



A human scFv antibody generation pipeline for proteome research

Michael Hust^{*}, Torsten Meyer, Bernd Voedisch¹, Torsten Rülker, Holger Thie², Aymen El-Ghezal³, Martina Inga Kirsch⁴, Mark Schütte, Saskia Helmsing, Doris Meier, Thomas Schirrmann⁵, Stefan Dübel⁵

Technische Universität Braunschweig, Institut für Biochemie und Biotechnologie, Abteilung Biotechnologie, Spielmannstr. 7, 38106 Braunschweig, Germany

ARTICLE INFO

Article history:

Received 30 March 2010
Received in revised form
13 September 2010
Accepted 16 September 2010
Available online 29 September 2010

Keywords:

Phage display
Antibody engineering
Panning
scFv
scFv-Fc
Immunoglobulin
mAb, Monoclonal antibody
CD71
Human proteome research

ABSTRACT

The functional decryption of the human proteome is the challenge which follows the sequencing of the human genome. Specific binders to every human protein are key reagents for this purpose. *In vitro* antibody selection using phage display offers one possible solution that can meet the demand for 25,000 or more antibodies, but needs substantial standardisation and minimalisation. To evaluate this potential, three human, naive antibody gene libraries (HAL4/7/8) were constructed and a standardised antibody selection pipeline was set up. The quality of the libraries and the selection pipeline was validated with 110 antigens, including human, other mammalian, fungal or bacterial proteins, viruses or haptens. Furthermore, the abundance of VH, kappa and lambda subfamilies during library cloning and the *E. coli* based phage display system on library packaging and the selection of scFvs was evaluated from the analysis of 435 individual antibodies, resulting in the first comprehensive comparison of V gene subfamily use for all steps of an antibody phage display pipeline. Further, a compatible cassette vector set for *E. coli* and mammalian expression of antibody fragments is described, allowing *in vivo* biotinylation, enzyme fusion and Fc fusion.

© 2010 Elsevier B.V. All rights reserved.

1. Introduction

To date, antibodies represent the fastest growing class of biological therapeutics on the market and are indispensable as detection reagents in research and diagnostics (Dübel, 2007; Hagemeyer et al., 2009; Nieri et al., 2009). For proteome research, antibodies or alternative scaffolds are a key tool for the decryption of the human proteome (Aebersold and Mann, 2003; Berglund et al., 2008; Dübel et al., 2010; Hust and Dübel, 2004; Ohara et al., 2006; Skerra, 2007; Taussig et al., 2007; Wingren et al., 2009). After sequencing of human genomes, including the genome of individual persons (International Human Genome Sequencing Consortium, 2004; Levy et al., 2007; Wheeler et al., 2008) is well established, the research focus shifted to the analysis of gene products. The estimated number of protein encoding human genes is about 20,000–25,000 (International Human Genome Sequencing

Consortium, 2004; Levy et al., 2007), due to alternative mRNA splicing variants and post-translational modifications including glycosylations, phosphorylation, sulfation etc. the number of different human proteins and protein variants is supposed to exceed the size of the genome severalfold (Harrison et al., 2002). The development of highly parallelized protein and tissue microarrays may revolutionise the proteome analysis in the same way nucleic acid microarrays accelerated the genome and transcriptome analysis (Stoevesandt et al., 2009).

For the human proteome atlas (www.proteinatlas.org), affinity purified polyclonal antibodies directed against protein expression signature tags (PRESTs) were generated in substantial number (Berglund et al., 2008). However, polyclonal antibodies are only a limited resource and without guarantee of constant quality. This can be overcome by the use of monoclonal antibodies. However, in respect of throughput, and despite becoming streamlined and automated to a considerable extent (De Masi et al., 2005), hybridoma technology still requires immunising of individual animals. Further, the immune system still defines the minimal time to obtain an immune response, whereas *in vitro* methods allow selection within a few days. Both, polyclonal and hybridoma technologies, do not provide direct access to the antibody gene sequence. Hence, subsequent genetic modifications to improve properties (e.g. affinity, stability etc.) or assay compatibility (e.g. fusion to alternative tags, switch between mono-, bi- or multivalency) are laborious. An

^{*} Corresponding author. Tel.: +49 531 391 5760; fax: +49 531 391 5763.

E-mail address: m.hust@tu-bs.de (M. Hust).

¹ Present address: Novartis Institutes for BioMedical Research, Basel, Switzerland.

² Present address: Miltenyi Biotec GmbH, Bergisch Gladbach, Germany.

³ Present address: Chemengineering Technology GmbH, Wiesbaden, Germany.

⁴ Present address: Sanofi-Aventis, Frankfurt, Germany.

⁵ Both senior authors contributed equally for this work.

added value is provided by the *in vitro* antibody selection methods which supply the antibody gene right from the start, for example to achieve functional knock down phenotypes of the protein bound by the antibody (Böldicke, 2007; Strebe et al., 2009).

In vitro technologies have demonstrated their potential to meet the demands of larger scale antibody generation projects (Deflorian et al., 2009; Dübel et al., 2010; Hallborn and Carlsson, 2002; Hust and Dübel, 2004; Konthur et al., 2005; Krebs et al., 2001; Mersmann et al., 2010; Ohara et al., 2006; Taussig et al., 2007), and still offer much potential to be further shortened and miniaturised, which mice and rabbits cannot. Here, the full potential of other *in vitro* selection display technologies, e.g. ribosome display (Hanes and Plückthun, 1997; He and Taussig, 2008), remains to be explored as well.

In the last two decades the development of phage display was driven mainly by the opportunity to make human antibodies for therapeutic applications. However, the technology is proven to yield valuable research tools, which start to appear on the market (Zitat: http://www.axxora.com/scripts/pdf/pdf_megadetail.php?PID=AG-27B-0001). To achieve that, high efforts were invested in the generation of a large number (i.e. “as many as possible”) of antibody candidates against few well-characterised antigens (Dübel, 2007; Hoogenboom, 2005; Taussig et al., 2007; Thie et al., 2008). In contrast, for research antibodies or alternative scaffolds have to be generated against thousands of poorly characterised proteins, many of them not yet known at all. Consequently, an integrated process combining antibody selection, screening and production which is amenable to parallelisation and miniaturisation as well as to the integration of different antigen sources, like peptides, proteins, fusion proteins, protein, protein complexes, PRESTs, is required (Dübel et al., 2010; Hust and Dübel, 2004; Konthur et al., 2005; Taussig et al., 2007).

The generation and application of very large standardised universal ‘single pot’ antibody gene libraries, which in principle contain binders against every possible antigen, is a key to the success of an universal antibody generation pipeline. Therefore, a naive antibody gene library was constructed for a simplified antibody selection pipeline designed to be able to generate antibodies for proteome research for reasonable cost. Beyond the use in research, since it employs a human antibody repertoire, antibodies selected from this library may further be developed for use in diagnostics and therapy with little effort. The quality of the library and process was validated by the selection of binders to 110 antigens.

2. Material and methods

2.1. Cell lines and bacterial strains

E. coli XL1-Blue MRF⁺ (Stratagene, Amsterdam, NL) (F⁺::Tn10(Tetr) proAB⁺ lacIq Δ(lacZ)M15/recA1 endA1 gyrA96 (Nal^r) thi hsdR17 (rK-mK⁺) glnV44 relA1 lac) was used for antibody gene library construction, phage display and screening.

The transformed human embryonic kidney (HEK) cell line 293T (American Type Culture Collection, ATCC, Rockwell, MD, No. CRL-11268) was cultured in DMEM (4.5 g/L glucose) supplemented with 2 mM L-glutamine, 10% (v/v) fetal calf serum (FCS) and 1% (v/v) penicillin/streptomycin (PAA, Parsing, Austria). The human cervix carcinoma cell line HeLa was grown in the same medium with additional supplementation of 1 mM sodium pyruvate and non-essential aminoacids (PAA). CHO-TRVb CD71-negative variant of chinese hamster ovarian cell line CHO and CHO-TRVb-1 transfected with human CD71 both kindly provided by Tim McGraw (Cornell University, New York, NY, USA) were cultured as described (McGraw et al., 1987). All mammalian cells were grown at 37 °C, 5% CO₂ atmosphere and 95% humidity.

2.2. Construction of the human antibody gene libraries

Lymphocytes were isolated from 44 donors with various ethnical backgrounds using the Lymphoprep Kit (Progen, Heidelberg) according to the manufacturers instructions. Total RNA was isolated using Trizol (Invitrogen, Karlsruhe) and mRNA (from all samples) was isolated using the QuickPrep micro mRNA Purification Kit (GE Healthcare, München) leading to 81 RNA preparations. Then cDNA was synthesised separately from each RNA sample using Superscript II (Invitrogen), random hexamer oligonucleotide primers and 50–250 ng mRNA or 2–20 µg total RNA, respectively.

To amplify the antibody gene fragments (kappa, lambda, VH) 27 individual first PCRs of each cDNA sample were performed separately for each cDNA preparation in a volume of 50 µL using Red Taq (Sigma, Hamburg) and 0.4 µM of each primer (Table 1) for 30 cycles (1 min 94 °C, 1 min 55 °C, 2 min 72 °C) followed by a 10 min final synthesis step. The PCR products were purified by agarose gel electrophoresis using the Nucleospin Extract 2 Kit (Macherey-Nagel, Düren). The VH, kappa or lambda PCR products were pooled separately. To add restriction sites for cloning, 100 ng of the purified and pooled VH, kappa or lambda PCR products of the first PCR for each of the 27 s PCR reactions were used in a volume of 100 µL using Red Taq and 0.2 µM of each primer (Table 1) for 20 cycles (1 min 94 °C, 1 min 57 °C, 2 min 72 °C) followed by a 10 min final synthesis step at 72 °C. The PCR products were purified by agarose gel electrophoresis using the Nucleospin Extract 2 Kit and VH, kappa or lambda were pooled separately. In total 2187 first PCR and 2187 s PCR reactions were performed.

