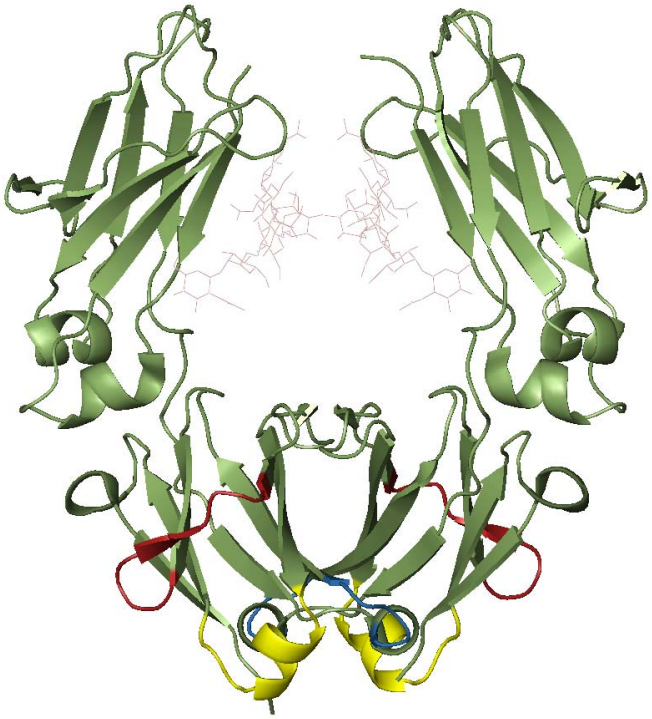
Master Thesis

Potential of IgG1-Fc libraries improved by stability-biased amino acid distributions for selection of Fcabs



BOKU - University of Natural Resources and Life Sciences
Department of Chemistry, Division of Protein Biochemistry
Christian Doppler Laboratory for Antibody Engineering

submitted by
Lisa-Sophie Handler, BSc
Vienna, Austria, 2015

Abstract

The engineering of immunoglobulin fragments generates the possibility to combine a small molecule size with precise antigen binding. Especially IgG1-Fc has proven to be a suitable engineering scaffold, where novel antigen binding properties can be added to the ability of the Fc fragment to elicit important effector functions and its high stability in serum. Since introduction of novel binding sites might impair the immunoglobulin fold and thus affect the biophysical properties of the molecule, strategies are needed to enhance efficiency of the selection. In the course of this thesis, the potential of stability-biased amino acid distributions in the EF loop design of IgG1-Fc libraries was evaluated.

Three different variant libraries based on IgG1-Fc were constructed and an optimized electroporation protocol of the EBY100 strain of *S. cerevisiae* for generation of large (10^9) yeast surface display libraries was established.
Determination of the binding ability to the structural markers FcγRI/CD64 and aC_H2 compared to surface displayed wild type IgG1-Fc revealed that the overall quality of a library with such calculated amino acid distributions in one of its loops was superior to two different designs. Also, temperatures of half-maximal irreversible denaturation ($T_{1/2}$) were assessed on a library scale.
Furthermore, the enrichment in antigen binding, resulting from screening and selection for IgG1-Fc binding to HER3 by using magnetic bead- and fluorescence-activated cell sorting (MACS/FACS), was evaluated by flow cytometry.
Due to various obstacles during the naïve selection rounds, the final validation of the concept through biochemical characterization of solubly expressed single clones, could not be finished in the course of this thesis. However, distinct ideas for the improvement of *de novo* selections of IgG1-Fc yeast surface display libraries were generated for future application.

Zusammenfassung

Das *Engineering* von Immunglobulin-Fragmenten bietet die Möglichkeit, kleine Molekülgrößen mit präziser Antigen-Erkennung zu kombinieren. Besonders IgG1-Fc hat sich als geeignetes Grundgerüst erwiesen, um wichtige Effektor-Funktionen und Eigenschaften wie eine hohe Serumstabilität mit Antigenspezifität zu kombinieren. Durch das Einfügen neuer Bindungsstellen wird die Proteinfaltung oftmals beeinträchtigt und die biophysikalischen Eigenschaften verschlechtert, weshalb Strategien benötigt werden um den Selektionsprozess effizienter zu gestalten. Die Evaluierung des Potenzials von stabilitäts-gewichteten Aminosäure-Verteilungen im *EF-Loop*-Design von IgG1-Fc Variantenbibliotheken war das Ziel dieser Arbeit.

