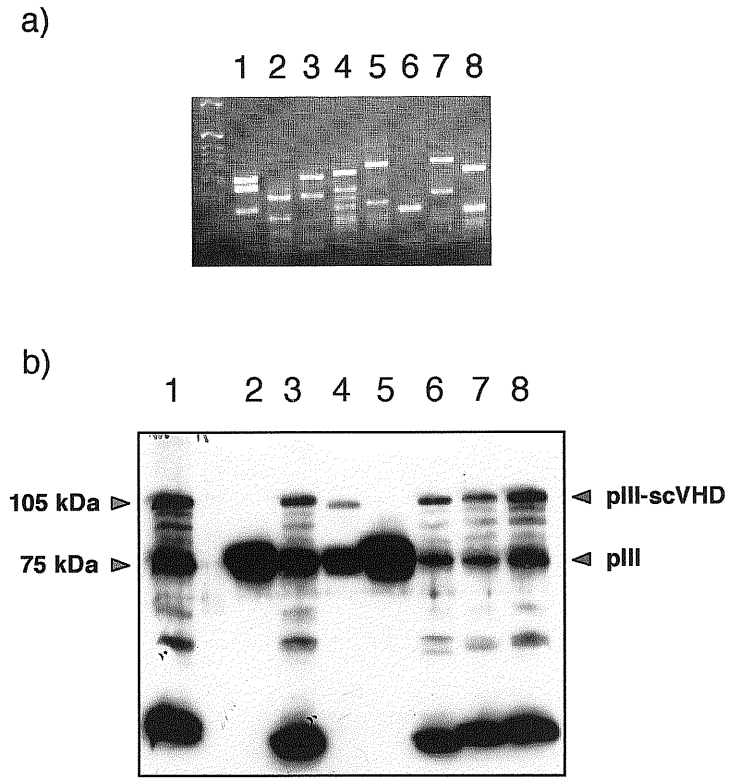


Figure 23 Analysis of scVHD using the 220 linker. a) Fingerprint of complete VH-VH segment digested with BstN1. b) Westernblot analysis of displayed scVHD-pIII protein reacted with mAb against the SV5 tag.

Structural analysis of functional VHD

With the aim of characterising the dimeric state of functional VHDs coming from phage display selections we expressed them in sp2/0 cells. Three different VHs were studied: VH^{H16} that binds wheat extracts, both as homo and heterodimers, VH^{GST} selected as a homo scVHD on GST (Hulin Jin) and VH^{DNA}, a homo scVHD with DNA binding activity (Hulin Jin). All three were expressed as double chain (dc) VHD fused to SV5 tag with and without the C-terminal cysteine (homo-dcVHD/SV5-cys and homo-dcVHD/SV5, respectively).

As shown in figure 24, each one of them was well expressed and secreted as covalently linked dimers in the format with terminal cysteine. This was confirmed by DSP crosslinking of the material secreted in the homo-dcVHD/SV5 format (without cysteine). In all three cases, dimers were clearly detected under non-reducing conditions. However, the ratios between monomer, dimers and oligomers (>60kDa) differed between the three constructs. This analysis was repeated several times, always obtaining dimers and monomers, eventhough with some variability.

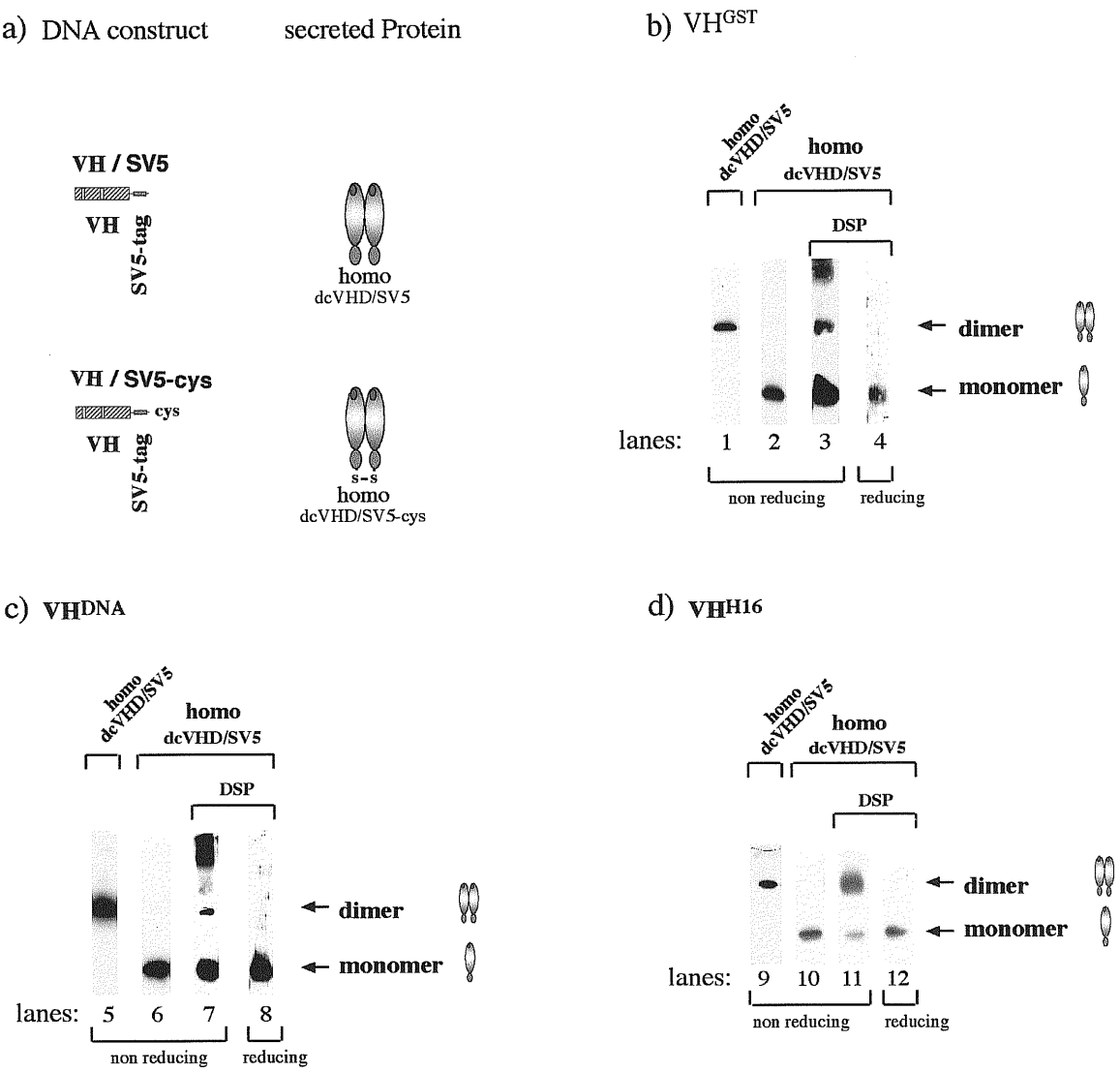


Figure 24 a) Scheme of the formats used to express the three different functional homo-VHDs (VH^{GST}, VH^{DNA} and VH^{H16}). Only proteins expressed in the homo-dcVHD/SV5 format were used for DSP crosslinking. Western blots of homo-dcVHD corresponding to VH^{GST} (b), VH^{DNA} (c) and VH^{H16} (d) are shown. Lanes 1, 5 and 9 correspond to expression of in dcVHD/SV5-cys format. Samples were crosslinked or not crosslinked with DSP and analysed under reducing and non-reducing condition as indicated.

In conclusion, these data confirmed the previous observation obtained with VH^{PGG} and suggested that VHs interact with each other forming dimers oriented as in the VL-VH scaffold.

All the VHDs produced in mammalian cells were tested in ELISA for their binding activity. Figure 25 shows that all of them retained the specificity observed in the phage displayed scVHD format.

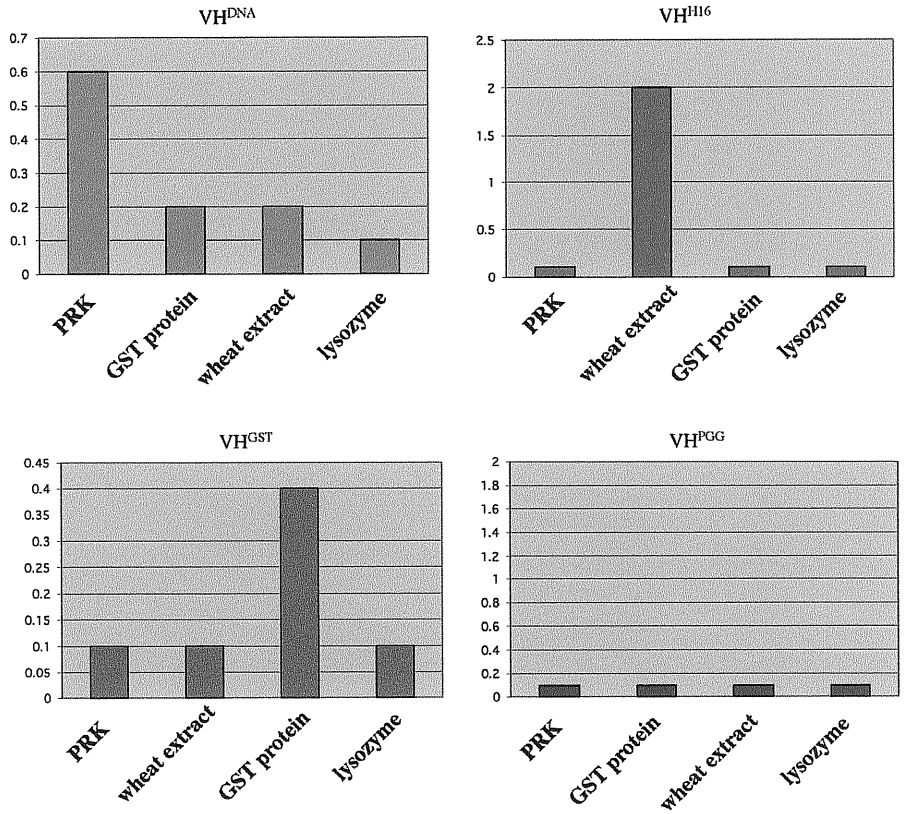


Figure 25. Specificity analysis of homo-VHDs (as homo-dcVHD/SV5-cys) by ELISA. PRK stands for a specific DNA sequence, used for VHD^{DNA} selection, and GST for Glutathione S-transferase.