For VL cloning 5 µg vector pHAL14 and 2 µg kappa or lambda PCR products were digested in a volume of 100 µL using 30 U *Mlu*I and *Not*I (NEB, Frankfurt) at 37 °C for 2 h. This step was performed multiple times to produce enough material for up to 20 transformations. For quality control each digestion was performed separately with each enzyme in parallel. The digestions were inactivated at 65 °C for 10 min, followed by adding 0.5 U calf intestine phosphatase (CIP) (MBI Fermentas, St. Leon-Rot) and incubation at 37 °C for 30 min, this step was repeated once. The vector and the PCR products were purified using the Nucleospin Extract 2 Kit, removing short stuffer fragments containing multiple stop codons between *Mlu*I and *Not*I in pHAL14. Each ligation was performed with 1 µg digested vector and 270 ng PCR product using 3 U T4 ligase (Promega, Mannheim) in a volume of 100 µL overnight at 16 °C. The ligations were inactivated at 65 °C for 10 min and precipitated with 10 µL 3 M sodium acetate pH 5.2 and 250 µL ethanol for 2 min at RT, followed by 5 min centrifugation at 16,000 × g and 4 °C. The pellet was washed two times with 500 µL 70% (v/v) ethanol. The dried pellet was resuspended in 35 µL dH₂O and mixed on ice with 25 µL electrocompetent XL1-Blue MRF⁺. The electroporation was performed with 0.1 cm prechilled cuvettes using a MicroPulser electroporator (BIO-RAD, München) at 1.7 kV. Immediately, 1 mL prewarmed 37 °C SOC medium pH 7.0 (2% (w/v) tryptone, 0.5% (w/v) yeast extract, 0.05% (w/v) NaCl, 20 mM Mg solution, 20 mM glucose) was added and incubated for 1 h at 600 rpm and 37 °C. The transformation was plated out on 25 cm square Petri dishes with 2xYT-GAT agar (1.6% (w/v) tryptone, 1% (w/v) yeast extract, 0.5% (w/v) NaCl, 100 mM glucose, 100 µg/mL ampicillin, 20 µg/mL tetracycline, 1.5% (w/v) agar-agar) and incubated overnight at 37 °C. In parallel 10⁻⁶ dilutions were plated out to determine the transformation rate. These plates were used also for quality control by colony PCR using the oligonucleotide primers MHLacZ-Pro.f and MHgIII.r (Table 1) to verify the insert size. The colonies were scraped from the 25 cm plates using 40 mL 2xYT medium (1.6% (w/v) tryptone, 1% (w/v) yeast extract, 0.5% (w/v) NaCl) and 5 mL were directly used for midi plasmid preparation using the Nucleobond Plasmid Midi Kit (Macherey-Nagel).

Table 1

Oligonucleotide primers used for first and second PCR of antibody genes. Restriction sites are underlined.

Primer	5' to 3' sequence
<i>First antibody gene PCR VH</i>	
MHVH1.f	cag gtb cag ctg gtc cag tct gg
MHVH1/7.f	car rts cag ctg gtr car tct gg
MHVH2.f	cag rtc acc ttg aag gag tct gg
MHVH3.f1	sar gtc cag ctg gtc gag tct gg
MHVH3.f2	gag gtc cag ctg ktc gag wcy sg
MHVH4.f1	cag gtc car ctg cag gag tgc gg
MHVH4.f2	cag stg cag ctr cag sag tss gg
MHVH5.f	gar gtc cag ctg gtc cag tct gg
MHVH6.f	cag gta cag ctg cag cag tca gg
MHlgMCH1.r	aag ggt tgg ggc gga tgc act
<i>First antibody gene PCR kappa</i>	
MHVK1.f1	gac atc cag atg acc cag tct cc
MHVK1.f2	gmc atc crg wtg acc cag tct cc
MHVK2.f	gat rtt gtc atg acy cag wct cc
MHVK3.f	gaa atw gtc wtg acr cag tct cc
MHVK4.f	gac atc gtc atg acc cag tct cc
MHVK5.f	gaa acg aca ctc acg cag tct cc
MHVK6.f	gaw rtt gtc mtg acw cag tct cc
MHkappaCL.r	aca ctc tcc cct gtt gaa gct ctt
<i>First antibody gene PCR lambda</i>	
MHVL1.f1	cag tct gtc ctg act cag cca cc
MHVL1.f2	cag tct gtc ytg acg cag ccg cc
MHVL2.f	cag tct gcc ctg act cag cct
MHVL3.f1	tcc tat gwg ctg acw cag cca cc
MHVL3.f2	tct tct gag ctg act cag gac cc
MHVL4.f1	ctg cct gtc ctg act cag ccc
MHVL4.f2	cag cyt gtc ctg act caa tcr yc
MHVL5.f	cag sct gtc ctg act cag cc
MHVL6.f	aat ttt atg ctg act cag ccc ca
MHVL7/8.f	cag rct gtc gtc acy cag gag cc
MHVL9/10.f	cag scw gkg ctg act cag cca cc
MHlambdaCL.r (9 parts)	tga aca ttc tgt agg ggc cac tg
MHlambdaCL.r2 (1 part)	tga aca ttc cgt agg ggc aac tg
<i>Second antibody gene PCR VH</i>	
MHVH1-NcoI.f	gtctctcgca <u>cc</u> atg gcc cag gtb cag ctg gtc cag tct gg
MHVH1/7-NcoI.f	gtctctcgca <u>cc</u> atg gcc car rts cag ctg gtr car tct gg
MHVH2-NcoI.f	gtctctcgca <u>cc</u> atg gcc cag rtc acc ttg aag gag tct gg
MHVH3-NcoI.f1	gtctctcgca <u>cc</u> atg gcc sar gtc cag ctg gtc gag tct gg
MHVH3-NcoI.f2	gtctctcgca <u>cc</u> atg gcc gag gtc cag ctg ktc gag wcy sg
MHVH4-NcoI.f1	gtctctcgca <u>cc</u> atg gcc cag gtc car ctg cag gag tgc gg
MHVH4-NcoI.f2	gtctctcgca <u>cc</u> atg gcc cag stg cag ctr cag sag tss gg
MHVH5-NcoI.f	gtctctcgca <u>cc</u> atg gcc gar gtc cag ctg gtc cag tct gg
MHVH6-NcoI.f	gtctctcgca <u>cc</u> atg gcc cag gta cag ctg cag tca gg
MHlgMCH1scFv-HindIII.r	gtctctcgca <u>aag</u> ctt tgg ggc gga tgc act
<i>Second antibody gene PCR kappa</i>	
MHVK1-MluI.f1	accgctctcc a <u>cgc</u> gta gac atc cag atg acc cag tct cc
MHVK1-MluI.f2	accgctctcc a <u>cgc</u> gta gmc atc crg wtg acc cag tct cc
MHVK2-MluI.f	accgctctcc a <u>cgc</u> gta gat rtt gtc atg acy cag wct cc
MHVK3-MluI.f	accgctctcc a <u>cgc</u> gta gaa atw gtc wtg acr cag tct cc
MHVK4-MluI.f	accgctctcc a <u>cgc</u> gta gac atc gtc atg acc cag tct cc
MHVK5-MluI.f	accgctctcc a <u>cgc</u> gta gaa acg aca ctc acg cag tct cc
MHVK6-MluI.f	accgctctcc a <u>cgc</u> gta gaw rtt gtc mtg acw cag tct cc
MHkappaCLscFv-NotI.r	accgctctcc <u>gc ggc</u> cgc gaa gac aga tgg tgc agc cac agt
<i>Second antibody gene PCR lambda</i>	
MHVL1-MluI.f1	accgctctcc a <u>cgc</u> gta cag tct gtc ctg act cag cca cc
MHVL1-MluI.f2	accgctctcc a <u>cgc</u> gta cag tct gtc ytg acg cag ccg cc
MHVL2-MluI.f	accgctctcc a <u>cgc</u> gta cag tct gcc ctg act cag cct
MHVL3-MluI.f1	accgctctcc a <u>cgc</u> gta tcc tat gwg ctg acw cag cca cc
MHVL3-MluI.f2	accgctctcc a <u>cgc</u> gta tct tct gtc ctg act cag gac cc
MHVL4-MluI.f1	accgctctcc a <u>cgc</u> gta ctg cct gtc ctg act cag ccc
MHVL4-MluI.f2	accgctctcc a <u>cgc</u> gta cag cyt gtc ctg act caa tcr yc
MHVL5-MluI.f	accgctctcc a <u>cgc</u> gta cag sct gtc ctg act cag cc
MHVL6-MluI.f	accgctctcc a <u>cgc</u> gta aat ttt atg ctg act cag ccc ca
MHVL7/8-MluI.f	accgctctcc a <u>cgc</u> gta cag rct gtc gtc acy cag gag cc
MHVL9/10-MluI.f	accgctctcc a <u>cgc</u> gta cag scw gkg ctg act cag cca cc
MHlambdaCLscFv-NotI.r	accgctctcc <u>gc ggc</u> cgc aga gga sgg ygg gaa cag agt gac
<i>Primer for colony PCR and sequencing</i>	
MHLacZPro.f	ggctctgatgtgtgtgtg
MHglIII.r	ctaaagttttgtctgtcttcc

For VH cloning 5 µg pHAL14-VL sublibraries and 2 µg VH PCR products were digested in a volume of 100 µL using 30 U *Nco*II and 100 U *Hind*III (NEB) at 37 °C for 2 h to 18 h. This step was performed several times to produce enough material for 80–120 transformations. The digestions, ligations and transformations were performed as described for VL cloning with following modifications. The ligation was performed with 250 ng VH PCR product. After transformation, overnight incubation on 25 cm 2xYT-GAT plates and resuspension in 40 mL 2xYT medium, glycerol stocks of each library were made using 800 µL bacteria solution (~10¹⁰ bacteria) and 200 µL glycerol (Roth, Karlsruhe) and stored at –80 °C. For the final library one glycerol stock of each sublibrary was thawed, mixed together, aliquoted (1 mL) and stored at –80 °C.

The antibody gene libraries HAL4/7/8 were packaged separately inoculating 400 mL 2xYT-GA (2xYT containing 100 mg/mL ampicillin and 100 mM glucose) with a library glycerol stock. The bacteria were grown to an optical density at 600 nm (OD 600) of 0.4–0.5 at 37 °C and 250 rpm. 25 mL bacterial culture (~1.25 × 10¹⁰ cells) were infected with 2.5 × 10¹¹ Hyperphage particles (Rondot et al., 2001), incubated at 37 °C for 30 min without shaking followed by 30 min at 250 rpm. The infected cells were harvested by centrifugation for 10 min at 3220 × g. The pellet was resuspended in 400 mL 2xYT-AK (2xYT containing 100 mg/mL ampicillin and 50 mg/mL kanamycin). The phage were produced at 30 °C and 250 rpm over night. The bacteria were centrifuged for 20 min at 10,000 × g. Phage particles in the supernatant were precipitated with 1/5 volume of 20% (w/v) polyethylene glycol (PEG)/2.5 M NaCl solution for one hour on ice with gentle shaking and pelleted by centrifugation for one hour at 10,000 × g at 4 °C. The precipitated phage were resuspended in 10 mL phage dilution buffer (10 mM TrisHCl pH7.5, 20 mM NaCl, 2 mM EDTA), 1/5 volume of PEG solution was added and incubated on ice for 20 min, followed by centrifugation for 30 min at 10,000 × g and 4 °C. The pellet was resuspended in 1 mL phage dilution buffer. Residual bacteria and cell debris were removed by additional centrifugation for 5 min at 16,000 × g at 20 °C. The supernatants containing the antibody phage were stored at 4 °C. Phage titration was done as described before (Hust et al., 2007a). The scFv display rate of the packaged libraries was controlled by 10% SDS-PAGE, Western-blot and anti-pIII immunostain (mouse anti-pIII 1:2000, goat anti-mouse IgG AP conjugate 1:10,000). Wildtype pIII has a calculated molecular mass of 42.5 kDa, but it runs at an apparent molecular mass of 65 kDa in SDS-PAGE (Goldsmith and Konigsberg, 1977). Accordingly, the scFv::pIII fusion protein runs at about 95 kDa.