Drei verschiedene IgG1-Fc Bibliotheken wurden konstruiert und ein optimiertes Elektroporationsprotokoll für *S. cerevisiae* EBY100 etabliert, welches die Generierung von großen (10^9) Hefeoberflächen-Display-Bibliotheken ermöglicht.
Die Bindung der Strukturmarker FcγRI/CD64 und aC_H2 im Vergleich zu Wildtyp IgG1-Fc zeigte, dass die Gesamtqualität des Bibliotheks-Designs mit der berechneten Aminosäureverteilung im EF *Loop* am Höchsten war. Zusätzlich wurden die *Temperaturen halb-maximaler irreversibler Denaturierung* ($T_{1/2}$) im Biobliotheks-Maßstab ermittelt.
Auch die Anreichung in der Antigenbindung, resultierend aus Selektionen von HER3-bindenden IgG1-Fc mittels magnetischen und fluoreszenz-aktivierten *cell sorting* (MACS/FACS) wurde durchflusszytometrisch evaluiert.
Aufgrund verschiedener Probleme im Zuge der naïven Selektionsrunden konnte die finale Validierung des Konzeptes durch biochemische Charakterisierung von löslich exprimierten Einzelklonen im Rahmen dieser Arbeit nicht mehr fertiggestellt werden. Dennoch konnten konkrete Ideen für Verbesserungen der *de novo* Selektionen von IgG1-Fc Hefeoberflächen-*Display*-Bibliotheken für zukünftige Anwendung formuliert werden.

Danke

… an Prof. Christian Obinger für die Möglichkeit in seiner Arbeitsgruppe unter großartigen Bedingungen an einem tollen Thema zu arbeiten.
… an meinen Betreuer Christoph Hasenhindl, von dem ich wirklich sehr viel habe lernen können.
… an alle lieben Kollegen der Biochemie B, mit denen die Zeit noch zusätzlich ganz viel Spaß gemacht hat.

… und an meine Mama Anna Handler und meinen Freund Florian Heinz - für generell Alles!!!

Table of Contents

Index of Figures

Index of Tables

1 Introduction

1.1 Immunoglobulins and protein engineering

Almost thirty years ago, the first monoclonal antibody (mAb), designed for minimization of immunologically triggered rejection against human transplantation tissue[1], was approved, serving as an important prototype for future therapeutic mAbs. Since then, the importance of immunoglobulins in medical applications has been growing steadily. Because of their diverse ways of action, they represent a class of therapeutics that can be used for treatment of cancer, infectious diseases or immunological disorders[2].

So far, more than thirty antibodies have been approved by the EMA and FDA[3]. Successful therapies are reported e.g. for antibodies against human epidermal growth factor receptor 2 (HER2) and epidermal growth factor receptor (EGFR)[4,5], which show elevated expression on the surface of ovarian and breast cancer cells. Furthermore, the combination of antibody therapeutics with conventional cytotoxic chemotherapeutics has proven to be superior to chemotherapy alone[6] as the molecules reach the target tissue with high specificity.

In vivo, this precise recognition is achieved by generation of a large set of different antibodies through gene segment rearrangement and somatic mutation of the hypervariable CDR-loops (see 1.1.1) forming the antigen binding sites. Through specific activation of a distinct B lymphocyte displaying one of these antibody variants on its surface, by a helper T lymphocyte sensitized to the same antigen, antibodies against this target can be produced in large quantities in the body[7].

An important field of biotechnological research, which is dealing with understanding and manipulation of such protein-protein interactions, is protein engineering. *Directed evolution* is one of its strategies and aims at mimicking evolutionary selection with increased efficiency *in vitro*[8]. Here, a combination of molecular biology- and biochemical techniques is employed in order to isolate novel molecules with desired functionalities. Immunoglobulins have proved to be a very suitable scaffold for engineering of their antigen binding properties.

In addition to that, directed evolution can be used to optimize these molecules for other important characteristics like thermal and conformational stability. Also prolongation of human serum-half life can be optimized, which is important for medical applications.

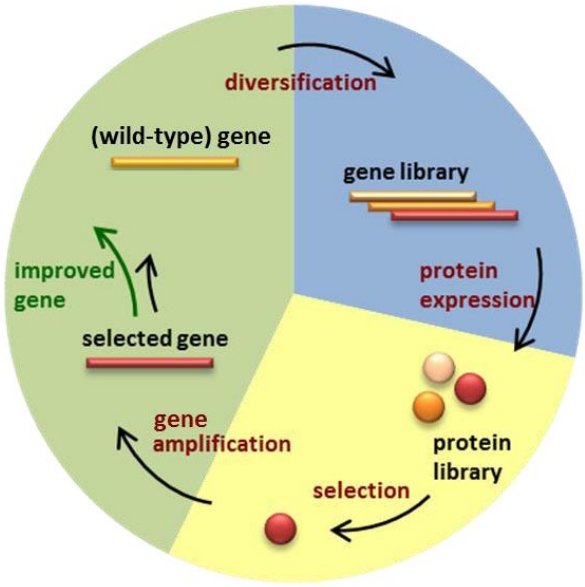


Figure 1-1 Protein engineering by directed evolution. Through diversification of a (wild-type) gene, an improved gene can be selected from a gene library that can be screened in the form of an expressed protein library

Selection is done by constructing variant libraries of the protein, mutated with methods like error prone PCR, DNA shuffling or randomization with degenerate oligonucleotides. Using display methods (e.g. yeast surface display) (see 1.3) these libraries can be screened for molecules with functional improvements[8].