VH^{H16}, with wheat extract specificity, was further analysed because of its relatively large H3 (16 aminoacids) and strong reactivity and specificity.

In addition to the previously described VHD/SV5-cys format, VH^{H16} was also expressed as both homo and hetero-scVHD as well as a complete Ig molecule. The two scVHD formats ($scVHD^{PGG/H16}$ and $scVHD^{H16/H16}$) were efficiently expressed and secreted from sp2/0 transfected cells (figure 26).

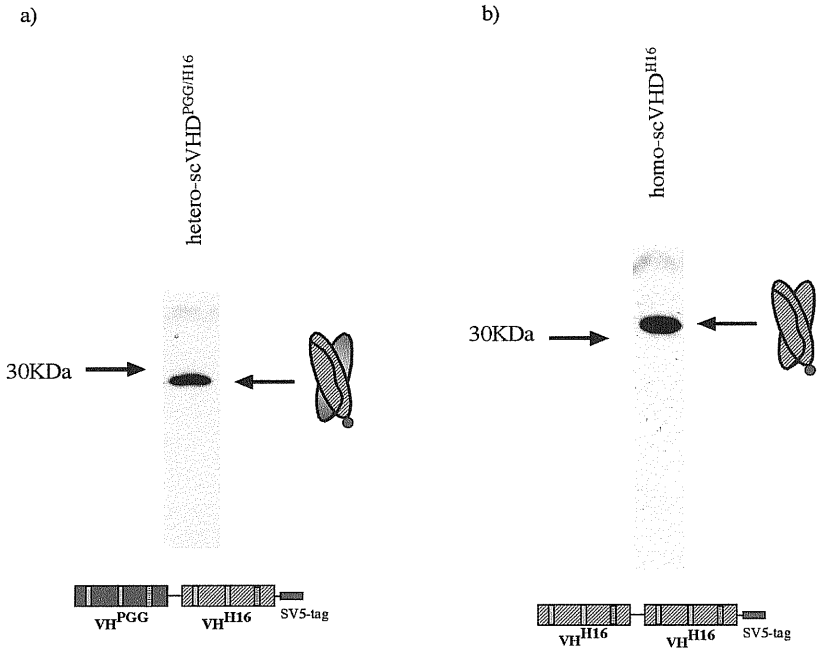


Figure 26 Western blot analysis of sp2/0 supernatant transfected with the hetero-$scVHD^{PGG/H16}$ and homo- $scVHD^{H16}$). Supernatant was loaded into a 15% PAGE-SDS gel and electrotransfer into a PVDF membrane. Visualisation of the secreted protein was performed through the SV5 tag using mAb against the SV5-tag. Schematic representation of the constructs and proteins are also shown.

To construct the full Ig molecule format, VH^{H16} was cloned as complete human γ1 heavy chain and human κ light chain. The light chain consisted of the VH^{H16}, instead of the VL fused to the C region of the κ–chain. Since light chains can dimerise we constructed the L chain vector without the selection marker for neomycin. Thus, it can not produce a stable clone when transfected alone and cultured in the presence of the antibiotic. On the other hand, heavy chains can not be secreted in the absence of L chain due to the presence of CH1, which acts as a retention element through interaction with BiP in the endoplamic reticulum. Therefore the only viable clones are those transfected with the 2 plasmids that should express the complete Ig molecule.

Analysis of several clones by Western blotting under non-reducing conditions showed a single band of 160kDa (correspond to the whole IgG). Moreover, in reducing condition the big complex disappeared and a smaller band could be observed, which corresponded to the H chain (approx. 55 kDa) (figure 27).

Figure 28 shows that all VHD formats of VH^{H16} including the complete IgG were able to specifically bind the wheat extract antigen originally used for its selection. This results was particularly important because it shows that the completely reconstituted molecule retains binding specificity and opens the possibility of using VHD binders in the context of a whole Ig exploiting effector functions of different isotypes.

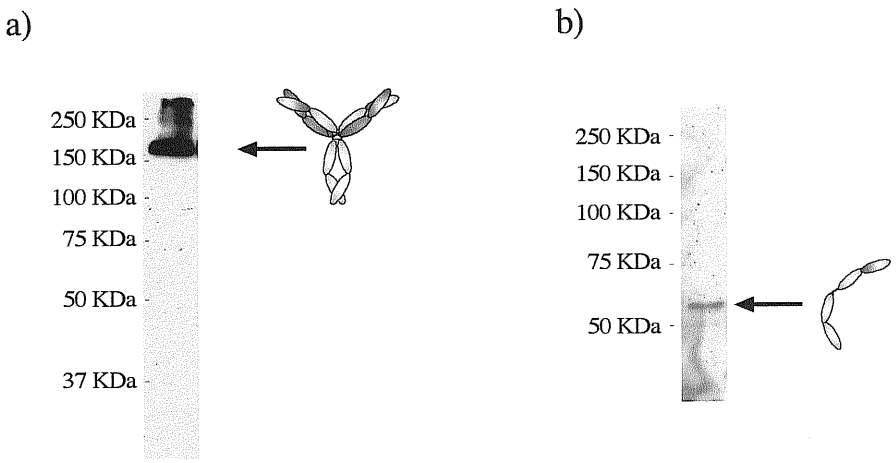


Figure 27 Western blot analysis of a VHD of VH^{H16} in the whole Ig format (IgG1-VHD^{H16}. Supernatant of transfected sp2/0 cells were analysed on 10% PAGE-SDS gel under non-reducing a) and reducing b) conditions. Detection was performed using anti-human IgG-HRP conjugate.

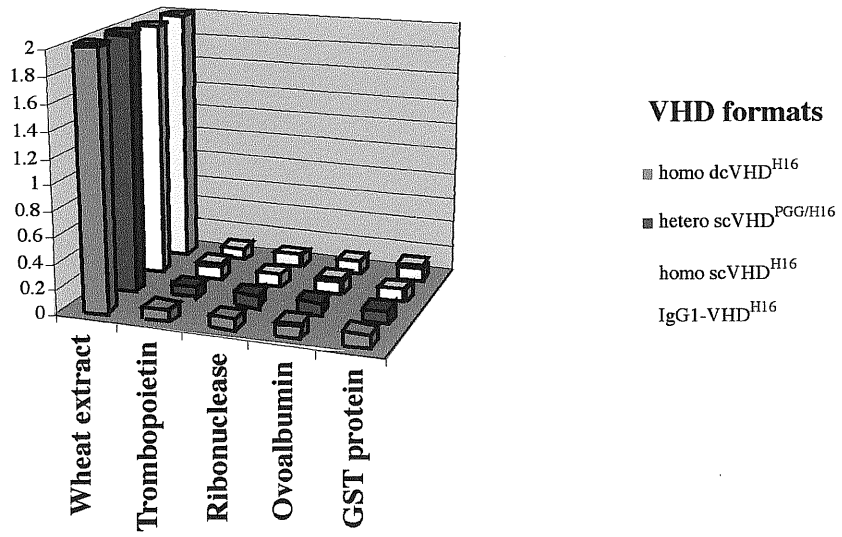


Figure 28: Specificity of the different formats used to express VHD containing VH^{H16}

CHAPTER 3

DISCUSSION

In this project a novel binding molecule that we called VHD was developed and characterised. It is based on two immunoglobulin heavy chain variable domains that form an antigen-binding site similar to the VL-VH scaffold. Eukaryotes and prokaryotes can secrete VHDs, maintaining their binding specificities. Thus, in the advent of the functional genomic era this thesis described a new tool of biotechnological interest to find alternative binders in order study new proteins.

VH domains should not be stable as monomers in an aqueous system, when normally folded, due to their hydrophobic β-bulges. Power et al. found by circular dichroism, a VH domain of a monoclonal antibody (NC41) preserving its β-sheet structure when expressed alone in *Escherichia coli*, suggesting that it should be quite stable even in the absence of VL (104). When trying to purify single expressed VH, as monomer, it was found that these domains tend to produce dimers under standard (lack of detergents) conditions (94), suggesting that the hydrophobic surfaces are preferentially neutralized by themselves in order to stabilize the molecule. After doing a simple computer model, no evident reason was found that could preclude VH dimerisation, through the neutralisation of the hydrophobic surface (around 1300 $Å^2$) that normally forms the interface between VL and VH (12, 13).

Mammalian cell secretion is accepted as a criteria of correctly folded and assembled immunoglobulin domains (99, 100). The results we obtained allowed us to conclude that VHs are correctly folded and assembled even in the absence of VL, preferentially as dimers or using the dimeric state as a basic unit to form large aggregates. So far, no evidences of VH interactions have been reported. We thus considered that these new immunoglobulin-like scaffolds based on the properties of VH-VH dimers (VHD) could result in a family of molecules of biotechnological interest.

In normal antibody molecules all the CDRs usually participate in antigen recognition. Moreover, Poljak and coworkers have shown that CDRs interact with the antigen through an induced fit rather than a 'lock and key' fit manner. The main forces stabilizing the complex arise from antigen-antibody hydrogen bonding, van der Waals interactions, enthalpy of hydration, and conformational stabilization rather than solvent entropy (hydrophobic) effects (105). However, diversity of antibodies is mainly given by the VH (58). In addition, studies analyzing the buried surface between CDRs and antigen have put in evidence that H3 was generally responsible for the major antigen-surface contact (13). The H3 region is much more variable in length and sequence than the other loops (5, 59) because it is formed by the rearrangement of the V-D-J genes during the immunoglobulin heavy chain gene assembly. Analysis of the relationship between the sequence and the three dimensional structure of antibody combining sites revealed that, except for H3, the other loops have a restricted number of main chain conformations or canonical structures, determined by its size and the presence of certain residues at key sites in the loop and in the framework (54, 55). These data suggest that, if CDRs in VHD are oriented as in the VL-VH scaffold, VH dimers might create more and different antigen-binding site conformations compared to VL/VH. Moreover, better affinities might be achieved if we consider H3 as the longest CDR and the one that generally interacts more with the antigen. However, this aspect needs to be demonstrated.