2.3. Panning

The pannings were mainly performed in 96 well microtitre plates (Maxisorb, Nunc, Wiesbaden, Germany). The antigens (10 µg or less) were coated in phosphate buffered saline (PBS) pH7.4 (Sambrook and Russell, 2001) or bicarbonate buffer (50 mM NaHCO₃) pH9.6 overnight at 4 °C. The concentrations were reduced each panning round (down to 100 ng/well). The antigen wells and wells for the preincubation of the library were blocked with MPBST (2% skim milk in PBS with 0.1% Tween 20). 2.5 × 10¹¹ phage particles of each library (e.g. HAL7/8) in PBST with 1% skim milk and 1% bovine serum albumin (BSA) were preincubated for 1 h. The supernatant, containing the depleted library, was incubated in the wells with the immobilised antigen at RT for 2 h followed by 10 washing steps with PBST. Afterwards the bound scFv phage particles were eluted with 200 µL trypsin solution (10 µg/mL trypsin in PBS) at 37 °C for 30 min. The supernatant containing the eluted scFv phage was transferred into a new tube. 10 µL of eluted scFv phage were used for titration as described before (Hust et al., 2007a). The remaining scFv phage were reamplified as described before (Hust et al., 2007a) and used for the next panning round. The

second and third panning round using the amplified phage were performed with the following modifications: the amount of antigen was reduced and washing cycles during panning were increased to 20, respectively 30.

2.4. Screening

For the identification of monoclonal binders, colonies from the titre plates of the eluted phage particles of the second and third panning round were picked and soluble scFvs were produced in microtitre plates as described before (Hust et al., 2009). The supernatants were directly used for an antigen ELISA as described before (Schütte et al., 2009). The monoclonal binders were sequenced using the oligonucleotide primer MHLacZ-Pro.f and the sequence was analysed using VBASE2 (www.vbase2.org, Fab/scFab/scFv analysis) (Mollova et al., 2010) to identify unique binders.

2.5. Application of recombinant antibody fragments

The isolated scFv antibody fragments and engineered variants were used in different applications. Antigen ELISA was performed as described above. SDS-PAGE followed by Immunoblot or

Coomassie staining were performed as described before (Kirsch et al., 2005). Flow cytometry was done as described (Menzel et al., 2008).

3. Results

3.1. Construction of the antibody gene libraries HAL 4/7/8

The first step in the generation of an universal antibody selection pipeline (Fig. 1) is the generation of a large antibody gene library using a phagemid which allows phage display, selection and screening without any subcloning steps. For the construction of the human naive antibody gene libraries HAL4/7/8, lymphocytes were isolated from the blood of 44 donors of Caucasian, African, Indian and Chinese origin. Total RNA and mRNA were isolated from the isolated lymphocytes and reverse transcribed to cDNA using random hexamer oligonucleotide primers.

The antibody gene repertoire was amplified in two steps. In the first step, VH and the full length light chains were amplified using a newly designed set of forward primers and IgM, kappa and lambda reverse primers (Table 1). These primer sets allow the amplification of all seven VH, all six kappa and 11 lambda subfamilies. In the

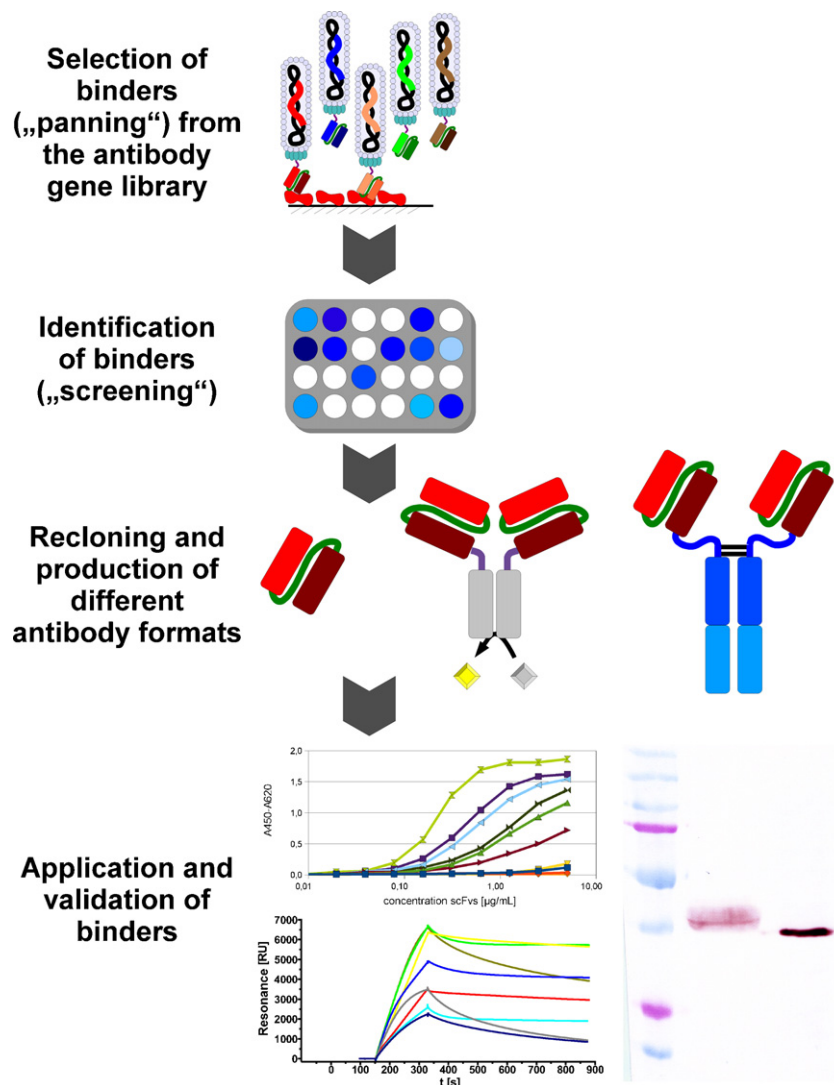


Fig. 1. Schematic overview of the antibody selection pipeline. At the bottom, typical validation experiments are shown, e.g. titration ELISA (comparison of different scFvs; left, top), surface plasmon resonance (determination of the affinity of different scFvs; left, bottom) or immunoblot (detection of recombinant and native *Salmonella typhimurium* OmpD using a scFv-phoA fusion; right).

second step, the restriction sites were added for the cloning into the pHAL14 vector. This phagemid has been used successfully for the generation of immune antibody gene libraries as well (Pelat et al., 2007, 2009; Schütte et al., 2009). The resulting libraries HAL4, 7 and 8 contain the same VH repertoire but differ in their light chain repertoire. For HAL4 only the kappa subfamilies 1, 3 and 4 are included, allowing the purification of these scFvs with protein L (Nilson et al., 1992). HAL7 includes all lambda subfamilies, HAL8 all kappa subfamilies with the exception of the pseudogen encoding VK7.

The cloning was done in two steps. First, the light chain repertoire was cloned between *MluI* and *NotI*. Here, libraries of average sizes between 2 and 8×10^8 independent clones were created. The heavy chain repertoire was cloned in the phagemid containing the light chain repertoire between *NcoI* and *HindIII*. Here, 69 (HAL4), 84 (HAL7) or 101 (HAL8) VH transformations were done. The final library sizes, determined as the number of independent clones of the VH transformations, was 2.2×10^9 for HAL4 (kappa), 2.8×10^9 for HAL7 (lambda), 2.4×10^9 for HAL8 (kappa).

The library quality was initially assessed by colony PCR, revealing 86–89% full size inserts. Sequencing of 265 randomly selected single clones from HAL7, respectively 279 from HAL8, revealed that 75%, respectively 65% of the clones contained VH and VL. In

the incorrect clones, mainly VH was missing and often replaced by other human gene fragments of equal size. The abundance of different VH, kappa and lambda subfamily genes and gene combinations is given for HAL7 (Table 3A) and HAL8 (Table 3B). Here, VH1 and VH3 are the dominating VH subfamilies, V λ 1–3 and 6, or V κ 1, 3 and 4 are most abundant light chain families. Unfortunately, a small fraction of HAL8 clones contained murine light chains derived from the antibody 215 (Schmiedl et al., 2000). These were not included in the analysis and, as expected, this sequence was not found in any clone after panning.

For panning, the library was packaged using Hyperphage (Rondot et al., 2001; Soltes et al., 2007). The quality of each packaging was analysed by immunoblot, where a strong antibody–pIII fusion band has to be apparent (data not shown). To investigate the bias of the *E. coli* phage display system on the abundance of different VH subfamilies, 275 Hyperphage packaged clones of HAL7, respectively 280 clones of HAL8, were sequenced (Table 3C and D) before doing any selections on antigen. 78% of the Hyperphage packed HAL7, respectively 67% of the packaged HAL8, contained VH and VL. Here, as for the initial libraries, the subfamilies heavy chain subfamilies VH1 and VH3, the lambda subfamilies V λ 1–3 and 6 and the kappa subfamilies V κ 1, 3 and 4 were dominating.

Table 2

Overview of the germline gene combinations found in (A) the HAL7 library, (B) the HAL8 library, (C) the Hyperphage packaged HAL7 library, (D) the Hyperphage packaged HAL8 library and (E) the selected scFvs from HAL4/7/8. The percentual amount of germline genes is given separately for VH, kappa and lambda.