1.1.1 Structure

One of the five immunoglobulin isotypes synthesized in human serum is IgG, with IgG1 being the most abundant subclass with a molecular mass of approximately 150 kDa.
The four polypeptide chains (two heavy chains and two light chains), forming the molecule, are linked by disulfide bonds and comprise constant and variable domains. Heavy chains (55 kDa) consist of three constant (C_H1, C_H2, C_H3) and one variable domain (V_H), while light chains (25 kDa) consist of one constant (C_L) and one variable domain (V_L) [7,9].

The typical Y-shaped structure of an IgG1 can be dissected into two types of antibody fragments. The two branches represent the antigen binding fragments (Fab), the stem region is called fragment crystallizable (Fc) (see Figure 1-2).

While neutralization and opsonization is achieved through binding of the hypervariable CDR-loops on the Fab fragments to surface structures of pathogens or malignant cells, the Fc fragment participates in important mechanisms of the adaptive and innate immune response (see 1.1.2).

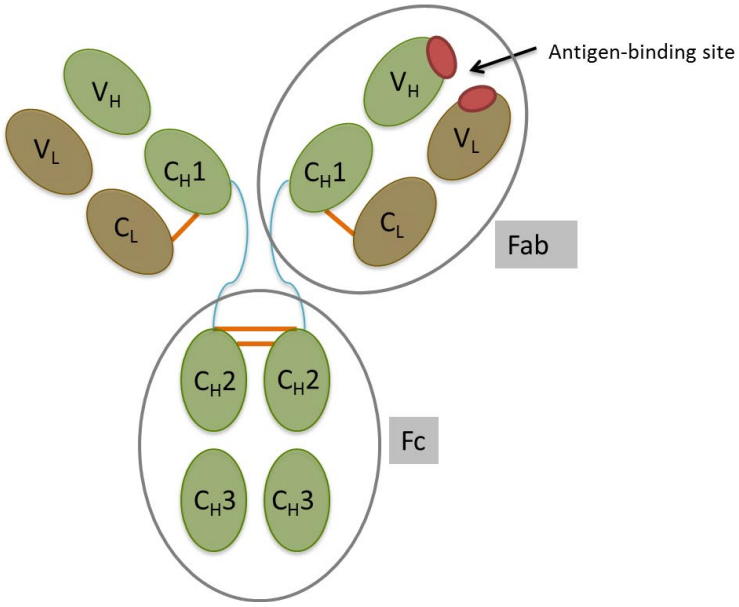


Figure 1-2 Structure of immunoglobulin G1.

1.1.2 Effector functions

The ability of the immunoglobulin constant regions to activate cells and molecules of the immune system and to mediate long *in vivo* half-life is referred to as antibody effector functions. There are two major mechanisms: ADCC and CDC.

ADCC (antibody-dependent cell-mediated cytotoxicity) is a mechanism of adaptive cell-mediated immune defense whereby an effector cell of the immune system actively lyses a target cell, whose membrane-surface antigens have been bound by specific antibodies. Classical ADCC is mediated by effector cells (e.g. natural killer (NK) cells, macrophages, neutrophils) and is triggered by interaction of the IgG1-Fc with the Fcγ-receptors (FcγRI/CD64, FcγRIII/CD16).[7]

CDC (complement-dependent cytotoxicity), consisting of a number of different small plasma proteins, is part of the innate immune system. However, it is brought into action by Fc-interaction with C1q, which starts an activation cascade resulting in the formation of cell-killing membrane attack complexes (MAC). Thus, the complement system assists antibodies and phagocytic cells to clear harmful cells from the organism.[7]