Therefore, our first goal was to study the feasibility of VH association; both in homo and heterodimer manner, in order to perform further functional analysis.

To study the problem of VH-VH interactions we transfected human VH domains alone in mammalian cells. We observed that in all cases expression was always achieved, although at different levels, confirming that the VH folding was correct. In order to have an indication about the relative orientation that VHs can adopt a cysteine residue was placed at the carboxy-terminal part of the construct. Mammalian cells consistently

secreted a protein with disulphide-bridge properly formed between two VHs (practically 100% of the material), supporting the idea of the proximity of the two carboxi-terminal ends of the monomers, like in the VL-VH pairing. In addition, no monomers were observed indicating that VHs rarely exist without interacting with each other, and most likely forming dimers. On the contrary, if this association would be at random oxidations of cysteines would have created very large multimers, covalently bound. There is another interaction possibility to keep the cysteines close to each other. This involves the hydrophobic patches that interact with the CH1 in the complete immunoglobulin. This region however has a variable location and a smaller extension of the surface (around 700 $Å^2$) (14). Most likely that these patches are related to the VH aggregation observed when trying the purification of VHD. Moreover, mammalian cells were able to secrete a recombinant homo VHD in the complete human IgG1 format. This construct had as variable domains only a VH (VH^{H16}) that was also replacing the VL, further suggesting that the interaction between VHs occur through their VL-neutralisation hydrophobic surfaces since CH1 and CL should hide the other hydrophobic patches in both V regions. Interestingly, this molecule retained the binding specificity of the original $scVHD^{H16}$.

In order to know if VHs can interact stably with each other, without any covalent link, we produce homodimers in both eukaryotic and prokaryotic systems. After trying to purify the molecule for further proving its association through a gel filtration column we found an aggregation state that could be due to the CH1-interface-hydrophobic patches. This was not observed with the complete IgG-VH^{H16}, which gave a single peak of approx. 150 kDa with no evidence of aggregates (data not shown). In addition, it is also known that VL-VH pairs without covalent links are unstable (103).

To solve the problem we decided to look for VH-VH interactions directly from mammalian cell culture supernatants by direct crosslinking of the molecules without any purification step. The idea of performing this analysis was to observe the oligomerisation

state that VHs could adopt when secreted. If we consider the two hydrophobic surfaces present in one domain it is possible to predict the formation of tetramers or higher oligomerisation states. However, dimer was also expected as the basic unit of any higher order. The results indicated that dimeric forms are the main oligomerisation-state, although the higher associated forms were sometimes observed. With this technique, monomers were present as the main product and which somehow contrast with the high efficacy of dimer formation of the molecules expressed with a C-terminal cysteine. We know that in part, the presence of the monomers after in vitro crosslinking reflects an uncontrolled chemical reaction, probably due to variability in the protein: crosslinker ratio. The setup of the reaction is difficult to perform. In titration experiments we observed that, for a fixed amount of protein, the amount of dimer recovered was highly dependent on the amount of reagent used. An excess of crosslinker would yield a lower recovery of dimer. In conclusion, our crosslinking assays confirmed that dimers are the preferred oligomerisation-state suggesting an assembly similar to VL-VH. Nevertheless, to fully understand how VHs interact and how large the interface surface is, structural analysis by X-ray diffraction of the molecule is necessary. This indeed is currently being performed in the lab with one homo VHD that shows good solubility properties. One possible way to avoid the tendency of this molecule to aggregate will be its expression as Fab (instead of scVHD), where the hydrophobic patches at the bottom of the β-barrel structure of the VH/VH pair would be neutralised by CH1 and CL.

On the light of our findings, it is interesting to discuss the results obtained in the natural assembling context. Each immunoglobulin domain behaves as an autonomous folding unit (106, 107, 108) but folding of CH1 is stopped by the endoplasmic reticulum chaperone protein BiP (37). Hendershot et al (36) demonstrated that the CH1 domain of heavy chains is necessary for their association with BiP. In the absence of association with BiP as in the case of H chain mutants lacking CH1 domain, free heavy chains and

assembling intermediates can be secreted, whereas complete heavy chains are retained in the absence of L chain. Kallof et al. (109) showed that CH1 is important in controlling the in vivo formation of a homogenous population of functional IgG antibody molecules. It is important to mention that the CH1 domain can not dimerise and therefore this also guarantees "in vivo" the impossibility of VH dimerisation. When the light chain is expressed, CL can efficiently displace BiP, allowing CH1 to complete its folding and stabilizing the formation of VL/VH pairs, and consequently creating the antigen-recognition-site. This mechanism may prevent the formation of VHDs in normal B-cells.

It is possible that in all these previous studies the dimerisation of VHs was taking place without anybody realising. In only one case Kallof et al. (109) suggested a possible and non-desirable VH homodimer formation in a secreted CH1-deleted IgG, which was ill assembled (lack of L chain) and non-functional.

Our results demonstrate that VHs can interact with each other, although the precise way in which they perform the association should be further addressed with more complex structural analysis. These may also reveal whether particular sequence/conformational characteristics are required for proper dimerisation.

In order to test the functionality of this new molecule as binders, a phage display methodology was used to perform selections against antigens. Our first attempt to get binders was to use naturally immunised patients (celiac disease) to construct a library against wheat extract. Initially we used VH^{PGG} (with a 3 aminoacid long H3) as a partner in the VHD in order to reduce possible incompatibilities of H3 lengths. Selecting VHs that were able to bind the wheat extract we wanted to simulate an affinity-maturation-like process. After we obtained few different binders, a VH (VH^{H16}) identified in one VHD with an apparent higher affinity was chosen as fixed partner in the position of VH^{PGG}. The homodimer VH^{H16} was apparently the best binder obtained. Nevertheless, binders were found containing different VHs, suggesting that VH^{H16} may be the main contributor for

binding. Interestingly, negatives VHD were also found which indicate that not all associations are permissive to retain antigen specificity. One possibility to explain these results is that VH^{H16} is able to interact only with some VHs, with the binding activity due only to the VH^{H16} domain. However, in mammalian cells the homo VHD^{H16} is expressed mainly as dimer and retains binding activity. Due to aggregation problems we could not study the comparative affinities between the homo and heterodimer VHD^{H16}. We are trying to get more information on this aspect using the whole immunoglobulin format. With this objective we will also use a VL-VH^{H16} anti wheat extract scFv that we selected from the original celiac patient library (89).

As previously mentioned, mammalian cells can not secrete heavy chains in the absence of light chains. We observed in one case a single VH that could not be expresses in the absence of its original VL partner. Indeed, VH^{6C6} that comes from an IgG monoclonal antibody and has a somatically mutated sequence, quite different to its germline, can not be expressed unless the appropriate VL^{6C6} partner is cotransfected (unpublished data). There are cases of reported VLs that can neither be secreted without their appropriate VH counterpart due to somatic mutations (110, 111). After realizing that VH^{PGG} has a very short CDR3 and that it was produced well by the cells, H3 grafting of the $H3^{PGG}$ to the VH^{6C6} was performed. The modified VH^{6C6} could not be expressed and secreted. We concluded from these results that some mutations are preventing the correct assembly of the VH^{6C6} dimer. Since VH^{6C6} was derived from an IgG after an immunization process it is possible that the somatic mutation process had introduced aminoacid replacements that favor the specific VL-VH pairing. Furthermore, since in some cases it has been shown that part of the H3 can contribute to VL-VH associations, mutations at that level selected to improve affinity to the antigen may have stabilized the selected antibody making a unique VL-VH pair. On the other hand, the sequence of VH^{H16} is very similar to the germline, as seen in other examples, suggesting that the somatic mutation might

complicate the search of functional VHDs. Thus, a germline condition or a very similar one is likely to be desirable to form VHDs.

In a second step, and in order to find more examples of this novel binding molecules we constructed a naïve library to further select against different antigens.

Phage display libraries have been introduced as a tool for making antibodies as an alternative to hybridoma technology. Advantageous appears the fact that directly human Fvs can be used bypassing immunization. However, the success of having antibodies with good affinities depends on the size and diversity of the library (84, 87, 79, 80, 81).

One limitation of the method is the reduced Fv diversity obtained in libraries of limited size. In vivo recombination methods have been used to overcome this problem (79, 89, 112, 113, 114). Library diversity is theoretically ruled by a fixed number of VL and VH genes. Since much more VHs than VLs exist, it is theoretically possible that our new hetero VHD could achieve higher diversity than current scFv (VL-VH). Nevertheless, there may still be limits to the potential diversity of VHD libraries, considering that as described for VLs (9), VHs might also have a preferential association to only a fraction of all VHs, thus reducing the combinatorial diversity.

Unfortunately our first attempt to construct a functional large diversity VHD library was unsuccessful because of the incompatibility of the lox511 linker in the scVHD context. Since this linker works well with scFvs (89) we can conclude that additional restriction exist between VHs for using it in the single chain VHD format.

Other interesting examples of homo VHD binders were selected in our lab using phage display methodology on a specially designed vector (Jin H, unpublished results). Binding against a specific DNA sequence, GST protein and lysozyme were achieved. In all these cases the specific binding capacity of the VHD selected was proved, even though, a precise mapping of the antigen-binding site is still in progress. In one case, a VH of a

homo scVHD was replaced with different VHs. The resulting mutants failed to bind the antigen, suggesting that both VHs contribute to the binding specificity.

One advantage of our novel VHD binders resides in the possibility of creating symmetric binding surfaces that might be useful in recognising symmetric epitopes such as palyndromic sequences on DNA. In contrast, one important limitation could be the source of VHs since hypermutated sequences might not be able to interact with each other.

A few attempts to use a single VH domain as binders have been described. In all cases the idea was to use them as monomers eliminating the VL and reducing the size of the scFv (8, 115, 116).

An interesting case of natural antibodies without VL is that of camels. One particular isotype of camels do not contain L-chain due to the lack of CH1 in H-chain which is accompanied by substitutions of aminoacids located in the interface region of the VH that interacts with he VL. The consequence of this is that the VHs are soluble avoiding aggregations.