A	HAL7 (LAMBDA) ANTIBODY GENE LIBRARY								
	VH1	VH2	VH3	VH4	VH5	VH6	VH7	SUM LIGHT CHAIN	% LIGHT CHAIN
V λ 1	17	0	27	8	1	3	0	56	28
V λ 2	9	0	26	3	2	2	0	42	21
V λ 3	17	0	18	2	6	3	0	46	23
V λ 4	0	0	3	2	0	0	0	5	3
V λ 5	0	0	0	0	0	0	0	0	0
V λ 6	8	0	13	6	0	4	0	31	16
V λ 7	4	0	3	2	1	0	0	10	5
V λ 8	1	0	3	0	0	0	0	4	2
V λ 9	0	0	0	0	0	0	0	0	0
V λ 10	2	0	2	2	0	0	0	6	3
SUM VH	58	0	95	25	10	12	0	200	
% VH	29	0	48	13	5	6	0		

B	HAL8 (KAPPA) ANTIBODY GENE LIBRARY								
	VH1	VH2	VH3	VH4	VH5	VH6	VH7	SUM LIGHT CHAIN	% LIGHT CHAIN
V κ 1	15	5	23	5	2	1	0	51	33
V κ 2	3	1	5	0	1	0	0	10	6
V κ 3	13	5	30	9	3	0	0	60	39
V κ 4	5	4	14	3	2	0	0	28	18
V κ 5	0	0	4	0	1	0	0	5	3
V κ 6	0	0	0	0	0	0	0	0	0
SUM VH	36	15	76	17	9	1	0	154	
% VH	23	10	49	11	6	1	0		

Table 2 (Continued).

C

HYPERPHAGE PACKAGED HAL7 (LAMBDA) ANTIBODY GENE LIBRARY									
	VH1	VH2	VH3	VH4	VH5	VH6	VH7	SUM LIGHT CHAIN	% LIGHT CHAIN
V λ 1	20	0	27	6	3	0	0	56	26
V λ 2	5	0	19	4	1	3	0	32	15
V λ 3	21	0	19	15	6	0	0	61	29
V λ 4	3	0	4	1	0	0	0	8	4
V λ 5	0	0	1	0	0	0	0	1	<1
V λ 6	12	0	11	4	2	3	0	32	15
V λ 7	2	0	2	2	0	0	0	6	3
V λ 8	1	0	4	3	0	0	0	8	4
V λ 9	0	0	1	0	0	0	0	1	<1
V λ 10	3	0	3	2	0	0	0	8	4
SUM VH	67	0	91	37	12	6	0	213	
% VH	31	0	43	17	6	3	0		

D

HYPERPHAGE PACKAGED HAL8 (KAPPA) ANTIBODY GENE LIBRARY									
	VH1	VH2	VH3	VH4	VH5	VH6	VH7	SUM LIGHT CHAIN	% LIGHT CHAIN
V κ 1	19	2	29	7	5	0	1	63	38
V κ 2	4	1	3	2	2	1	0	13	8
V κ 3	21	2	32	2	3	0	0	60	36
V κ 4	9	3	14	3	0	0	0	29	16
V κ 5	0	0	0	1	0	0	0	1	1
V κ 6	0	0	1	0	0	0	0	1	1
SUM VH	53	8	79	15	10	1	1	167	
% VH	32	5	47	9	6	1	1		

E

SELECTED SCFVS FROM HAL4/7/8									
	VH1	VH2	VH3	VH4	VH5	VH6	VH7	SUM LIGHT CHAIN	% LIGHT CHAIN
V κ 1	9	0	16	0	0	1	0	26	68
V κ 2	1	0	3	0	0	0	0	4	11
V κ 3	2	0	4	0	0	0	0	6	16
V κ 4	1	0	1	0	0	0	0	2	5
V κ 5	0	0	0	0	0	0	0	0	0
V κ 6	0	0	0	0	0	0	0	0	0
V λ 1	46	0	76	6	10	2	1	141	39
V λ 2	10	0	54	0	8	3	0	75	21
V λ 3	24	0	37	7	4	1	0	73	20
V λ 4	0	0	0	0	0	0	0	0	0
V λ 5	0	0	0	0	0	0	0	0	0
V λ 6	21	0	20	2	1	3	1	48	13
V λ 7	6	0	9	0	0	0	0	15	4
V λ 8	2	0	2	0	1	0	0	5	1
V λ 9	0	0	0	0	0	0	0	0	0
V λ 10	3	0	2	0	0	0	0	5	1
SUM VH	125	0	224	15	24	10	2	400	
% VH	31	0	56	4	6	3	< 1		

3.2. Selection and screening of antibodies

All steps in the selection procedure were analysed and if necessary, modified to establish a standardised antibody generation pipeline. The panning was mainly performed on directly immobilised antigen in microtitre plates (MTP). The amount of antigen was reduced in each subsequent panning round to increase the stringency. To remove unspecific binders, the library was preincubated in MTP wells containing skim milk powder and BSA. When fusion proteins were used for panning, non-relevant fusion proteins were used for preincubation and sometimes competition during panning. For panning, a phage particle number ensuring at least 100 copies of each antibody gene were used to ensure the utilization of the entire library diversity in every panning. All washing steps were performed in automated ELISA washers to achieve reproducible washing conditions. The number of washing steps were increased which each panning round to enhance stringency. The phage display vector pHAL14 encodes a trypsin cleavage site between the scFv fragment and pIII, allowing to elute binders by trypsin digestion. This elution method has shown to be more efficient than the elution by pH shift or direct infection of *E. coli*, in particular when using polyvalent antibody display. For reamplification after the first round of panning, packaging of the eluted antibody phage particles was performed with M13K07 to switch to monovalent display and therefore to select for antibody fragments with a high monovalent affinity. In general, two to three panning rounds were sufficient to generate a panel of primary hits.

The pHAL14 vector is designed to allow both phage display and the production of soluble antibody fragments in MTPs without the necessity of recloning. Usually, for the identification of primary hits, 92 *E. coli* clones infected with trypsin eluted phage were picked from the second or third panning round and cultivated in 96 well plates. For functional analysis of individual scFv, monoclonal production was performed in 96 well MTPs overnight (Hust et al., 2009). The supernatant was directly used for an antigen ELISA to identify monoclonal binders. Individual binders, limited to the best ones in the initial ELISA, were subjected to DNA sequencing (GATC, Konstanz, Germany) to identify unique antibody clones and analyze the V gene families using the database VBASE2 (www.vbase2.org) which is adapted to standardised analysis of scFv gene sequences (Mollova et al., 2010). The analysis process was narrowed to the binders with the best signal to noise ratio in order to streamline the selection procedure.

3.3. Validation of the antibody gene libraries

The HAL antibody gene libraries were successfully validated with 91 targets. These targets included human, other mammalian, fungi and bacterial proteins as well as viruses and haptens. All antigens resulted in the generation of binders. Table 2 shows the numbers of validated unique scFvs obtained. The antibodies were validated by ELISA, titrations ELISA, immunoblot, immunohistochemistry (IHC), Biacore, flow cytometry, microarray, immunoprecipitation or neutralisation depending on the target protein. Minimum two methods were used for each validated antibody. Since only a fraction of the identified different clones could be analysed in detail, the number of individual scFvs isolated from the library was much higher. Although one kappa library (HAL4 or HAL8) was combined with the lambda library (HAL7) for most pannings, most isolated binders originated from the lambda library. When performing a panning using only a kappa library, kappa binders were also isolated in cases where no kappa binders were initially selected from the combined libraries.

Table 3

Panning results on 81 different antigens. Validated scFvs are defined as antibodies which were functional in at least one secondary assay.

Origin of antigen	Number of different antigens	Number of unique, validated binders
Human	72	297
Other mammalia	11	30
Fungi	2	10
Bacteria	15	63
Virus	6	27
Haptens	4	8
Total	110	435

The subfamily combination used in the selected scFvs was evaluated for the validated binders (Table 3E). Most of the VHs were identified to originate from VH1 or VH3 subfamilies. The lambda light chains mainly originated from VL1, VL2, VL3 and VL6 subfamilies. For kappa, the Vk1 subfamily is dominating (Fig. 2). Several scFvs contained variable domains encoded by pseudogenes.

3.4. Enhancing scFv properties by one step cloning into additional expression vectors

The application of recombinant scFv antibodies can be facilitated by fusion with tags, enzymes or multimerisation domains as well as by changing the expression system. For this purpose, a set of *E. coli* and mammalian expression vectors, allowing one step cassette subcloning from the phage display libraries, but also from other single chain antibody vectors carrying the appropriate *NcoI* and *NotI* sites, were constructed (Fig. 3).

The vector pOPE101-XP (Hust et al., 2009), derived from the vector pOPE101-215(Yol) (Schmiedl et al., 2000), allows an improved scFv production in *E. coli*. The vector pOPE107-XP contains 6xhis, c-myc and an additional 6xhis tag, allowing improved affinity purification by immobilised metal affinity chromatography. The vector pHT121, derived from pOPE101-XP and phagemids containing the biotin acceptor (BAD) sequence, allows the production of biotinylated scFvs in *E. coli* strains expressing biotin ligase (BirA) (Thie et al., 2009). The fusion to the alkaline phosphatase (phoA) gene in pOPE201-XP allows a dimerisation of scFvs and a direct detection without secondary antibodies, e.g. in immunoblots.

The mammalian expression vector pCMX2.5-hlgG1-Fc-XP allows the fusion of scFv antibodies to the gene fragment encoding the human IgG1 Fc part (hinge-CH2-CH3) and allows the production of homodimeric scFv-Fc fusion proteins in HEK293T cells. This vector is derived from pCMV vectors described before (Kirsch et al., 2008; Schütte et al., 2009) but allows higher secretion levels. Moreover, additional pCMX vectors containing alternative Fc moieties like mouse IgG1Fc or IgG2a Fc are also available which can be used for sandwich ELISA applications.