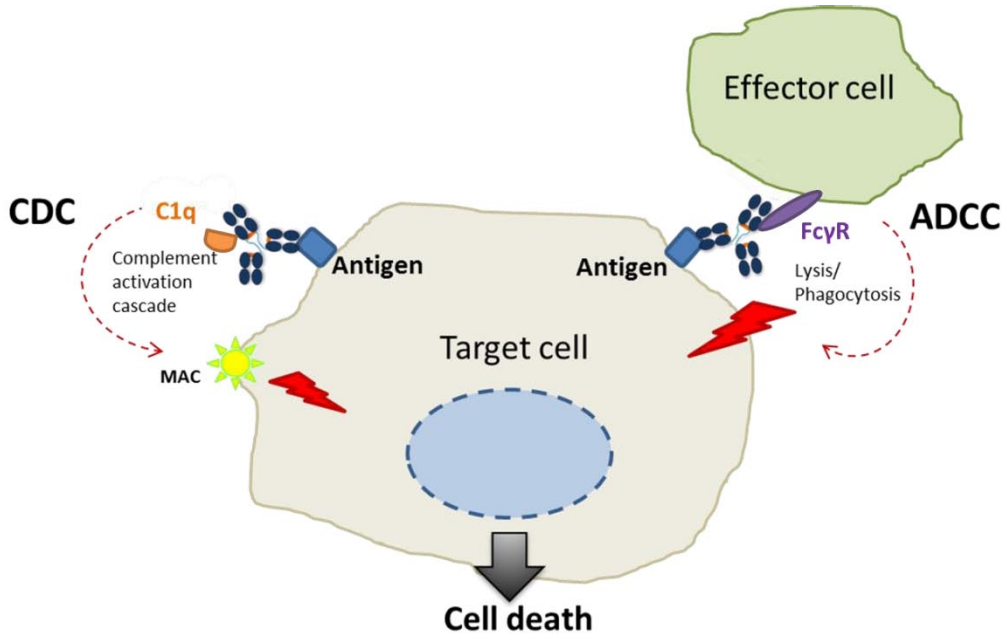


Figure 1-3 Antibody-dependent cell-mediated cytotoxicity (ADCC) and complement dependent cytotoxicity (CDC). Interaction of IgG1-Fc with Fcγ-receptors or C1q initiates effector functions.

1.1.3 Stability in serum

Another important characteristic of IgG, mediated by the Fc fragment is interaction with the neonatal Fc receptor (FcRn).

In contrast to the majority of proteins, the half-life of IgG is about 20 days[10]. This is caused by protection of proteolysis through binding of FcRn in endosomes of catabolic cells which facilitates recycling of IgG.

1.2 Antibody fragments as engineering scaffold

The tertiary structure of the IgG1 domains, the so-called immunoglobulin fold, which can also be found in many other proteins, makes the protein interesting for engineering. Its barrel-like structure is composed by two β-sheets, linked via a disulfide bond and built of anti-parallel β-strands[11]. While hydrophobic core of this β-barrel has shown to be conserved in shape, the N-terminal CDR-loops between two β-strands building the antigen binding sites are highly variable in sequence and length[12]. Thus, they offer a great mutation target in directed evolution.

However, the size of big multi-domain proteins might cause poor penetration of tissues like solid tumors[4] and are difficult to produce efficiently, so current research is done towards the development of smaller antibody formats. Different variants such as Fabs, single-chain variable fragments (scFvs)[13] or variable fragments of a heavy chain antibody (VHH)[14] are constructed and engineered successfully for improved properties[15].
Nevertheless, the need of a molecule that combines the ability to elicit the important effector functions with highly specific antigen binding, low off-target effects and high stability in serum, maintains.

1.2.1 Antigen binding fragment crystallizable of IgG1

Since 2009 modular antibody technology exploits the structural similarity between the immunoglobulin variable domains to the constant domains for introduction of antigen binding sites into the C-terminus of IgG1-Fc[16]. This allows for the design of antigen-binding IgG1-Fcs (Fcabs) through mutation of the structural loops AB, CD and EF of the C_H3 domain (see Figure 1-4)[17] .

As shown by Wozniak-Knopp et al., 2010[18] and Traxlmayr et al., 2011[19], this technology enables the generation of novel functions in therapeutic antibodies while the structural integrity, antibody effector functions and *in-vivo* half-life are preserved. This also offers possibilities for the design of bispecific or multivalent antibodies. Furthermore, IgG1-Fc is a third in size of a full-size IgG1, and therefore can be produced at higher yields and is expected to better reach its intended targets.

IgG1-Fc is a homodimer of the two C-terminal domains of the heavy chains (C_H2 and C_H3) linked in the N-terminal Hinge region by two disulfide bonds. Its conformation is determined by interactions between residues of the two C_H3 domains and an N-linked glycan on the N297 residue of each C_H2 domain, providing an open arrangement between the two C_H2 domains. Through binding to different ligands, IgG1-Fc is responsible for mediating ADCC and CDC (see 1.1.2). Interaction with the neonatal Fc-receptor (FcRn) provides the long serum half-life typical for immunoglobulins.