This natural selection has been exploited with biotechnological interest by specifically mutating residues in the VH interface ($G^{44}E$, $L^{45}R$, $W^{47}G$) allowing the monomeric expression in soluble form of defined VHs in bacteria (8, 115).

Similarly, Reiter et al proposed the use of a VH from a monoclonal antibody that has shown to be stable as a monomer and thus useful as a scaffold to graft different H3s (116).

In contrast to these examples our VHD exploit the ability to dimerise spontaneously offering the possibility of obtaining binders based on two efficient domains, which could also create symmetric binding surfaces as in the case of homo VHDs.

In conclusion, VHs are suitable candidates to work with in order to find new sources of binders. They can dimerise similarly to VL-VH couples, most likely allowing all CDRs to form the antigen-binding surface. The advantages in relation to VL-VH pairs are based

on the different non-natural surfaces that can be formed and the large diversity with undefined properties. In addition, the fact that VHs generally contribute more than VLs in the interaction with the antigen suggests that VHDs might achieve higher affinities than VL-VH pairs. The potentiality of these molecules not only involves different and non-natural binding sites, but it also comprises a source of symmetric binder's surfaces that could target symmetric epitopes. The posibility to express this molecule with different formats in several systems makes also VHDs a biotechnological tool with a high potentiality.

CHAPTER 4

ABBREVIATIONS

aa	Amino Acids
Amp	ampicillin
bp	base pair
cDNA	complementarity DNA
cfu	colony forming units
DMSO	Dimethylsulfoxide
EDTA	ethylenediamine tetraacetic acid
IPTG	isopropylthio-_-D-galactoside
isoAA	isoamyl alcohol
Kan	Kanamycin
LB	Luria-Bertani
PAGE	polyacrilamide gel electrophoresis
PBS	Phosphate buffer saline
PEG	polyethylene glycol
pfu	plaque forming units
rcf	relative centrifugal field
SDS	Sodium dodecyl sulfate
Ab	antibody
APC	antigen presenting cells
C	constant
CDR	complementary determining region
D	diversity
DC	dendritic cells
dcVHD	double chain VH dimer
DP	double positive
ER	endoplasmic reticulum
FDC	follicular dendritic cell
FW	framework
GC	germinal center
H	heavy
IFN	interferon
Ig	immunoglobulin
IL	interleukin
J	joining
L	light
LPS	lipopolysaccaride

ScFv	single chain fv
scVHD	single chain VH dimer
SIP	small immune protein
TCR	T-cell receptor
TD	T-cell dependent
TGF	tumor growth factor
Th	T-helper
V	variable

Materials and methods

Cell lines and media

Mouse myeloma Sp2/0 is available from American Type Culture Collection (Rockville, MD, USA). Sp2/0 was grown and maintained in RPMI-1640 medium supplemented with 10% fetal calf serum

Bacterial strains

DH5alfaF' (Gibco BRL): F'/endA1 hsdR17 (rK- mK+) supE44 thi-1 recA1 gyrA (Nalr) relA1 D (lacZYA-argF)U169 deoR (F80dlac delta(lacZ)M15) was used for standard phage propagation.

BS1365 BS591 F' kan (BS591: recA1 endA1 gyrA96 thi-1 D lacU169 supE44 hsdR17 [lamda1mm434 nin5 X1-cre].

HB2151 (K12, ara Δ(lac-pro), thi/F' proA+B+, lackIqZΔM15) a non-suppressor strain was used to make soluble scFv.

Modeling

VHs dimers were simulated using the Swiss-PDB viewer programme, available in the Web site "http://www.expasy.ch/spdbv/text/getmac.htm", and the following immunoglobulin structures (PDB entry): 1ADQ (homo sapiens), 1VFB(mus musculus), 1IGM (homo sapiens).

First, VL domains from three different VL-VH couples were replaced by the same VH of the pair used, forming homodimers. For simulating the heterodimer two different VHs were combined, keeping the 1IGM structure fixed and replacing its VL with the 1ADQ VH. Since both tertiary structures are very similar the replacement was done aligning the VH over the VL backbone.

The PGG homodimer was modelled using the Swiss-Model package (95, 96, 97) and the following templates (PDB entry): 1AD9H, 1AXSB, 1DHAH, 1GC1H and 1MLBB. The resulted VH (called PGG due to the aminoacids of the H3) structure was placed instead of VH and VL of a VH-VH dimer scaffold which H3s clash to each other (PDB entry: 1ADQ).

Eukaryotic constructs

SIPs:

The first epsilon CH4 construct (SIP1) was done by Martin Ivanovski (unpublished data). Shortly, pUT-SEC (100) was digested with Pst1 and Kpn1, and oligos SIP-linker 1 and 2 were cloned. After amplifying εCH4 with the primers 5'εCH4/Apa1 and εCH4/b, it was cloned between Apa1 and Kpn1 sites.

For constructing SIP2, oligos SIP-LINK XHO1 and 2 were annealed and clone between Spe1 and BspE1 cloning sites, after proper digestion of SIP1. The resulting vector contain the tetracycline resistance, an ApaL1 and Spe1 sites for cloning VLs and a Xho 1 and BspE1 sites for cloning VHs. The linker between VLs and VHs has the sequence GSTSGSGKPGSGEGSTKG and between VH and epsilon CH4 we introduce a minilinker (GGSG) (101).

VH^{6C6} and VL were cloned using the primer 215-H5V/6C6BspE1 and 215-L5V/3LL, respectively, into SIP2.

Homodimers constructs:

PGG-SIP2 was constructed amplifying VH^{PGG} with the primers PGG-ApaL1 and PGG-Sal1 and cloning the digested fragment into SIP2 between ApaL1 and BspE1. To eliminate the εCH4, two oligos, PGGSV5 B and PGGSV5 f were annealed and further

cloned in the vector PGG-SIP2, digested with BspE1 and Kpn1, to generate VH^{PGG}-SV5-Cys. In parallel, the same construct but without cysteine at the c-terminal was done annealing to oligos, SIP-HIS A and SIP-HIS B. After completing the fragment with klenow enzyme it was cut with BspE1 and Kpn1 to be cloned in PGG-SIP2. The rest of the homodimers were done following the latter approach but using specific primers for each sequence. The only construct that was also done as VH^{PGG}-SV5-Cys was VH^{H16} always cloning with ApaL1 and BspE1.

ScFvh

To clone the single chain format of VH^{PGG}-VH^{H16}, pUT-SV5 was done annealing PUTSV5 BACK and FOR and cloning the fragment in SIP2, previously digested with Spe1 and Kpn1. From pDAN3 (see below) the scFvh was amplify with PGG-ApaL1 and fdseq (provided by Sblattero and Bradbury), digest with ApaL1 and Nhe1 and clone in the vector using ApaL1 and Spe1 sites.

Eukaryotic expression vector

After checking the sequences, all the final constructs were excised with HindIII/KpnI and cloned in the eukaryotic expression vector pcDNA3 (Invitrogen, Leek, The Netherlands)

under the control of the cytomegalovirus promoter.

VH^{H16} whole immunoglobulin

The whole immunoglobulin format of VH^{H16} was done using a strategy already developed in our lab. First, VH^{H16} was amplify using H16-ApaL1 and H16 C Gamma Xho primers to be cloned in pUT-SEC (101), which contains the secretion signal, using the ApaL1 and Xho1 cloning sites, for further transfer it in pUC-IgG1-Xho, digesting with HindIII and Xho1. The whole construct consists in a VH, VH^{H16}, and the whole heavy chain of the IgG1, but for expressing it in eukaryotes must be passed to pCDNA3. For constructing the light chain, VH^{H16} was amplified with H16-Apal1 and H16 C kappa Ban/Bs. The resulting fragment was directly cloned in the expression vector pCMV-2/Delta-Nhe, which contains the C kappa for completing the whole light chain, and a cytomegalovirus promoter but not the selection marker as pcDNA3.

Table of primers used

constructs	Primers (5'-3')
SIP1	SIP-linker 1: GTAGAGCTCCGGAGGCTCTGGGGGCCCGCGGGTAC
	SIP-linker 2: CCGCGGGCCCCCAGAGCCTCCGGAGCTCTACTGCA
	5'epsilon CH4/ Apa1: GATAGCTACAGGGCCCCGTGCTGCCCCGGAA
	epsilon CH4/b: GAGGCTCTGGGGGCCCGCGTGCT
SIP/2	SIP-LINK XHO1: CTAGTGGTAGCGGCAAACCAGGTTCCGGCGAAGGCTCGAGCAAAGGTT
	SIP-LINK-XHO2: CCGGAACCTTTGCTCGAGCCTTCGCCGGAACCTGGTTTGCCGCTACCA
VL^{6C6}-SIP/2	215-L5V: GTGTGCACTCGGATATTGTGATG
	3LL:CAGGACTAGTGCTGCCTTTTATCTCCAGCTTGG
VH^{6C6}-SIP/2	215-H5V: GTGTGCACTCTGAGGTGCAGCTG

	6C6 Bspe1: CGAATCCGGAAGAGACAG
PGG-SIP2	PGG-ApaL1: ACAGGTGTGCACTCGCAGGTGCAGCTGGTGCAGTCTGG
	PGG-Sal1: GAAGTTATGGTCGACCCTCCGGAGGAGACGGTGACCAGGGTT
VH^{PGG}-SV5-Cys	PGGSV5 B: CCGGAGGCAAACCAATCCCAAACCCACTGCTGGGCCTGGATGGTGGCTGC TAGGTAC
	PGGSV5 f: CTAGCAGCCACCATCCAGGCCCAGCAGTGGGTTTGGGATTGGTTTGCCT
VH^{PGG}-SV5-His6	SIP-HIS A: AGTTCCGGAGGCAAACCAATCCCAAACCCACTGCTGGGCCTG
	SIP-HIS B: AGTGGTACCTAATGGTGATGGTGATGGTGAGTACTATCCAGGCCCAGCAG
VH^{H16} SV5 with and without Cys	H16-ApaL1: ACAGGTGTGCACTCGCAGGTGCAGCTGCAGGAGTCGG
	H16-Sal1: GAAGTTATGGTCGACCCTCCGGAGGAGACGGTGACCGTGGTC
VH^{DNA}	DNA818ApaL1: ACAGGTGTGCACTCGCAGGTCACCTTGAAGGAGTCTGG
	DNA818BspE1: same as PGG-Sal1
VH^{GST}	GST58ApaL1: ACAGGTGTGCACTCGGAGGTGCAGCTGGTGGAGACTGG
	GST58BspE1: GAAGTTATGGTCGACCCTCCGGAGGAGACGGTGACCATTGTCC
PUT-SV5	PUTSV5 BACK: CTAGTGGCAAACCAATCCCAAACCCACTGCTGGGCCTGGATGGTGGCTAGGTAC
	PUTSV5 FOR: CTAGCCACCATCCAGGCCCAGCAGTGGGTTTGGGATTGGTTTGCCA
Ig VH^{H16}	H16 C Gamma Xho: CTGACTCGAGACGGTGACCGTGGTCC
	H16 C kappa Ban/Bs: CTCTSGTACSTGAGGAGACGGTGACCGTG