3.5. Validation and test of application

More than 300 unique scFvs were isolated from the HAL4/7/8 libraries, validated and many of them were further characterised and used in a variety of applications. As an example, the isolation of several unique scFv antibody clones against human transferrin receptor (CD71) is shown. Recombinant human transferrin receptor (extracellular domain) was used as a panning antigen. Selected monoclonal scFv antibodies were produced in *E. coli* XL1-Blue-MRF' for further analysis (Fig. 4). SDS-PAGE and immunoblot analysis of periplasmic preparation confirmed the individual molecular mass of the scFvs and already revealed differences in productivity (Fig. 4A). The scFvs were purified by immobilised metal affinity chromatography (IMAC) (Fig. 4A). Antigen ELISA with serial dilu-

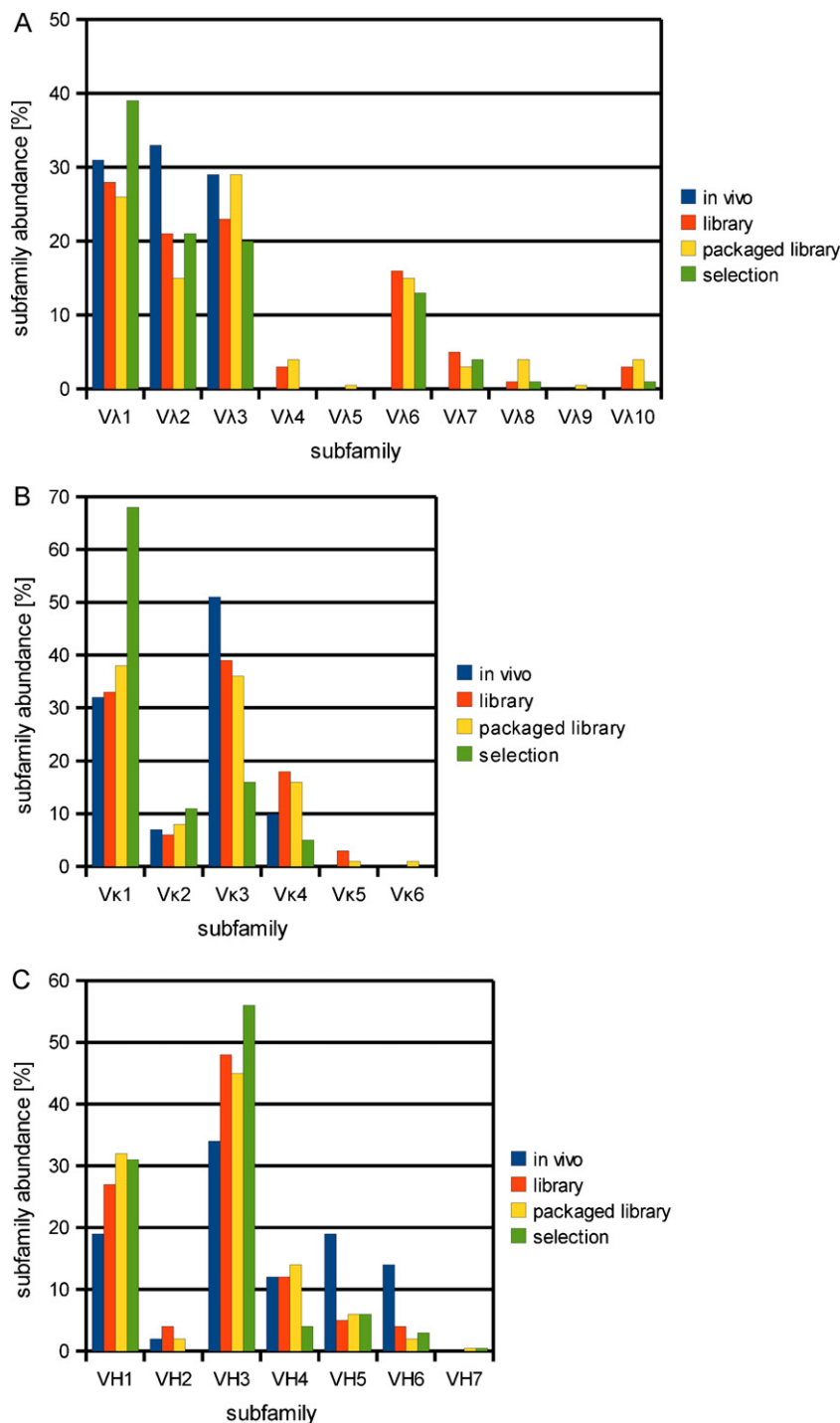


Fig. 2. Representation of antibody subfamilies found *in vivo* (unknown for subfamily VH7, VL4–10 comprises in total 8%) (Knappik et al., 2000), subfamilies in the HAL7/8 library, subfamilies after Hyperphage packaging of the HAL7/8 libraries and subfamilies of the selected scFvs from the HAL4/7/8 libraries. (A) Abundance of kappa subfamilies, (B) abundance of lambda subfamilies and (C) abundance of VH subfamilies.

tions of scFv confirmed binding to recombinant CD71 immobilized on plastic (Fig. 4B). Binding to CD71 on living cells was tested by flow cytometry (Fig. 4C). Here, all tested scFvs bound to CD71 positive cells but not to CD71 negative controls. The reactivity in ELISA was not correlated to activity on living cells, once again underlining the limited value of ELISA for selection of the best scFv antibodies. For example, the scFv SH83-D8 showed lowest cell binding in flow cytometry but was the best binder in antigen ELISA. The scFv SH83-A7 was further tested for internalisation via CD71 by fluorescence microscopy (Fig. 4D). Incubation of CD71 positive cells with this

scFv at 4 °C resulted in surface staining whereas incubation at 37 °C revealed intracellular uptake.

4. Discussion

Various methods have been applied to clone the genetic diversity of antibody repertoires. Naive or immune antibody gene libraries are mostly constructed by one, two or three separate cloning steps. In the “one cloning step” method VH and VL are

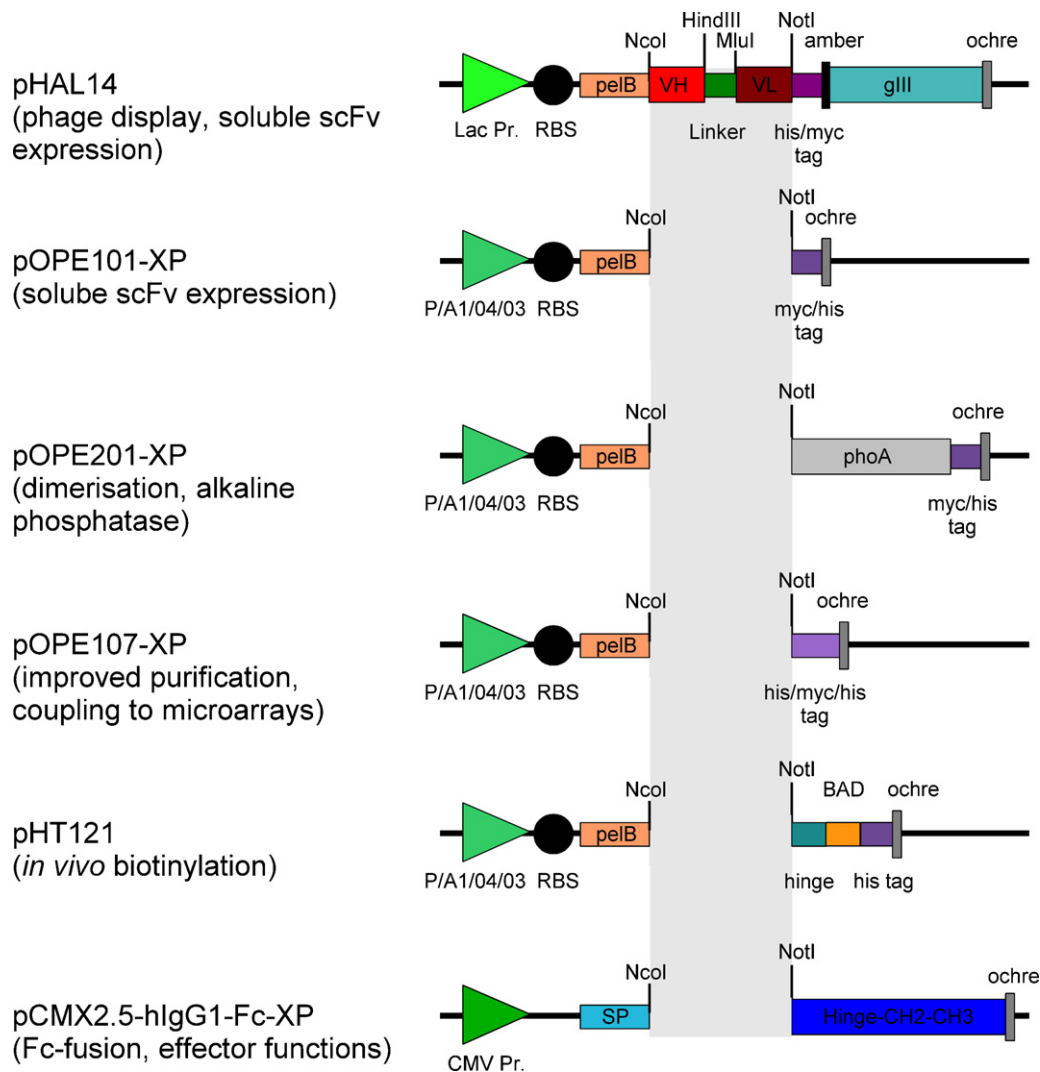


Fig. 3. *E. coli* and mammalian cassette expression vectors for scFv subcloning.

combined by assembly PCR and the PCR product is cloned into the vector (Clackson et al., 1991; McCafferty et al., 1994; Vaughan et al., 1996). In the “two step cloning” strategy, the amplified repertoire of light chain genes is cloned into the phage display vector first. In the second step the heavy chain gene repertoire – as the heavy chain contributes more to diversity due to its highly variable CDRH3 – is cloned into the phagemids containing the light chain gene repertoire (Johansen et al., 1995; Little et al., 1999; Løset et al., 2005; Welschhof et al., 1997). For the “three step cloning strategy”, separate heavy and light chain libraries are engineered, then the VH gene repertoire is excised and cloned into the phage display vector containing the repertoire of VL genes (de Haard et al., 1999). The “two step cloning” strategy was applied for the generation of HAL4/7/8, because this approach minimises the steps which may lead to loss of diversity. The scFv antibody format was chosen because of its better display on phage, due to better *E. coli* expression (Hust et al., 2007b). Moreover, scFvs can easier be engineered compared to Fabs because the latter usually require two cloning steps to construct other formats.