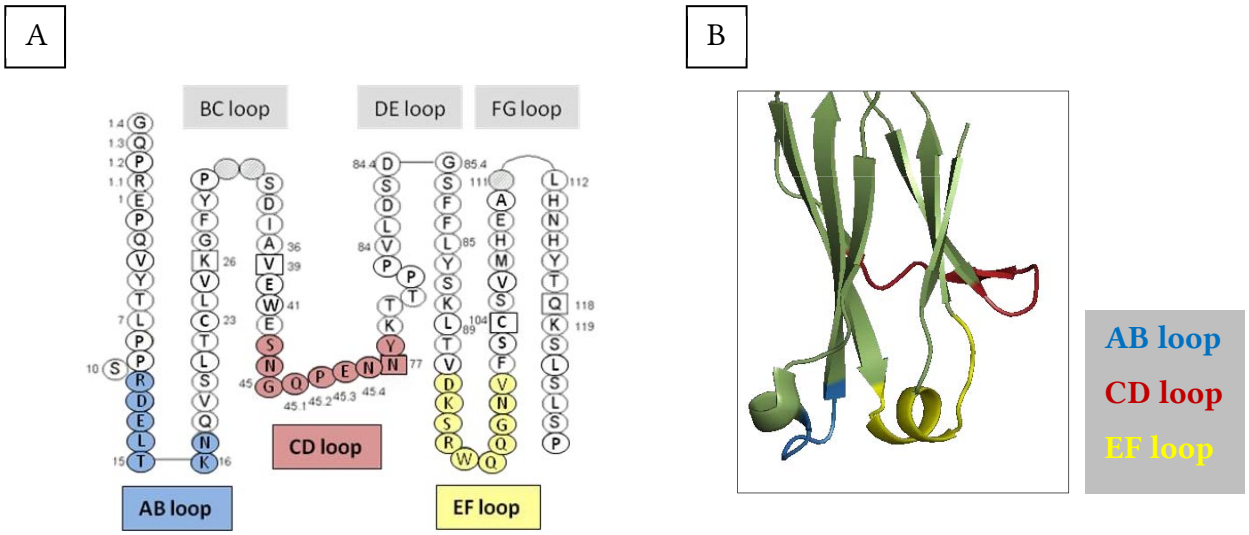


Figure 1-4 Structural loops of the IgG1 constant domain C_H3 A: Pearl chain depiction B: Schematic representation of the secondary structures (derived from PDB-ID: 1OQO)

1.2.2 Stability Engineering of IgG1-Fc

IgG1-Fc has shown to be a suitable engineering scaffold, as the overall fold is hardly impaired by mutation of the structural loops. However, the thermal stability is possibly decreased[19–21], indicating the need for optimization of the biophysical properties of these molecules.

Different ideas have been pursued in order to stabilize antigen-binding IgG1-Fcs. Inter- and intradomain disulfide bonds were introduced and engineered, respectively[20,21].
Through combination of experimental and computational means, progress was also made in understanding the effect of loop mutation or insertion of additional residues on the molecule structure[22]. This *in silico* loop reconstruction gave information about the effect of single point mutations on the stability of the C_H3 domain and also revealed regions of high

local flexibility that thus are thought as suitable for engineering. In addition, molecular dynamics simulations proved the beneficial combination of insertion of stabilizing stem residues and elongation of the EF loop.[23]
Moreover, a stability landscape of the C_H3 domain of IgG1-Fc was constructed[24], giving insight into the structural relevance of single amino acid residues.

1.2.3 Stability-biased amino acid distributions in library design

Another approach for stability improvements is to focus on the library design. As integral part of modular antibody technology, it offers a wide range for further improvements, as a well-designed library is able to provide a larger pool of correctly folded and stable clones to select from. Thus, the efficiency of the discovery process can be largely increased.

Previous to this work, an IgG1 library with a fully randomized EF loop was heat shocked at 60 °C and stained with the structurally specific ligand FcγRI/CD64. Several rounds of these selections were performed in combination with high throughput sequencing after each sorting step and multiple sequence alignments, revealing distinct amino acids at each position of the loop that contribute to stabilization.
It was hypothesized, that the residual binding to the conformational epitope of FcγRI/CD64 is strongly decreased for incorrectly folded clones and that certain amino acids are depleted at critical positions. The experiment resulted in a calculated, stability-biased amino acid distribution for each residue of the EF-loop which was then translated into a library design where an ideal loop variability shall meet a stable protein fold.

In the present study, the potential of library design derived from the steps described above, EF-CARA1, is evaluated through comparison to two other libraries. In the EF-CARA2 design, the amino acid distribution for the EF loop residues is biased towards the wild type sequence. The UP+0_RW design has a fully NNK randomized EF loop (see 3.1.3).
In all three libraries, the AB loop is NNK randomized. Because of steric distance to an antigen and thus less impact on target-binding interactions, the CD loop sequence was not altered. The aim of this setup is to validate, whether clones resulting from EF-CARA1 show biophysical properties superior to clones resulting from EF-CARA2 and UP+0_RW and, if so, that this effect is not due to the absence of stop-codons and cysteines in the loops, resulting from the use of degenerate ELLA oligonucleotides for DNA library generation (see 3.1.3.2).