Cell transfection

About 2-3 x 10^6 cells were resuspended in 0.5 ml cold PBS (10.1 mM Na_2HPO_4, 1.8 mM KH_2PO_4, 137 mM NaCl, 3 mM KCl, pH 7.2) and put in a electroporation cuvette, which has an electrode gap of 0.4 cm. 10 µg Pvu I- linearized plasmid was added to the cells and electroporation was performed with a single pulse at 960 µF, 290 V, in a Bio-Rad Gene Pulser equipped with a capacitance extender. After electroporation cells were kept 5 min on ice, passed to 45 ml culture medium and seeded in four 96 well plates. After 24 h selective medium containing G-418 (Geneticin, Gibco-BRL, Gaithersburg, MD, USA) at a final concentration of 400 µg/ml was added. Selected clones were screened by

ELISA of supernatants on plates coated with anti-human IgE to detect the presence of secreted SIPs or by western blotting.

ELISA

ELISA plates were coated with a solution of anti human IgE, gliadin extract or another antigen in carbonate buffer (50 mM Na_2CO_3/$NaHCO_3$ buffer, pH 9.5), overnight at 4°C. After washing twice with PBS-Tween 20 0.1%, it was blocked for 60 min at room temperature, using a solution of 2% milk in PBS. Samples were incubate for 1 hour at room temperature and second antibody was added against IgE (rabbit anti IgE conjugated with HRP,DAKO) and developed with OPD (orthophenilendiamine, SIGMA) in 0,1 M KH_2PO_4/K_2HPO_4, pH 6.0. The reaction was stopped with 2 N H_2SO_4 solution and read at 490 nm. Alternatively, TMB (Pierce) was used and read at 450 nm. After stopping with 2N H_2SO_4.

Western blotting

To detect the presence of scFv, IgG or SIPs in the supernatants of transfected Sp2/0 cells, 10 to 20 μL of supernatants of transfected cells were resuspended in reducing SDS sample buffer and loaded onto a 12% SDS-polyacrilamide gel. After electrophoresis the

gel was transferred to nitrocellulose and incubated with HRP-conjugated antibodies: IgG (DAKO) or IgE (DAKO) and visualised by chemioluminicence (ECL kit, Amersham). In case of analysing through the SV5 tag, incubation with the monoclonal antibody anti-SV5 (117) was carried out before the incubation with the HRP conjugate.

Crosslinking

Media without fetal calf serum was added to cells. After an overnight-incubation, surnatant was taken, dialysed against PBS and DSP (dithiobis[succinimidylpropionate], Pierce) added to a final concentration of 600 μM. After 1 hour-reaction it was taken and further dialysed against PBS. Samples were analysed by western blotting in reducing and non-reducing condition and developed through their SV5 tag.

Phagemid constructs

VH^{PGG}-VH library

Total cellular RNA was isolated from PBMC by the procedure of Chomczynski and Sacchi (118) from normal individuals 0.5-1 μg of RNA was reverse transcribed (RT) using random hexamer and the GeneAmp RNA/PCR kit (Perkin Elmer Cetus, Norwalk, CT), following the procedure recommended by the manufacturer. The cDNA was

amplified using a set of primers (117) for the VH gene and cloned into pDAN vector (89) cutting with Nhe I and Xho I. Once the smallest CDR was identified (see fingerprint of CDR3) it was reamplified with the primers:

PGG/BssH2:

5'AGCAAGCGGCGCGCATGCCCAGGTGCAGCTGGTGCAGTCTGG3' and

PGG/Sal1: 5'GAAGTTATGGTCGACCCTCCGGAGGAGACGGTGACCAGGGTT3'.

Digestion with BssH2 and Sal I were performed for cloning this VH (VH^{PGG}) into the phagemid vector (PDAN-PGG). Different scFVHD were constructed introducing again the cDNA library into VH region of PDAN-PGG using Nhe I and Xho I with primers already published (119).

Antigliadin library

The cDNA from celiac disease patients was given by Bradbury A and Sblattero D. The VH of the anti gliadin library was cloned into PDAN-PGG vector with Nhe I and Xho I, using the same primers as before. After two rounds of selection (see above) one positive clone was selected and introduced replacing VH^{PGG}, and the VHs of the antigliadin library were again cloned into the new vector.

Fingerprint of CDRs 3 (120)

Each clone was amplified with the above VH set of primers. The PCRs were done with 10 pmol of each primer using 30 cycles of denaturation at 94°C for 1 minute, annealing at 60°C for 1 min and synthesis at 72°C for 1 min. 2μl of the PCR product were next labelled by primer extension with the primer hV2 (5'CTGAGGACACGGCC-GTGTATTACTG 3', codon 84 to 92 in FW3). The reaction mixture (2 μl of the PCR sample, 10 pM of ^{32}P end-labelled primer, 200 μM dNTP, 2.5 μl of GeneAmp 10xPCR Buffer and 1.25 U of Taq polymerase [both from Perkin Elmer Cetus], in a final volume of 25 μl) was subjected to denaturation at 94°C for 5 min, annealing at 64°C for 1 min, and extension at 72°C for 15 min. 2 μl of each reaction were analysed on denaturing 6M urea polyacrylamide sequencing gels, with sequencing reactions used as markers.

Phage preparation

Colonies grown after overnight incubation were collected from plates and resuspended in 10 ml 2YT media. Using 100 μl aliquots of these cell suspensions, infection with M13 helper phage, phage propagation and purification were carried out following standard protocols. The incubation temperature for phage propagation was 30°C. Phages were analysed by ELISA or western blotting.

Preparation of soluble scFv

1) Phages from individual colonies were used to infect HB2151 (non-suppressor) E. coli strain. Bacteria were grown in 2xYT amp medium at 37°C to 0.5 O.D.

2) Induced with 0.1 mM IPTG and incubate at 25°C for an additional 5h.

3) The periplasmic scFv fraction was prepared with osmotic shock. Pelleted bacteria were resuspended in PPB buffer (200mg/ml sucrose, 1mM EDTA, 30 mM TrisHCl pH 8) and left on ice for 20 minutes.

4) After centrifugation the supernatant was collected and the cells resuspended in 5mM $MgSO_4$ buffer for 20 minutes

5) The solution was centrifuged and both supernatants were pooled and dialysed against PBS.

Batch Purification on Ni2+ NTA Agarose

All the following steps are carried out at 4°C or on ice:

1) 100 µl of 50% slurry of NTA-agarose (Qiagen cat. no. 30210) was preequilibrated with PBS pH 7.2 in a column.

2) 2 ml periplasmic extract was added to 100 µl pre-equilibrated NTA-agarose and leave running out by gravity.

3) 1 volume column was washed with 5 times volumes PBS and 5 times PBS-30mM imidazol.

4) Elute scFv with 50 µl PBS-250mM imidazole and analyse sample by SDS-PAGE/ western blotting.

5) If required the sample could be dialysed into PBS to remove the imidazole.

Gel filtration

Periplasmic purified protein or concentrated tissue culture media were passed through a gel filtration column (Superdex 75 HR 10/30 Pharmacia) using buffer PBS and a flow rate of 0.5 ml/min. 250 or 500 µl aliquots were collect and analysed by western blotting.

Fingerprinting PCR bands

To make a good library many V regions are required. When amplifying V regions from cDNA it is not obvious from an examination of the amplified band whether a single V region (which may be a contaminant) or many are present. This can be resolved by

carrying out fingerprinting of the V regions using restriction enzymes with 4bp recognition sequences. Each individual V region will give a 'fingerprint'. A diverse collection of V regions will give a smear when cut with such enzymes (caused by an overlap of many different fingerprints), whereas a restricted collection of V regions will give a number of discrete bands. Fingerprinting may be carried out on V regions amplified from cDNA, or a complete library (in which case many different V regions are present), or on V regions amplified from individual clones (each one of which should be different in a good library).

The primers used to amplify VHs were:

a) pDAN3: position 1=pel Bseq 2/ link fort 2
position 2=link back 2/ fd seq

b) pDAN5: position 1=pelB seq 2/ P5 link for
position 2=P5 link back/ fd seq
pelB seq 2= 5'-CAGTCATAATGAAATACC-3'
fd seq= 5'-GAATTTTCTGTATGAGG-3'
link back 2= 5'-TTCTGGGAAACCAGGTTCC-3'
link fort 2= 5'-TACCGCTCGAGCCTTCACC-3'
P5 link bac= 5'-GTATACTATACGAAGTTATCC-3'
P5 link for= 5'-GGATAACTTCGTATAGTATAC-3'

PCR were performed like previously described but with a final volume of 20 μL.

If fingerprinting n clones, a restriction enzyme mix should be made containing:

- (n+1) x 17.8μl water

- (n+1) x 4μl restriction enzyme buffer (New England Biolabs buffer 2)

- (n+1) x 0.2μl restriction enzyme BstNI

- (n+1) x 0.4μl BSA (10mg/ml)

20μl of the mix has to be added to 20 µl PCR mix containing the amplified clone.

Each sample was covered with sellotape and put to incubate at the appropriate temperature (60°C) for 2-3 hours.