The construction of large naive and semi-synthetic libraries (Glanville et al., 2009; Hoet et al., 2005; Little et al., 1999; Løset et al., 2005; Schofield et al., 2007; Sheets et al., 1998; Vaughan et al., 1996) requires significant effort to tunnel the genetic diversity

through the bottleneck of *E. coli* transformation, e.g. 600 transformations were described to be necessary for a 3.5×10^{10} Fab phage library (Hoet et al., 2005). In this work, about 250 transformations resulted in three libraries with a total size of about 7.4×10^9 independent clones. The abundance of gene families in the sequenced libraries HAL7/8 was compared to the distribution of subfamilies found *in vivo* (Knappik et al., 2000). For VH, the subfamilies VH1 and VH3 were over-represented, whereas the VH5 and VH6 were under-represented in the libraries. Interestingly, VH2 was found only in the kappa library, but not in the lambda library. For kappa, Vκ3 is under-represented and Vκ4 is over-represented. The lambda subfamilies Vλ4–Vλ10 are rare *in vivo*, whereas these subfamilies, especially Vλ6, are included in the HAL7 library. Hence, the abundance of subfamilies in HAL7/8 are not congruent with the *in vivo* found subfamilies. This is not interpreted to be a disadvantage, since rare subfamilies provide different structural features and random combination to the other chains provide improved structural diversity. The V gene coverage (data not shown) of the HAL libraries is better balanced as in the “Pfizer” library which is lacking the most of the lambda V genes (Glanville et al., 2009).

The library was initially packaged with Hyperphage for oligovalent display to reduce the loss of specific binders in the most critical first panning round where nonspecific binders are always present

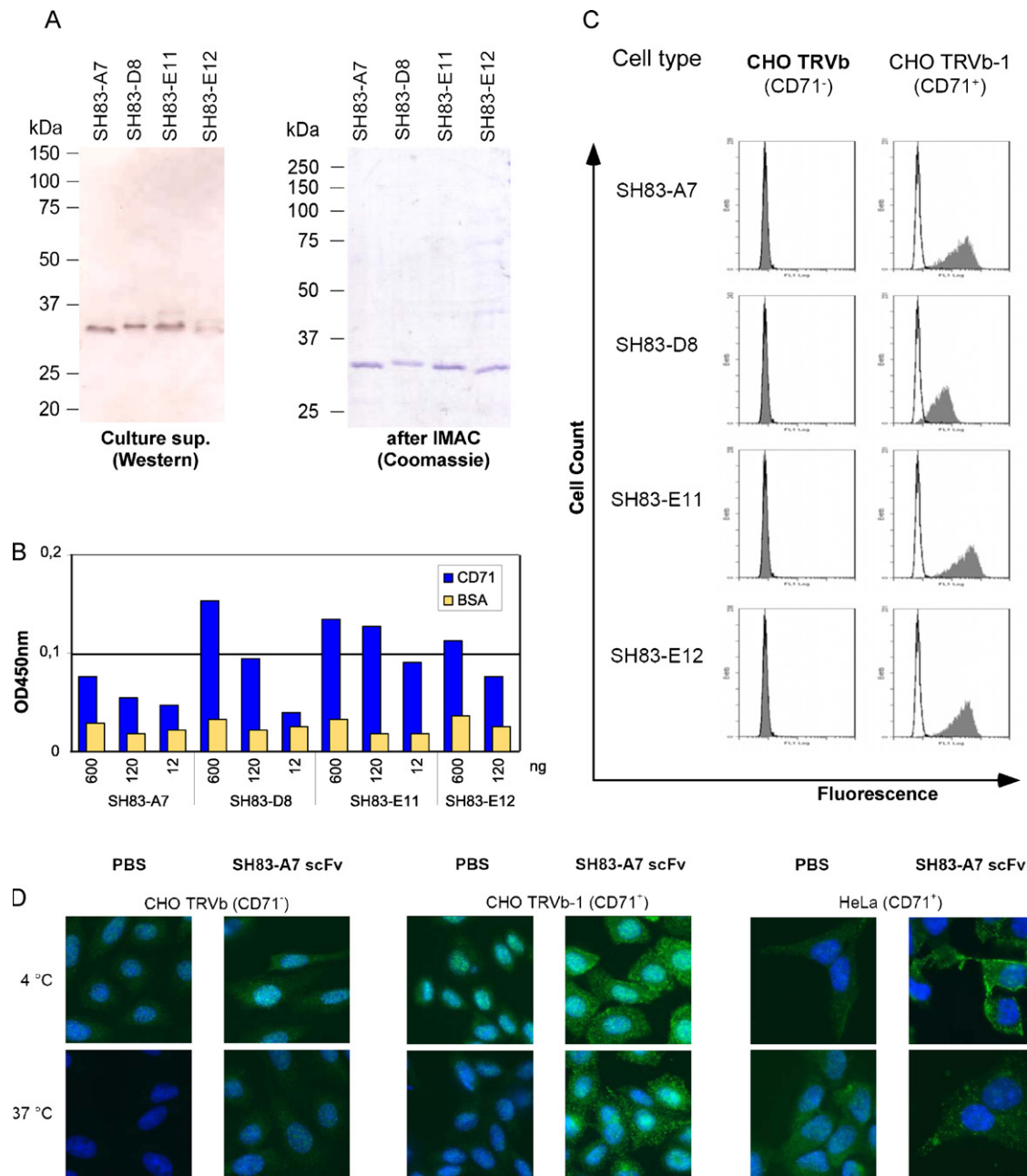


Fig. 4. CD71 specific scFvs selected from the universal library. ScFv clones (SH83-A7, -D8, -E11 and -E12) were isolated by panning against human CD71, produced in *E. coli* XL1-Blue-MRF' and purified by IMAC. (A) Periplasmic preparations of scFv productions were analyzed by SDS-PAGE (left) followed by Western blot and immunostaining with penta-his antibody and goat-anti-mouse IgG (Fc-specific) AP conjugate. Sample volumes were adjusted to production cell density to compare periplasmic yields. After scFv purification by IMAC, samples of elution fractions were analysed by SDS-PAGE and Coomassie staining (right). These scFvs were used in different tests: (B) antigen ELISA on human CD71 (blue) or BSA as control (red). Bound scFvs were detected by staining with mouse anti-penta-his IgG and goat anti-mouse IgG (Fab specific) HRP conjugate. (C) Specific cell binding of scFvs was tested by flow cytometry using CD71 negative (left) and CD71 positive (right) CHO-TRV cell variants (gray filled histograms). For negative control, a non-relevant scFv was used (black line). Bound scFvs were detected by mouse anti-hexa-his IgG and goat-anti-mouse IgG (Fc specific)-FITC conjugate. (D) Immunofluorescence analysis of CD71 negative and positive CHO-TRVB cells after incubation with scFv SH83-A7 at 4 °C and 37 °C to test cell binding and internalization. Incubation with PBS was used as negative control. Cells were fixed with formaldehyde and permeabilised with Triton X-100. After blocking cells were stained with a hexa-his tag specific antibody followed by goat anti-mouse IgG(H+L) Alexa488 conjugate. Nuclei were counterstained with DAPI.

in very large excess (Rondot et al., 2001; Soltes et al., 2007). In addition, Hyperphage enriches intact open reading frames (ORFs) during packaging (Hust et al., 2006; Kügler et al., 2008; Naseem et al., 2009), which may also help to enhance the quality of the antibody gene library. The sequencing of the Hyperphage packaged libraries showed no significant differences in the abundance of different subfamilies for VH and kappa compared to the cloned libraries. Whereas, for lambda Vλ3 and Vλ8 are higher represented

in the packaged library and Vλ2 is lower represented compared the cloned library. Hence, *E. coli* may lead to a bias in the distribution of antibody subfamilies due to expression or folding properties which varies between lambda subfamilies.

For the validation of the HAL4/7/8 libraries, scFvs against 110 antigens were selected and 435 antibodies were analysed in detail. In all cases specific monoclonal binders were successfully isolated. Interestingly, in average more binders were selected from

the lambda library (HAL7) when compared to the kappa libraries (HAL4/8), particularly when a lambda and kappa libraries were used in combination during the panning. This is in accordance with Løset et al. (2005), who described that lambda scFvs are expressed in higher yields in *E. coli* compared to kappa scFvs. This may have led to an advantage in phage display.

The variable domains of several functional scFvs are encoded by V pseudogene elements. Since the V genes resulting from open reading frames of somatically recombined V(D)J gene elements and isolated from mRNA or RNA, the pseudogene classification of these V gene elements is doubtful.

The representation of the antibody subfamilies within the group of selected scFvs is largely in accordance with Schofield et al. (2007). There, VH1 and VH3 were the dominating variable heavy chain domains which is matching the distribution *in vivo* where VH3 is the most abundant family followed by VH1, VH5, VH4, VH6, and VH2 (Knappik et al., 2000). In the VH library and the packaged library VH4 were often found compared to the amount of selected VH4 binders. No antigen specific scFv with VH from the VH2 subfamily has been isolated from any HAL library. For V λ , mainly the V λ 1 subfamily was isolated, followed by V λ 3, V λ 2 and V λ 6 which is similar to the *in vivo* distribution where V λ 1, V λ 2 and V λ 3 are dominating but V λ 6 is rare *in vivo*. Compared to that, Schofield et al. (2007) isolated mainly scFvs with V λ 6, followed by V λ 1, V λ 2 and V λ 3. No scFv with V λ 4 was selected despite the fact, that this subfamily was found in the cloned and the packaged library. The small number of isolated scFvs from the kappa antibody libraries does not allow a reliable and comprehensive evaluation of V κ distribution after panning. So far, V κ 1 was dominating among the scFv clones isolated from the HAL libraries which is comparable to the data of Schofield et al. (2007). Whereas in the kappa library and the packaged library V κ 3 was equal to V κ 1, the amount of selected binders with V κ 3 was decreased.

The data suggest some bias of the *E. coli* expression and folding system towards an increased abundance of some antibody subfamilies.

The generation of very large naive antibody gene libraries of high quality is the initial requirement for an antibody generation process with the goal to isolate binders against many antigens of different kinds. Here, all protocols for panning and screening were optimised for robustness and throughput. Standardized panning procedures with high success rates were considered to be more important than individually optimised protocols for each antigen. Therefore, all steps during antibody selection and screening were adjusted to microtitre scale, which made the process amenable to parallelisation and automation. The complete pipeline was applied for the SH2 pilot project initiated by the Structural Genomics Consortium. Here, more than binders to 20 different SH2 domains were generated in a short time (Mersmann et al., 2010; Pershad et al., 2010). In principle, the whole antibody generation process can be accelerated by introducing robotics.