Table 1-1 Libary design

Library name	AB	CD	EF	Aa distribution
EF-CARA1	LTKNQ **NNK**	wild type	DKSRWQQGNV **CARA1**	**experimentally determined and calculated**
EF-CARA2	LTKNQ **NNK**	wild type	DKSRWQQGNV **CARA2**	**wild type bias**
UP+0_RW	LTKNQ **NNK**	wild type	DKSRWQQGNV **NNK**	**random**

1.3 Yeast surface display

Yeast surface display is an effective tool for isolation of novel proteins with improved affinity, specificity or stability from naïve libraries. The most abundant variant is the Aga1p/Apa2p system of *Saccharomyces cerevisiae* in which the yeast agglutinin mating receptor subunits, which are covalently linked via two disulfide bonds, are used as anchor for surface presentation of proteins[15].

The display system consists of genetically modified yeast, carrying the *aga1* gene under the control of the *GAL1* galactose inducible promoter, and of the expression vector with the *aga2* gene, also under the control of the *GAL1* promoter. In the setup used in the current thesis, the gapped expression vector is reconstituted via gap-repair in *S. cerevisiae*. This is achieved by adding upstream and downstream homology sequences to the insert fragments carrying the partially altered IgG1-Fc gene, which enables homologous recombination.
The N-terminus of the protein is expressed as a fusion protein to Aga2p, presenting the C-terminal structural loops for selection staining. Aga1p is anchored within the cell wall via a glycophosphatidylinositol-(GPI)-anchor ensuring a correct genotype-phenotype coupling. For normalization, an N-terminal Xpress expression tag can be detected via antibody staining as depicted in Figure 1-5.

Typically, each yeast cell carries 10^4 to 10^5 copies of the respective protein variant on its surface[25]. Screening of the libraries can be achieved by using *magnetic bead cell sorting* (MACS) for fast elimination of non-binders or incorrectly folded clones (see 3.3.2) and *fluorescence activated cell sorting* (FACS) for isolation of higher affinity binders (see 3.3.3).

Also binding kinetics of single clones can be determined on the yeast surface via titration with the antigen, reducing the need for soluble expression and purification of the protein[26]. Kinetic constants have been shown to strongly correlate with values derived from other methods[27]. Another advantage of yeast surface display is the fact that *S. cerevisiae* is a eukaryotic organism which is able to correctly fold most complex human proteins and perform post-translational modification like disulfide bond formation or glycosylation.[28]