To 10 µL mix 2μl 6 X gel loading was added and loaded on 2% Agarose TBE gel.

VH-VH library

Primary library construction by standard cloning techniques

1) cDNA was prepared from peripheral blood lymphocytes as before (VH^{PGG}-VH library)

2) IgM VH genes were first amplified from 0.5μl cDNA reaction, using IgMfor and the individual VHback primers separating each family. Reaction volumes were 20 μl, using 0.5 μl of cDNA reaction, 10 pmol of each primer, 200 μM dNTPs, 2μl 10X PCR buffer, and 0.5 μl (2.5 U) of Taq DNA polymerase (Perkin Elmer). Cycling parameters were

93°C for 30 seconds (denaturation), 55°C for 1 min (annealing) and 72°C for 30 seconds (extension) for thirty-five cycles. Subsequently, VH genes were reamplified using the VHfor and VHback primers in 50μl volumes using 1μl of VH, gel-purified and cloned between Xho1 and Nhe1 sites in pDAN5.

3) For cloning VH in the first position, i.e., VL position, it was similarly amplified, using the same priming region but VL primer tails. Purification was performed as above and pDAN5-VH library (second position) was digested with BssH2 and Sal1 to clone it.

4) The ligation mix was electroporated into electrocompetent DH5alfaF' and plated on 2XTY 100μg/ml ampicillin/ 1% glucose plates to obtain a primary library. Both electroporations achieved an efficiency of 10^7.

5) The colonies were scraped up in 2XTY 10% glycerol and frozen down in 1ml aliquots.

Recombination

1) $3x10^{12}$ phagemid particles obtained from the primary library were added to 20 ml of exponentially growing BS1365 (10^{10} cells) in phagemid particles excess with a MOI of 300:1. After overnight recombination, bacteria were diluted 1/20 into 400ml and phagemid particles prepared by standard techniques (121).

2) As these phagemid particles arise from bacteria containing many different scFvs, there is no coupling between phenotype and genotype. This was overcome by infecting 1 liter of DH5alfaF' ($5x10^8$ cells) at OD600 0.5 at an MOI ≤1 and growing the culture overnight at 30 °C.

3) Phagemid particles were purified from the culture with double PEG precipitation and resuspended in 20ml PBS. This constitutes the final antibody phagemid library, which was used for selections. Aliquots of 10^{12}-10^{13} phagemid particles in 15% glycerol were stored at -80°C.

Selection of phage-antibody libraries by panning in 'immunotubes'

1) To a 75 x12 mm Nunc-immunotube (Maxisorp; Cat. No. 4-44202) 2 ml antigen in carbonate buffer (up to 0.1 mg/ml) was added and left overnight at 4 °C to coat.

2) Next day the tube was washed 2x with PBS-Tween-20 (0.1%) and 2x with PBS

3) For pre-blocking the tube was filled with PBS containing 2% Marvel (2% MPBS). Cover with parafilm and incubate at room temperature for at least 30 min to block (ideally 2 hours).

4) Phage mix was prepared with 500 μl of PEG concentrated phage (~10^{11}-10^{12} TU), 500μl PBS and 1 ml of 4% MPBS; and left for 1 h at room temperature.

5) The tube was washed as step 2.

6) The phage mix from step 4 was added to the Immunotube, incubate 30 min at room temp on under and over turntable; and then stand for at least a further 1.5 hrs at room temp.

7) Tube was washed with 5 washes PBS-Tween-20 (0.1%), then 5 washes PBS. Each washing step is performed by immediately pouring buffer in and out. For subsequent selections more stringent washes are required (for example: 10 times washes with PBS-Tween 20, 1 wash with the same buffer for 30 min, and 10 times washes with PBS alone).

8) To elute the phages from the tube 1 ml of 100 mM HCl was introduced and left rotating for 10 min on an under and over turntable. Phage viability decreases with longer elution times.

9) The elution was transferred to an eppendorf tube with 0.5 ml 1.0 M Tris-HCl, pH 8.0 and mix by inversion.

10) Stored at 4°C for later re-infection.

Displayed antibody/p3 fusion protein is gradually proteolysed during storage, so we generally perform our phage rescues and selections on two consecutive days.

Rescuing phagemid libraries

1) The inoculum size should be 10x library size in number of bacteria at the start, but should not exceed 0.05 O.D.600nm.

The inoculum was transferred into 10 ml of 2 x TY, 100 μg/ml ampicillin, 2% glucose, in a 50ml Falcon tube, or 20ml in a 250ml flask, generally an inoculation of 10 µl of concentrated bacterial stock.

2) After growing with shaking (270 rpm) for 1,5-2,5 hours at 37°C, to an O.D.600nm of 0.5, 5 ml (2 x 10^9 bacteria) was transferred to a 50-ml Falcon tube containing an appropriate amount of helper phages, in a ratio between 10:1 and 20:1, phage: bacteria.

3) Incubation at 37°C for 30 minutes with occasional agitation was carried out.

4) For a first rescue of a large repertoire, also a shaking at 37°C for 30 min, at 100 rpm, is recommended.

5) After the infection event, cells were pelleted for 10 min. at 3.000 rcf, and the supernatant removed.

6) The bacterial pellet was resuspended in 10-20 ml of 2xTY, 100 μg/ml ampicillin, 25 μg/ml kanamycin, transferred to a 50ml Falcon (10ml) or 250ml flask (20ml) and further incubated with shaking (270 rpm) overnight at 30°C (scFv).

7) The culture was spinned in a 50-ml Falcon tube for 20 minutes at 3.000 rcf to pellet the bacteria and keeping the surnatant.

7) To the supernatant a PEG solution (20 % Polyethylene glycol 6000, 2.5 M NaCl) (1/5th of the surnatant volume) was added and left on ice for 1 hour.

8) Phages were pelleted by spinning for 15 min., 3.000 rcf at 4°C throwing away the supernatant. The pellet was resuspended in 1.0 ml PBS with a blue 1-ml filter-tip and transferred to a 1.5 ml eppendorf tube.

9) A further spinning in microcentrifuge (2 min, max speed), to remove the remaining bacteria, was performed and the supernatant transferred to a new tube.

11) The phage is now ready for selection. Phage can be stored at 4°C without much loss of titre; the antibodies may however proteolytically be removed by contaminating proteases, and, for selection, phages should be used within a week.

The standard yield is about 2-10 x 10^{12} phages from a 25 ml culture.

Growth and rescue of phage(mid) particles in 96-well micotitre plates.

1) Colonies were toothpicked into 150 μl 2xTY, 100 μg/ml ampicillin, 2% glucose in 96-well flat-bottomed plates (Costar, Cat No. 3595) and grown with shaking (270 rpm.) overnight at 30°C.

2) Next day, 2 μl was inoculated from this plate to a fresh 96-well plate containing 150 μl 2xTY, 100 μg/ml ampicillin, 1% glucose per well. To the wells of the master plate, 50 μl 60 % glycerol per well were added and after use stored at -70°C.

3) After growing 2.5 hrs at 37°C with shaking, infection with M13K07 phage was carried out (50 μl 2xTY, 100 μg/ml ampicillin, 1% glucose containing $2x10^9$ pfu) at 37°C for 30 min. The ratio of phage to bacterium should be around 20:1.

4) Supernatant was removed by spinning at 600 rcf for 10 min and pellet resuspended in 150 μl 2 x TY, 100 μg/ml ampicillin, 25 μg/ml kanamycin for a further overnight growing at 30°C (scFv), with shaking.

5) Next day, 50 μl supernatant per well was taken for phage ELISA after spinning at 600 rcf for 10 min.

Induction of soluble antibody fragments in 96-well plates

This method relies on the low levels of glucose present in the starting medium being metabolised by the time the inducer (IPTG) is added (122).

1) Single clones were inoculated into 150 μl 2xTY, 100 μg/ml ampicillin, 1% glucose in 96-well plates (Costar 3595) and grown with shaking overnight at 30°C.

2) A small inocula (2 μl) from the previous plate was transferred to a second 96-well plate containing 100 μl fresh 2 x TY, 100 μg/ml ampicillin, 0.1% glucose per well, and was grown at 37°C with shaking, until O.D. 600nm is approximately 0.7 (about 2-3 hrs). The master plate was stored at -70°C after adding 50 μl 60% glycerol per well.

3) The expression of scFv was induced with a final concentration of 1 mM IPTG in 2x TY, 100 μg/ml ampicillin, and shaking at 30°C for a further 16 to 24 hrs.

4) After spinning at 600 rcf for 10 min 50 μl supernatant was used for ELISA screening.

ELISA for detection of soluble or phage antibody fragments

1) Plate was coated with 100 μl per well of protein antigen used for selections. Coating is carried out overnight at 4°C in 100 mM sodium hydrogen carbonate, pH 9.6.

2) Wells were rinsed 3x with PBS-Tween-20 (0.1%), 3x with PBS, and blocked with 150 μl per well of 2% Marvel/PBS (MPBS), for at least 30 min at room temperature.

3) After one more round of washing with PBS-Tween-20 (0.1%) and with PBS alone, 50 μl of 4% MPBS to all wells was added.

4) 50 μl culture supernatant containing soluble antibody fragment or phage antibody to the appropriate wells was adjoined, mixed by pipetting up and down and left approx. for 1.5 hrs at room temperature.

5) After discarding the solution, wells were washed out 3x with PBS/Tween-20 (0.1%) and 3x with PBS alone.

6) For soluble antibody detection: 100 μl of SV5 antibody in 2% MPBS was introduced into each well and left incubating at room temp. for 1 hr. Final concentration of SV5 antibody should be in the 1 μg/ml range.

For phage antibody detection: 100µl anti fd-phage HRP conjugated (1/5000 dilution in 2% MPBS) was used.

7) After discarding the antibody, wells were washed as step 5.

8) For soluble antibody detection an additional step is required adding 100 μl of 1:1000 dilution (in 2% MPBS) of anti-mouse antibody (peroxidase-conjugated anti-mouse immunoglobulins and incubating at room temp. for 1 hr.

9) Washing step as 5 was again performed

10) To develop the ELISA 100 µl of TMB (Pierce) was used and left in the dark at room temperature for 10-30 min.