To facilitate application of the scFv antibodies in research, a number of expression vectors were developed that allow one-step restriction subcloning. Here, the main goals were to improve the recombinant production and to extend properties of the antibody fragments in respect of final user applications. The vector pOPE101-XP (Hust et al., 2009) allows an improved scFv expression compared to the phage display vector pHAL14. The variant pOPE107-XP contains a double 6xhis tag which can improve the purification and the coupling of antibody microarray (Khan et al., 2006; Shukla et al., 2009). The vector pHT121 contains the BAD sequence, allowing the production of biotinylated scFvs in *E. coli* strains expressing biotin ligase (BirA). *In vivo* biotinylated scFvs can be further modified by formation of multimers with streptavidin also termed streptabodies (Thie et al., 2009). The fusion of the phoA gene in pOPE201-XP allows a dimerisation of scFvs and a direct use in e.g.

immunoblots without any further enzyme coupled secondary antibodies (Schofield et al., 2007). Moreover, a number of mammalian expression vectors were constructed for the fusion of scFvs with different immunoglobulin Fc moieties. ScFv–Fc fusion proteins have IgG-like properties like homodimeric structure including improved antigen binding by avidity effect and can potentially mediate effector functions like antibody dependent cellular cytotoxicity (ADCC) or complement dependent cytotoxicity (CDC) (Cao et al., 2009). The fusion of scFvs to Fc fragments of different species, e.g. human and mouse, allows the development of sandwich or competition ELISAs.

Acknowledgments

We would like to thank Dr. Henk S.P. Garritsen (Inst. for Clinical Transfusion Medicine, Städtisches Klinikum Braunschweig) for kindly providing blood samples and support. We gratefully acknowledge the financial support by the German ministry of education and research (BMBF, SMP “Antibody Factory” in the NGFN2 program), the German Research Foundation (DFG) Grant DU 337/3-1, the EU FP6 supported coordination action funded activities “Proteome Binders” (contract 026008) and the EU FP7 collaborative projects “Affinity Proteome” (contract 222635). The HAL libraries are dedicated to HAL9000 (2001: a space odyssey).

References

- Aebersold, R., Mann, M., 2003. Mass spectrometry-based proteomics. *Nature* 422, 198–207.
- Berglund, L., Björling, E., Oksvold, P., Fagerberg, L., Asplund, A., Szizyarto, C.A., Persson, A., Ottosson, J., Wernérus, H., Nilsson, P., Lundberg, E., Sivertsson, A., Navani, S., Wester, K., Kampf, C., Hober, S., Pontén, F., Uhlén, M., 2008. A gene-centric human protein atlas for expression profiles based on antibodies. *Mol. Cell Proteomics* 7, 2019–2027.
- Bödicke, T., 2007. Blocking translocation of cell surface molecules from the ER to the cell surface by intracellular antibodies targeted to the ER. *J. Cell. Mol. Med.* 11, 54–70.
- Cao, M., Cao, P., Yan, H., Lu, W., Ren, F., Hu, Y., Zhang, S., 2009. Construction, purification, and characterization of anti-BAFF scFv–Fc fusion antibody expressed in CHO/dhfr-cells. *Appl. Biochem. Biotechnol.* 157, 562–574.
- Clackson, T., Hoogenboom, H.R., Griffiths, A.D., Winter, G., 1991. Making antibody fragments using phage display libraries. *Nature* 352, 624–628.
- Deflorian, G., Cinquanta, M., Beretta, C., Venuto, A., Santoriello, C., Baldessari, D., Pezzimenti, F., Aliprandi, M., Mione, M., de Marco, A., 2009. Monoclonal antibodies isolated by large-scale screening are suitable for labeling adult zebrafish (*Danio rerio*) tissues and cell structures. *J. Immunol. Methods* 346, 9–17.
- De Masi, F., Chiarella, P., Wilhelm, H., Massimi, M., Bullard, B., Ansorge, W., Sawyer, A., 2005. High throughput production of mouse monoclonal antibodies using antigen microarrays. *Proteomics* 5, 4070–4081.
- Dübel, S., 2007. Recombinant therapeutic antibodies. *Appl. Microbiol. Biotechnol.* 74, 723–729.
- Dübel, S., Stoevesandt, O., Taussig, M.J., Hust, M., 2010. Generating recombinant antibodies to the complete human proteome. *Trends Biotechnol.* 28, 333–339.
- Glanville, J., Zhai, W., Berka, J., Telman, D., Huerta, G., Mehta, G.R., Ni, L., Mei, L., Sundar, P.D., Day, G.M.R., Cox, D., Rajpal, A., Pons, J., 2009. Precise determination of the diversity of a combinatorial antibody library gives insight into the human immunoglobulin repertoire. *Proc. Natl. Acad. Sci. U.S.A.* 106, 20216–20221.
- Goldsmith, M.E., Konigsberg, W.H., 1977. Adsorption protein of the bacteriophage fd: isolation, molecular properties, and location in the virus. *Biochemistry* 16, 2686–2694.
- de Haard, H.J., van Neer, N., Reurs, A., Hufton, S.E., Roovers, R.C., Henderikx, P., de Bruijn, A.P., Arends, J.W., Hoogenboom, H.R., 1999. A large non-immunized human Fab fragment phage library that permits rapid isolation and kinetic analysis of high affinity antibodies. *J. Biol. Chem.* 274, 18218–18230.
- Hagemeyer, C.E., von Zur Muhlen, C., von Elverfeldt, D., Peter, K., 2009. Single-chain antibodies as diagnostic tools and therapeutic agents. *Thromb. Haemost.* 101, 1012–1019.
- Hallborn, J., Carlsson, R., 2002. Automated screening procedure for high-throughput generation of antibody fragments. *Biotechniques Suppl.*, 30–37.
- Hanes, J., Plückthun, A., 1997. In vitro selection and evolution of functional proteins by using ribosome display. *Proc. Natl. Acad. Sci. U.S.A.* 94, 4937–4942.
- Harrison, P.M., Kumar, A., Lang, N., Snyder, M., Gerstein, M., 2002. A question of size: the eukaryotic proteome and the problems in defining it. *Nucleic Acids Res.* 30, 1083–1090.
- He, M., Taussig, M.J., 2008. Selection of recombinant antibodies by eukaryotic ribosome display. *Methods Mol. Biol.* 484, 193–205.
- Hoet, R.M., Cohen, E.H., Kent, R.B., Rookey, K., Schoonbroodt, S., Hogan, S., Rem, L., Frans, N., Daukandt, M., Pieters, H., van Heggelsom, R., Neer, N.C., Nastri, H.G., Rondon, I.J., Leeds, J.A., Hufton, S.E., Huang, L., Kashin, I., Devlin, M., Kuang,