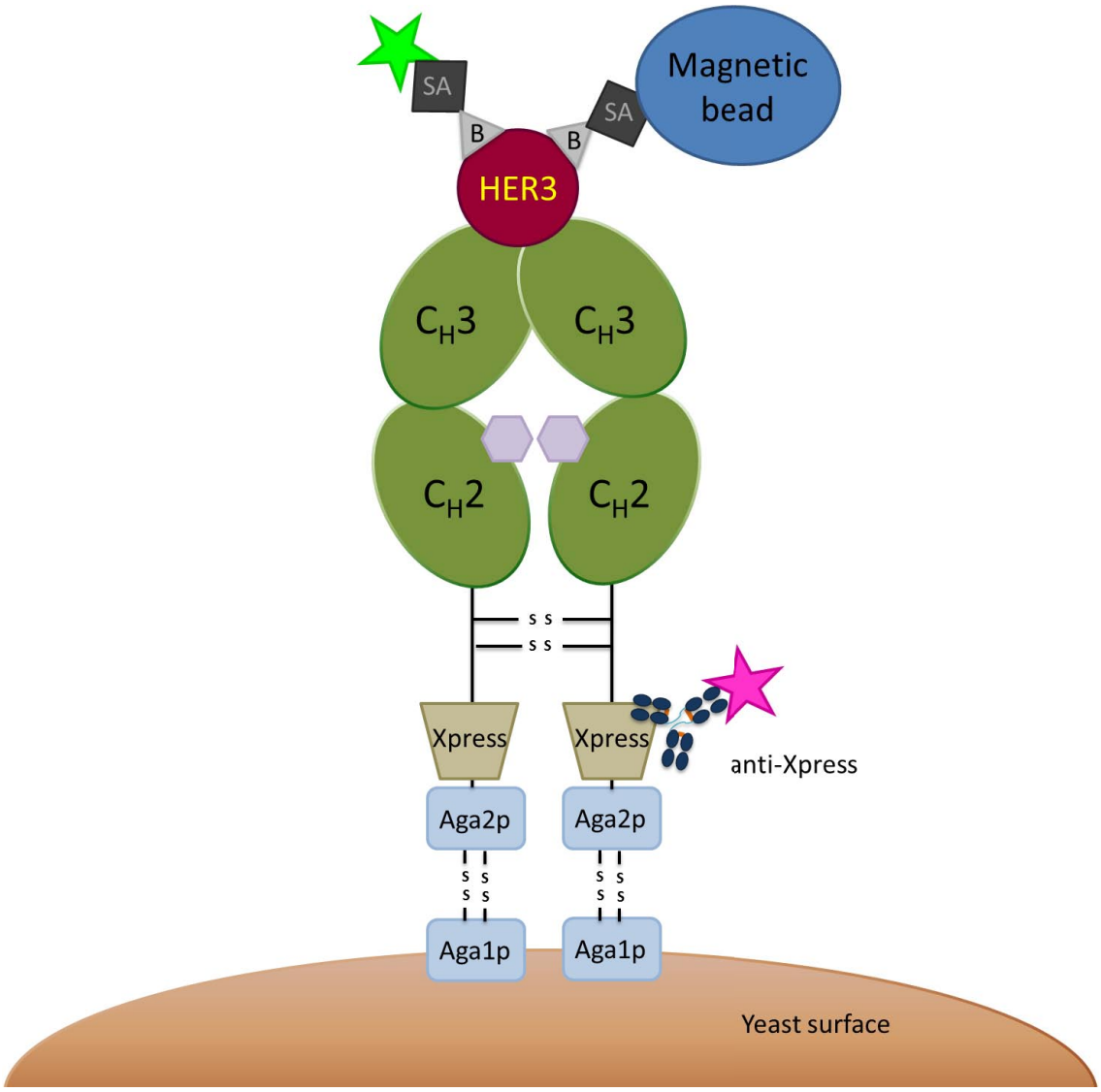


Figure 1-5 Schematic representation of the yeast surface display system used for screening
Green is the engineering scaffold IgG1-Fc, flanked by the Xpress tag (olive) for normalization. The N-terminal Aga2p as well as the cell-wall-linked Aga1p are depicted in blue. Glycan structures on the C_H2 domain are shown in violet. The biotinylated target protein HER3 (red) is detected by fluorophore-conjugated proteins (e.g. streptavidin-phycoerythrin (SA-PE), bright green) or streptavidin-coated magnetic microbeads (blue). Displaying cells are stained e.g. by an anti-Xpress*APC antibody (pink).

1.3.1 Biomarker HER3 as selection target

HER3 (human epidermal growth factor receptor 3) is a tryrosine kinase receptor from a family of growth factor receptors known since the early 1980s to play a role in cancer development. These proteins elicit intracellular signaling promoting cell proliferation and are normally responsible for cell growth and differentiation[29].

HER3 consists of an extracellular, a transmembrane and an intracellular area. Through binding of a ligand (e.g. neuregulin) to the extracellular domain, the receptor changes its conformation and thus is able to form dimers. Especially the heterodimerisation with the human epidermal growth factor receptor 2 (HER2) has shown to result in a very high signaling activity[30] (see Figure 1-6).

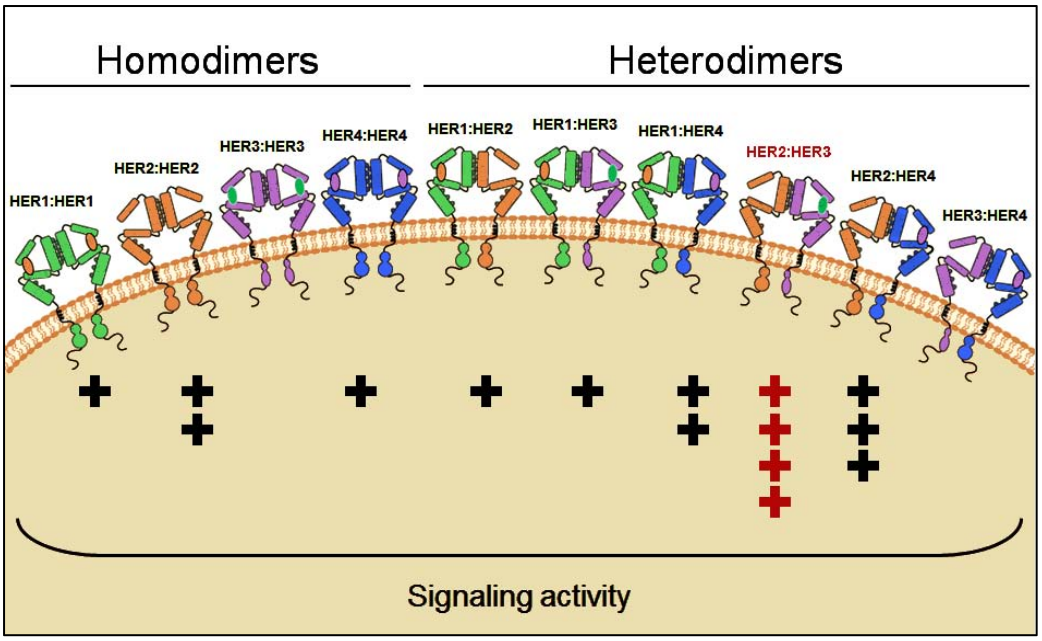


Figure 1-6 Signalling activiy of tyrosine kinase receptor dimers of the HER family
(Adapted from http://www.biooncology.com/research-education/her/dimer)