12) The reaction was quenched by adding 50 μl 2 N H_2SO_4, and the plate read at 450 nm.

Western blotting for phages

Western blotting was performed using purified phages prepared as described above. 10 µl of resuspended phages was blotted onto nitrocellulose following SDS PAGE using standard protocols. After blocking with MPBS the nitrocellulose sheets were sequentially incubated with anti pIII-HRP or anti SV5 tag monoclonal antibody. Anti mouse Igs-horseradish peroxidase diluted 1:5000 (Dako) was used as second antibody for anti SV5 monoclonal antibody. The positive bands were revealed by ECL (Amersham).

CHAPTER 5

References

1. Wolfgang WA *et al.* (2000) **Monomeric and oligomeric complexes of the B-cell antigen receptor.** Immunity 13: 5.

2. Randall TD *et al.* (1992) **Direct evidence that J chain regulates the polymeric structure of IgM in antibody-secreting B cells.** J Biol Chem 267: 18002.

3. Paul W (1999) **Fundamental Immunology. Fourth Edition.** Lippincott-Raven Publishers. Philadelphia, NY, USA.

4. Rees A *et al.* (1994) **Antibody design:beyond the natural limits.** TIBTECH 12: 199.

5. Kabat E *et al.* (1987) **Sequences of proteins of immunological interest.** US Dept. Health & Human Services, Washington, D.C.

6. Novotny J *et al.* (1985) **Structural invariants of antigen binding:Comparison of immunoglobulin VL-VH and VL-VL domain dimers.** Proc Natl Acad Sci USA 82: 4592.

7. Chatellier *et al.* (1996) **Functional mapping of conserved residues located at the VL and VH domain interface of Fab.** J Mol Biol 264: 1.

8. Hamers-Casterman C *et al.* (1993) **Naturally occurring antibodies devoid of light chains.** Nature 363: 446.

9. Primi D *et al.* (1987) **Highly preferential VH-VL pairing in normal B cells results in antigen-independent selection of the available repertoire.** J Immunol. 138: 1607.

10. Colman PM (1988) **Structure of antibody-antigen complexes: implications for immune recognition.** Adv Immunol 43: 99.

11. Stanfield RL *et al.* (1994) **Antigen-induced conformational changes in antibodies: a problem for structural prediction and design.** Trends Biotechnol 12: 275.

12. Saul FA *et al.* (1978) **Preliminary refinement and structural analysis of the Fab fragment from human immunoglobulin New at 2.0 Angstrom resolution.** J Biol Chem 253: 585.

13. Wilson IA *et al.* (1994) **Antibody-antigen interactions: new structures and new conformational changes.** Curr Opin Struct Biol 4: 857.

14. Nieba *et al.* (1997) **Disrupting the hydrophobic patches at the antibody variable/constant domain interface: improved in vivo folding and physical characterization of an engineered scFv fragment.** Prot Eng 10: 435.

15. Matsuda *et al.* (1998) **The complete nucleotide sequence of the human immunoglobulin heavy chain variable region locus** J Exp Med 188: 1.

16. Cook GP *et al.* (1995) **The human immunoglobulin VH repertoire.** Immunol Today 16: 237

17. Corbett, SJ *et al.* (1997) **Sequence of the human immunoglobulin diversity (D) segment locus: a systematic analysis provides no evidence for the use of DIR segments, inverted D segments, "minor" D segments or D–D recombination.** J Mol Biol 270: 587.

18. Mattila PS *et al.* (1995) **Extensive allelic sequence variation in the J region of the human immunoglobulin heavy chain gene locus.** Eur J Immunol 25: 2578.

19. Rathbun G *et al.* (1989) **Organization and expression of the mammalian heavy-chain variable-region locus.** In immunoglobulin genes page: 63 Academic Press, New York.

20. Maki R *et al.* (1981) **The role of DNA rearrangement and alternate RNA processing in the expression of immunoglobulin δ genes.** Cell 24: 353.

21. Blattner F *et al.* (1984) **The molecular biology of immunoglobulin D.** Nature 307: 417.

22. Tonegawa S *et al.* (1983) **Somatic generation of antibody diversity.** Nature 302: 575

23. van Gent *et al.* (1997) **Stimulation of V(D)J cleavage by high mobility group proteins.** EMBO J 16: 2665.

24. McBlane JF *et al.* (1995) **Cleavage at a V(D)J recombination signal requires only RAG1 and RAG2 proteins and occurs in two steps.** Cell 83: 387.

25. Melek M *et al.* (1998) **Rejoining of DNA by the RAG1 and RAG2 proteins.** Science 280: 301.

26. Murray S *et al.* (2000) **Crystal structure of the Xrcc4 DNA repair protein and implications for end joining.** EMBO J 19: 5962.

27. Grawunder *et al.* (1997) **A complex of RAG1 and RAG2 proteins persists on DNA after single strand cleavage at V(D)J recombination signal sequences.** Nucleic Acids Res 25: 1375.

28. Pelanda R *et al.* (1996) **A prematurely expressed Ig (kappa) transgene, but not V(kappa)J(kappa) gene segment targeted into the Ig(kappa) locus, can rescue B cell development in lambda5-deficient mice.** Immunity 5: 229.

29. Mostoslavsky R *et al.* (1998) **Kappa chain monoallelic demethylation and the establishment of allelic exclusion.** Genes Dev 15: 1801.

30. Banerji, J. *et al.* (1983) **A lymphocyte specific cellular enhancer is located downstream of the joining region in immunoglobulin heavy chain genes.** Cell, 33: 729.

31. Gillies S *et al.* (1983). **A tissue-specific transcription enhancer element is located in the major intron of a rearranged immunoglobulin heavy chain gene.** Cell 33: 717.

32. Mills FC *et al.* (1997) **Enhancer complexes located downstream of both human immunoglobulin C genes.** J Exp Med 186: 845.

33. Picard, D. *et al.* (1984) **A lymphocyte specific enhancer in the mouse immunoglobulin k gene.** Nature 307: 80.

34. Asenbauer H *et al.* (1999) **The immunoglobulin lamda light chain enhancer consists of three modules which sinergize in activation of transcription.** Eur J Immunol 29: 713.

35. Haas IG *et al.* (1983) **Immunoglobulin heavy chain binding protein.** Nature 306: 387.

36. Hendershot L *et al.* (1987) **Assembly and secretion of heavy chains that do not associate posttranslationally with immunoglobulin heavy chain-binding protein.** J Cell Biol 104: 761.

37. Lee YK *et al.* (1999) **BiP an immunoglobulin light chain cooperate to control the folding of heavy chain and ensure the fidelity of immunoglobulin assembly.** Mol Biol Cell 10: 2209.

38. Melnick J *et al.* (1994) **Sequential interaction of the chaperone BiP and GRP94 with immunoglobulin chains in the endoplasmic reticulum.** Nature 370: 373.

39. Nicchitta CV (1998) **Biochemical, cell biological and immunological issues surrounding the endoplasmic reticulum chaperone GRP94/gp96.** Curr Opin Immunol 10:103.

40. Batista F *et al.* (1996) **Characterisation of a second secreted igE isoform and identification of an asymmetric pathway of IgE assembly.** Proc Natl Acad Sci USA 93: 3399.

41. Zhang K *et al.* (1994) **Complex alternative RNA splicing of epsilon-immunoglobulin transcripts produces mRNAs encoding four potential secreted protein isoforms.** J Biol Chem 269: 31322.

42. Peng C *et al.* (1992) **A new isoform of human membrane-bound IgE.** J Immunol 148: 129.

43. Zhang K *et al.* (1992) **Two unusual forms of human immunoglobulin E encoded by alternative RNA splicing of epsilon heavy chain membrane exons.** J Exp Med 176: 233.

44. Batista FD *et al.* (1995) **Characterization and expression of alternatively spliced IgE heavy chain transcripts produced by peripheral blood lymphocytes.** J Immunol 154: 209.

45. Batista FD *et al.* (1996) **The two membrane isoforms of human IgE assemble into functionally distinct B cell antigen receptors.** J Exp Med 184: 2197.

46. Bestagno M *et al.* (2001) **Membrane immunoglobulins are stabilized by inter-chain disulphide bonds occurring within the extracellular membrane-proximal domain.** Sumitted.

47. MacLennan I *et al.* (1994) **Germinal Centers.** Annual Rev Immunol 12: 117.

48. Berek C *et al.* (1987) **Mutation drift and repertoire shift in the maturation of the immune response.** Immunol Rev 96: 23.

49. Williams GT *et al.* (2000) **The contribution of somatic hypermutation to the diversity of serum immunoglobulin: dramatic increase with age.** Immunity 13:409.

50. Jack H *et al.* (1988) **Circular DNA is excised by immunoglobulin class switch recombination.** Cell 62: 143.

51. De Pinho R *et al.* (1984) **Molecular basis of of heavy chain class switching and switch region deletion.** Mol Cell Biol 4: 2905.

52. Han S *et al.* (1996) **Neoteny in lymphocytes: Rag1 and Rag2 expression in germinal center B cells.** Science 20: 2094.

53. Muramatsu M (2000) **Class swith recombination and hypermutation require activation-induced cytidine deaminase (AID), a potential RNA editing enzyme.** Cell 102: 53.

54. Chothia C *et al.* (1989) **Conformations of immunoglobulin hypervariable regions.** Nature 342: 877.

55. Chothia C *et al.* (1992) **Structural repertoire of the human VH segments.** J Mol Biol 227: 799.

56. Al-Lazikani B *et al.* (1997) **Standard conformations for the canonical structures of immunoglobulins.** J Mol Biol 273: 927.

57. Morea V *et al.* (1998) **Conformations of the third hypervariable region in the VH domain of immunoglobulins.** J Mol Biol:275: 269.

58. Davies DR *et al.* (1996) **Interactions of protein antigens with antibodies.** Proc Natl Acad Sci USA 93: 7.

59. Rock EP *et al.* (1994) **CDR3 length in antigen-specific immune receptors.** J Exp Med 179: 323.

60. Sanz I (1991) **Multiple mechanisms participate in the generation of diversity of human H chain CDR3 regions.** J Immunol 147: 1720.