- G., Steukers, M., Viswanathan, M., Nixon, A.E., Sexton, D.J., Hoogenboom, H.R., Ladner, R.C., 2005. Generation of high-affinity human antibodies by combining donor-derived and synthetic complementarity-determining-region diversity. *Nat. Biotechnol.* 23, 344–348.
- Hoogenboom, H.R., 2005. Selecting and screening recombinant antibody libraries. *Nat. Biotechnol.* 23, 1105–1116.
- Hust, M., Dübel, S., 2004. Mating antibody phage display with proteomics. *Trends Biotechnol.* 22, 8–14.
- Hust, M., Dübel, S., Schirrmann, T., 2007a. Selection of recombinant antibodies from antibody gene libraries. *Methods Mol. Biol.* 408, 243–255.
- Hust, M., Jostock, T., Menzel, C., Voedisch, B., Mohr, A., Brenneis, M., Kirsch, M.I., Meier, D., Dübel, S., 2007b. Single chain Fab (scFab) fragment. *BMC Biotechnol.* 7, 14.
- Hust, M., Meysing, M., Schirrmann, T., Selke, M., Meens, J., Gerlach, G., Dübel, S., 2006. Enrichment of open reading frames presented on bacteriophage M13 using hyperphage. *Biotechniques* 41, 335–342.
- Hust, M., Steinwand, M., Al-Halabi, L., Helmsing, S., Schirrmann, T., Dübel, S., 2009. Improved microtitre plate production of single chain Fv fragments in *Escherichia coli*. *New Biotechnol.* 25, 424–428.
- International Human Genome Sequencing Consortium, 2004. Finishing the euchromatic sequence of the human genome. *Nature* 431, 931–945.
- Johansen, L.K., Albrechtsen, B., Andersen, H.W., Engberg, J., 1995. pFab60: a new, efficient vector for expression of antibody Fab fragments displayed on phage. *Protein Eng.* 8, 1063–1067.
- Khan, F., He, M., Taussig, M.J., 2006. Double-hexahistidine tag with high-affinity binding for protein immobilization, purification, and detection on ni-nitrotriacetic acid surfaces. *Anal. Chem.* 78, 3072–3079.
- Kirsch, M., Zaman, M., Meier, D., Dübel, S., Hust, M., 2005. Parameters affecting the display of antibodies on phage. *J. Immunol. Methods* 301, 173–185.
- Kirsch, M., Hülseweh, B., Nacke, C., Rülker, T., Schirrmann, T., Marschall, H., Hust, M., Dübel, S., 2008. Development of human antibody fragments using antibody phage display for the detection and diagnosis of Venezuelan equine encephalitis virus (VEEV). *BMC Biotechnol.* 8, 66.
- Knappik, A., Ge, L., Honegger, A., Pack, P., Fischer, M., Wellenhofer, G., Hoess, A., Wölle, J., Plückthun, A., Virnekäs, B., 2000. Fully synthetic human combinatorial antibody libraries (HuCAL) based on modular consensus frameworks and CDRs randomized with trinucleotides. *J. Mol. Biol.* 296, 57–86.
- Konthur, Z., Hust, M., Dübel, S., 2005. Perspectives for systematic in vitro antibody generation. *Gene* 364, 19–29.
- Krebs, B., Rauchenberger, R., Reiffert, S., Rothe, C., Tesar, M., Thomassen, E., Cao, M., Dreier, T., Fischer, D., Höss, A., Inge, L., Knappik, A., Marget, M., Pack, P., Meng, X.Q., Schier, R., Söhleemann, P., Winter, J., Wölle, J., Kretzschmar, T., 2001. High-throughput generation and engineering of recombinant human antibodies. *J. Immunol. Methods* 254, 67–84.
- Kügler, J., Nieswandt, S., Gerlach, G.F., Meens, J., Schirrmann, T., Hust, M., 2008. Identification of immunogenic polypeptides from a *Mycoplasma hyopneumoniae* genome library by phage display. *Appl. Microbiol. Biotechnol.* 80, 447–458.
- Levy, S., Sutton, G., Ng, P.C., Feuk, L., Halpern, A.L., Walenz, B.P., Axelrod, N., Huang, J., Kirkness, E.F., Denisov, G., Lin, Y., MacDonald, J.R., Pang, A.W.C., Shago, M., Stockwell, T.B., Tsiamouri, A., Bafna, V., Bansal, V., Kravitz, S.A., Busam, D.A., Beeson, K.Y., McIntosh, T.C., Remington, K.A., Abril, J.F., Gill, J., Borman, J., Rogers, Y., Frazier, M.E., Scherer, S.W., Strausberg, R.L., Venter, J.C., 2007. The diploid genome sequence of an individual human. *PLoS Biol.* 5, e254.
- Little, M., Welschof, M., Braunagel, M., Hermes, I., Christ, C., Keller, A., Rohrbach, P., Kürschner, T., Schmidt, S., Kleist, C., Terness, P., 1999. Generation of a large complex antibody library from multiple donors. *J. Immunol. Methods* 231, 3–9.
- Løset, G.A., Løbersli, I., Kavlie, A., Stacy, J.E., Borgen, T., Kausmally, L., Hvattum, E., Simonsen, B., Hovda, M.B., Brekke, O.H., 2005. Construction, evaluation and refinement of a large human antibody phage library based on the IgD and IgM variable gene repertoire. *J. Immunol. Methods* 299, 47–62.
- McCafferty, J., Fitzgerald, K.J., Earnshaw, J., Chiswell, D.J., Link, J., Smith, R., Kenten, J., 1994. Selection and rapid purification of murine antibody fragments that bind a transition-state analog by phage display. *Appl. Biochem. Biotechnol.* 47, 157–171 (discussion 171–173).
- McGraw, T.E., Greenfield, L., Maxfield, F.R., 1987. Functional expression of the human transferrin receptor cDNA in Chinese hamster ovary cells deficient in endogenous transferrin receptor. *J. Cell Biol.* 105, 207–214.
- Menzel, C., Schirrmann, T., Konthur, Z., Jostock, T., Dübel, S., 2008. Human antibody RNase fusion protein targeting CD30+ lymphomas. *Blood* 111, 3830–3837.
- Mersmann, M., Meier, D., Mersmann, J., Helmsing, S., Nilsson, P., Gräslund, S., Structural Genomics Consortium, Colwill, K., Hust, M., Dübel, S., 2010. Towards proteome scale antibody selections using phage display. *New Biotechnol.* 27, 118–128.
- Mollova, S., Retter, I., Hust, M., Dübel, S., Müller, W., 2010. Analysis of single chain antibody sequences using the VBASE2 Fab analysis tool. In: *Antibody Engineering*, second ed. Springer Verlag, Heidelberg/New York, pp. 3–10.
- Naseem, S., Meens, J., Jores, J., Heller, M., Dübel, S., Hust, M., Gerlach, G., 2009. Phage display-based identification and potential diagnostic application of novel antigens from *Mycoplasma mycoides* subsp. *mycoides* small colony type. *Vet. Microbiol.*, <http://www.ncbi.nlm.nih.gov/pubmed/19900769>.
- Nieri, P., Donadio, E., Rossi, S., Adinolfi, B., Podestà, A., 2009. Antibodies for therapeutic uses and the evolution of biotechniques. *Curr. Med. Chem.* 16, 753–779.
- Nilson, B.H., Solomon, A., Björck, L., Akerström, B., 1992. Protein L from *Peptostreptococcus magnus* binds to the kappa light chain variable domain. *J. Biol. Chem.* 267, 2234–2239.
- Ohara, R., Knappik, A., Shimada, K., Frisch, C., Ylera, F., Koga, H., 2006. Antibodies for proteomic research: comparison of traditional immunization with recombinant antibody technology. *Proteomics* 6, 2638–2646.
- Pelat, T., Hust, M., Hale, M., Lefranc, M., Dübel, S., Thullier, P., 2009. Isolation of a human-like antibody fragment (scFv) that neutralizes ricin biological activity. *BMC Biotechnol.* 9, 60.
- Pelat, T., Hust, M., Laffly, E., Condemine, F., Bottex, C., Vidal, D., Lefranc, M., Dübel, S., Thullier, P., 2007. High-affinity, human antibody-like antibody fragment (single-chain variable fragment) neutralizing the lethal factor (LF) of *Bacillus anthracis* by inhibiting protective antigen–LF complex formation. *Antimicrob. Agents Chemother.* 51, 2758–2764.
- Pershad, K., Pavlovic, J.D., Gräslund, S., Nilsson, P., Colwill, K., Karatt-Vellatt, A., Schofield, D.J., Dyson, M.R., Pawson, T., Kay, B.K., McCafferty, J., 2010. Generating a panel of highly specific antibodies to 20 human SH2 domains by phage display. *Protein Eng. Des. Sel.* 23, 279–288.
- Rondot, S., Koch, J., Breitling, F., Dübel, S., 2001. A helper phage to improve single-chain antibody presentation in phage display. *Nat. Biotechnol.* 19, 75–78.
- Sambrook, J., Russell, D., 2001. *Molecular Cloning: A Laboratory Manual*, third ed. Cold Spring Harbor Laboratory Press, New York.
- Schmiedl, A., Breitling, F., Winter, C.H., Queitsch, I., Dübel, S., 2000. Effects of unpaired cysteines on yield, solubility and activity of different recombinant antibody constructs expressed in *E. coli*. *J. Immunol. Methods* 242, 101–114.
- Schofield, D.J., Pope, A.R., Clementel, V., Buckell, J., Chapple, S.D., Clarke, K.F., Conquer, J.S., Crofts, A.M., Crowther, S.R.E., Dyson, M.R., Flack, G., Griffin, G.J., Hooks, Y., Howat, W.J., Kolb-Kokocinski, A., Kunze, S., Martin, C.D., Maslen, G.L., Mitchell, J.N., O'Sullivan, M., Perera, R.L., Roake, W., Shadbolt, S.P., Vincent, K.J., Warford, A., Wilson, W.E., Xie, J., Young, J.L., McCafferty, J., 2007. Application of phage display to high throughput antibody generation and characterization. *Genome Biol.* 8, R254.
- Schütte, M., Thullier, P., Pelat, T., Wezler, X., Rosenstock, P., Hinz, D., Kirsch, M.I., Hasenberg, M., Frank, R., Schirrmann, T., Gunzer, M., Hust, M., Dübel, S., 2009. Identification of a putative Crf splice variant and generation of recombinant antibodies for the specific detection of *Aspergillus fumigatus*. *PLoS One* 4, e6625.
- Sheets, M.D., Amersdorfer, P., Finnern, R., Sargent, P., Lindqvist, E., Schier, R., Hemmingsen, G., Wong, C., Gerhart, J.C., Marks, J.D., Lindqvist, E., 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. U.S.A.* 95, 6157–6162.
- Shukla, J., Khan, M., Tiwari, M., Sannarangaiah, S., Sharma, S., Rao, P.V.L., Parida, M., 2009. Development and evaluation of antigen capture ELISA for early clinical diagnosis of chikungunya. *Diagn. Microbiol. Infect. Dis.* 65, 142–149.
- Skerra, A., 2007. Alternative non-antibody scaffolds for molecular recognition. *Curr. Opin. Biotechnol.* 18, 295–304.
- Soltes, G., Hust, M., Ng, K.K.Y., Bansal, A., Field, J., Stewart, D.I.H., Dübel, S., Cha, S., Wiersma, E.J., 2007. On the influence of vector design on antibody phage display. *J. Biotechnol.* 127, 626–637.
- Stoevesandt, O., Taussig, M.J., He, M., 2009. Protein microarrays: high-throughput tools for proteomics. *Expert Rev. Proteomics* 6, 145–157.
- Strebe, N., Guse, A., Schüngel, M., Schirrmann, T., Hafner, M., Jostock, T., Hust, M., Müller, W., Dübel, S., 2009. Functional knockdown of VCAM-1 at the posttranslational level with ER retained antibodies. *J. Immunol. Methods* 341, 30–40.
- Taussig, M.J., Stoevesandt, O., Borrebaeck, C.A.K., Bradbury, A.R., Cahill, D., Cambillau, C., de Daruvar, A., Dübel, S., Eichler, J., Frank, R., Gibson, T.J., Gloriam, D., Gold, L., Herberg, F.W., Hermjakob, H., Hoheisel, J.D., Joos, T.O., Kallioniemi, O., Koegl, M., Konthur, Z., Korn, B., Kremmer, E., Krobatsch, S., Landegren, U., van der Maarel, S., McCafferty, J., Muyltermans, S., Nygren, P., Palcy, S., Plückthun, A., Polic, B., Przybylski, M., Saviranta, P., Sawyer, A., Sherman, D.J., Skerra, A., Templin, M., Ueffing, M., Uhlén, M., 2007. Proteomebinders: planning a European resource of affinity reagents for analysis of the human proteome. *Nat. Methods* 4, 13–17.
- Thie, H., Binius, S., Schirrmann, T., Hust, M., Dübel, S., 2009. Multimerization domains for antibody phage display and antibody production. *New Biotechnol.* 26, 314–321.
- Thie, H., Meyer, T., Schirrmann, T., Hust, M., Dübel, S., 2008. Phage display derived therapeutic antibodies. *Curr. Pharm. Biotechnol.* 9, 439–446.
- Vaughan, T.J., Williams, A.J., Pritchard, K., Osbourn, J.K., Pope, A.R., Earnshaw, J.C., McCafferty, J., Hodits, R.A., Wilton, J., Johnson, K.S., 1996. Human antibodies with sub-nanomolar affinities isolated from a large non-immunized phage display library. *Nat. Biotechnol.* 14, 309–314.
- Welschof, M., Terness, P., Kipriyanov, S.M., Stanescu, D., Breitling, F., Dörsam, H., Dübel, S., Little, M., Opelz, G., 1997. The antigen-binding domain of a human IgG-anti-F(ab')₂ autoantibody. *Proc. Natl. Acad. Sci. U.S.A.* 94, 1902–1907.
- Wheeler, D.A., Srinivasan, M., Egholm, M., Shen, Y., Chen, L., McGuire, A., He, W., Chen, Y., Makhijani, V., Roth, G.T., Gomes, X., Tartaro, K., Niaz, F., Turcotte, C.L., Irzyk, G.P., Lupski, J.R., Chinault, C., Song, X., Liu, Y., Yuan, Y., Nazareth, L., Qin, X., Muzny, D.M., Margulies, M., Weinstock, G.M., Gibbs, R.A., Rothberg, J.M., 2008. The complete genome of an individual by massively parallel DNA sequencing. *Nature* 452, 872–876.
- Wingren, C., James, P., Borrebaeck, C.A.K., 2009. Strategy for surveying the proteome using affinity proteomics and mass spectrometry. *Proteomics* 9, 1511–1517.