61. Yamada M *et al.* (1991) **Preferential utilization of specific immunoglobulin heavy chain diversity and joining segments in adult human peripheral blood B lymphocytes.** J Exp Med 173: 395.

62. Xu Jl *et al.* (2000) **Diversity in the CDR3 region of V(H) is sufficient for most antibody specificities.** Immunity 13:37

63. Madan MK *et al.* (2000) **Specific antibody production by V(H)-gene replacement.** Eur J Immunol 30: 2404.

64. Alt FW *et al.* (1982) **Joining of immunoglobulin heavy chain gene segments: implications from a chromosome with evidence of three D-JH fusions.** Proc Natl Acad Sci U S A 79: 4118.

65. Ohlin M *et al.* (1998) **Insertions and deletions in hypervariable loops of antibody heavy chains contribute to molecular diversity.** Mol Immunol 35: 233.

66. de Wildt RM *et al.* 1(999) **Somatic insertions and deletions shape the human antibody repertoire.** J Mol Biol 294: 701.

67. Stewart AK *et al.* (1993) **High-frequency representation of a single VH gene in the expressed human B cell repertoire.** J Exp Med 177: 409.

68. Cox JP *et al.* (1994) **A directory of human germ-line V kappa segments reveals a strong bias in their usage.** Eur J Immunol 24: 827.

69. Suzuki I *et al.* (1995) **Representation of rearranged VH gene segments in the human adult antibody repertoire.** J Immunol 154: 3902.

70. Tomlinson IM *et al.* (1996) **The imprint of somatic hypermutation on the repertoire of human germline V genes.** J Mol Biol 256: 813.

71. Ignatovich O *et al.* (1997) **The creation of diversity in the human immunoglobulin V(lambda) repertoire.** J Mol Biol 268: 69.

72. Kohler G *et al.* (1975) **Continuous cultures of fused cells secreting antibody of predefined specificity**. Nature 256: 495.

73. Bird RE *et al.* (1988) **Single-chain antigen-binding proteins**. Science 242: 423.

74. Huston JS *et al.* (1988) **Protein engineering of antibody binding sites: recovery of specific activity in an anti-digoxin single-chain Fv analogue produced in Escherichia coli.** Proc Natl Acad Sci USA 85: 5879.

75. Better M *et al.* (1988) **Escherichia coli secretion of an active chimeric antibody fragment**. Science 240: 1041.

76. Skerra A *et al.* (1988) **Assembly of a functional immunoglobulin Fv fragment in Escherichia coli**. Science. 240: 1038.

77. Barbas CF *et al.* (1991) **Assembly of combinatorial antibody libraries on phage surfaces: the gene III site**. Proc Natl Acad Sci USA 88: 7978.

78. Griffiths AD *et al.* (1993) **Human anti-self antibodies with high specificity from phage display libraries**. EMBO J 12:725.

79. Griffiths AD *et al.* (1994) **Isolation of high affinity human antibodies directly from large synthetic repertoires**. EMBO J 15;13: 3245.

80. Vaughan TJ *et al.* (1996) **Human antibodies with sub-nanomolar affinities isolated from a large non-immunized phage display library**. Nat Biotechnol 14: 309.

81. Sheets MD *et al.* (1998) **Efficient construction of a large nonimmune phage antibody library: the production of high-affinity human single-chain antibodies to protein antigens**. Proc Natl Acad Sci USA 95: 6157.

82. McCafferty J *et al.* (1990) **Phage antibodies: filamentous phage displaying antibody variable domains**. Nature 348: 552.

83. Hoogenboom HR *et al.* (1991) **Multi-subunit proteins on the surface of filamentous phage: methodologies for displaying antibody (Fab) heavy andlight chains**. Nucleic Acids Res 19: 4133.

84. Marks JD *et al.* (1991) **By-passing immunization. Human antibodies from V-gene libraries displayed on phage**. J Mol Biol 222: 581.

85. Winter G *et al.* (1994) **Making antibodies by phage display technology**. Annu Rev Immunol 12: 433

86. Hoogenboom HR *et al.* (1992) **By-passing immunisation. Human antibodies from synthetic repertoires of germline VH gene segments rearranged in vitro.** J Mol Biol 227: 381.

87. Nissim A *et al.* (1994) **Antibody fragments from a 'single pot' phage display library as immunochemical reagents.** EMBO J 13: 692.

88. Kruif J *et al.* (1995) **Selection and application of human single chain Fv antibody fragments from a semi-synthetic phage antibody display library with designed CDR3 regions.** J Mol Biol 248: 97.

89. Sblattero D *et al.* (2000) **Exploiting recombination in single bacteria to make large phage antibody libraries.** Nat Biotechnol 18: 75.

90. Hendershot LM (1990) **Immunoglobulin heavy chain and binding protein complexes are dissociated in vivo by light chain addition.** J Cell Biol 111: 829.

91. Morrison SL *et al.* (1975) **Heavy chain-producing variants of a mouse myeloma cell line.** J Immunol 114: 655.

92. Melnick J *et al.* (1992) **The endoplasmic reticulum stress protein GRP94, in addition to BiP, associates with unassembled immunoglobulin chains.** J Biol Chem 267: 21303.

93. Brandt CR *et al.* (1984) **Loss of a consensussplice signal in a mutant immunoglobulin gene eliminates the CH1 domain exon from the mRNA.** Mol Cell Biol 4: 1270.

94. Kortt AA *et al.* (1995) **Solution Properties of Escerichia coli-expressed VH domain of antineuraminidase antibody NC41.** J Prot Chem 14: 1671.

95. Guex N *et al.* (1997) **SWISS-MODEL and the Swiss-PdbViewer: An environment for comparative protein modelling.** Electrophoresis 18: 2714.

96. Peitsch, MC(1996) ProMod and Swiss-Model: **Internet-based tools for automated comparative protein modelling.** Biochem Soc Trans 24: 274.

97. Peitsch, MC (1995) **Protein modeling by E-mail** Bio/Technology 13: 658.

98. Chevet E *et al.* (2001) **The endoplasmic reticulum: integration of protein folding, quality control, signaling and degradation** Current Opinion in Structural Biology 11: 120.

99. Dul JL *et al.* (1990) **A single aminoacid substitution in the variable region of the light chain specifically blocks immunoglobulin secretion.** Proc Natl Acad Sci USA 87: 8135.

100. Bole DG *et al.* (1986) **Posttranslational association of immunoglobulin heavy chain binding protein with nascent heavy chains in non-secreting and secreting hybridomas** J Cell Biol 102: 1558.

101. Li E *et al.* (1997) **Mammalian cell expression of dimeric small imune proteins (SIP)** Protein Eng 10: 731.

102. Knappik A *et al.* (1995) **Engineered turns of a recombinant antibody improve its in vivo folding.** Protein Eng 8: 81.

103. Glockshuber R *et al.* (1990) **A comparison of strategies to stabilize immunoglobulin Fv-fragments.** Biochemistry 29: 136.

104. Power *et al.* (1992) **High-level temperature-induced synthesis of an antibody VH-domain in Escherichia coli using the PelB secretion signal.** Gene 113: 95.

105. Bhat TN et al. (1994) **Bound water molecules and conformational stabilization help mediate an antigen-antibody association.** Proc Natl Acad Sci U S A 91: 1089.

106. Goto Y *et al.* (1982) **Unfolding and refolding of the reduced constant fragment of the light chain.** J Mol Biol 156: 911.

107. Chothia C *et al.* (1985) **Domain association in immunoglobulin molecules.** J Mol Biol 186: 651.

108. Lilie H *et al.* (1995) **Association of antibody chains at different stages of folding:prolyl isomerization occurs after formation at quaternary structure.** J Mol Biol 248: 190.

109. Kaloff CR *et al.* (1995) **Coordination of immunoglobulin chain folding and immunoglobulin chain assembly is essential for the formation of functional IgG.** Immunity 2: 629.

110. Knittler MR *et al.* (1995) **Molecular chaperones involved in protein degradation in the endoplasmic reticulum: quantitative interaction of BiP with partially folded immunoglobulin light chains that are degradaded in the endoplasmic reticulum.** Proc Natl Acad Sci USA 92: 1764.

111. Dul JL *et al.* (1992) **A conditional secretory mutant in an Ig L chain is caused by replacement of tyrosine/phenylalanine 87 with histidine.** J Immunol 149: 1927.

112. Geoffroy FR *et al.* (1994) **A new phage display system to construct multicombinatorial libraries of very large antibody repertoires.** Gene 151: 109.

113. Tsurushita NH *et al.* (1996) **Phage display vectors for in vivo recombination of immunoglobulin heavy and light chain genes to make large combinatorial libraries.** Gene 172: 59.

114. Waterhouse PAD *et al.* (1993) **Combinatorial infectionand in vivo recombination: a strategy for making largephage antibody repertoires.** Nuc Acids Res 21: 2265.

115. Davies J *et al.* (1994) **Camelising human antibody fragments: NMR stuides on VH domains.** FEBS Letters 339: 285.

116. Reiter Y *et al.* (1999) **An antibody single-domin phage display library of a native heavy chain variable region: isolation of functional single domain VH molecules with a unique interface.** J Mol Biol 290: 685.

117. Hanke TP *et al.* (1992) **Construction of solid matrix-antibody-antigen complexescontaining simian immunodeficiency virus p27 using tag-specific monoclonal antibody and tag-linked antigen.** J gen Virol 73: 653.

118. Chomczynski P *et al.* (1987) **Single-step method of RNA isolation by acid guanidinium thiocyanate-phenol-chloroform extraction.** Anal Biochem 162: 156.

119. Sblattero D *et al.* (1998) **A definitive set of oligonucleotide primers for amplifying human V regions.** Immunotechnology 3: 271.

120. Efremov *et al.* (1993) **Molecular analysis of IgE heavy chain transcripts expressed in vivo by peripheral blood lymphocytes from normal and atopic individuals.** J Immunol 151: 2195.

121. Marks JD *et al.* (1992) **By-passing immunization: building high affinity human antibodies by chain shuffling.** Biotechnology (N Y) 10: 779.

122. De Bellis D *et al.* (1990) **Regulated expression of foreign genes fused to lac: control by glucose levels in growth medium.** Acids Res 18: 1311.