Unlike the other receptors of the HER family, the intracellular domain of HER3 has no intrinsic kinase activity[31]. Therefore, it can only be transphosphorylated through other HER receptor monomers and cannot be targeted by tyrosine kinase inhibitors. This provides an escape-mechanism for cancer cells through upregulation of HER3 expression which causes a compensatory shift towards HER3 phosphorylation upon incomplete loss of e.g. HER2 kinase activity[32].

In the current thesis, a construct consisting of the extracellular domain of human HER3 with a C-terminal polyhistidine tag (Sino Biological, BDA Beijing, CN) was used as antigen for *de novo* selections of IgG1-Fc yeast surface display libraries. This extracellular domain comprises of 635 amino acids and has a calculated molecular weight of 70.2 kDa. Selections were performed against the biotin-tag of the antigen, which was added in house (see 3.2.3.2). Antibodies directed against the extracellular HER3 domain possibly hinder heterodimerization with HER2. Thus HER3 is an interesting target for antibody treatment.

1.4 Aim of the thesis

The objective of this work is to isolate novel binding molecules, based on the IgG1-Fc scaffold, specific for the human epidermal growth factor receptor 3, which was chosen as target protein because of its therapeutic potential.

Firstly, three different protein libraries were to be constructed via homologous recombination in *S. cerevisiae* and functionally characterized on the yeast surface ($T_{1/2}$ values). Secondly, HER3-binding molecules should be selected by magnetic bead- and fluorescence activated cell sorting. The question to be answered was, whether the concept of stability-biased amino acid distributions is applicable for an efficient discovery process of Fcabs.

2 Material

2.1 Bacteria and yeast strains

Escherichia coli

DH5α™ derivative (New England Biolabs, Ipswich, MA)

fhuA2 Δ(argF-lacZ)U169 phoA glnV44 Φ80 Δ(lacZ)M15 gyrA96 recA1 relA1 endA1 thi-1 hsdR17

Saccharomyces cerevisiae

EBY100 (Boder & Wittrup, 1997)[15]

(a GAL1-AGA1::URA3 ura3-52 trp1 leu2Δ1 his3Δ200 pep4::HIS2 prb1Δ1.6R can1 GAL), Trp-Leu-

2.2 Plasmid - pYD1dH3s

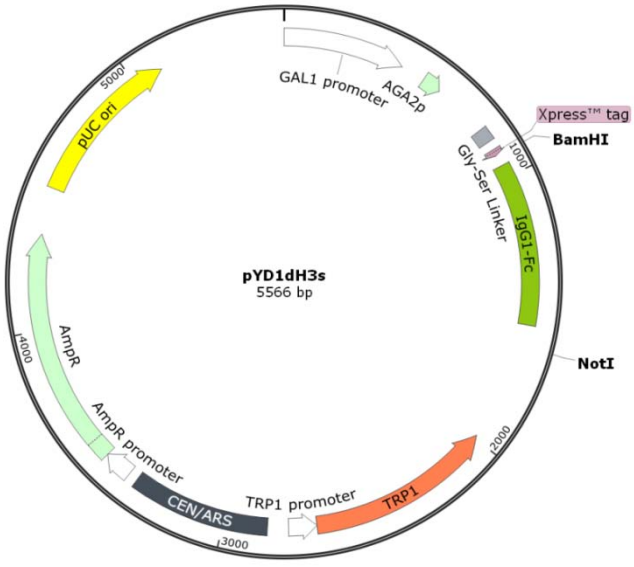


Figure 2-1 Schematic representation of the vector pYd1dH3s with features: origins or replication for yeast (CEN6/ARS 4) and *E .coli* (pUC ori), selection markers for yeast (TRP1 – auxotrophic marker coding for tryptophan synthetase) and *E. coli* (AmpR – gene conding for beta-lactamase→ ampicillin resistance), GAL1 promoter, a glycin/serine linker for spacing of the IgG1-Fc from Aga2p and the Xpress tag

A yeast-codon-optimized IgG1-Fc gene cloned into an adapted pYD1 vector (Invitrogen) for *S. cerevisiae* surface expression as Aga2p fusion protein, using BamHI and NotI, was used as expression vector. Adaption of the vector was described by Hasenhindl et al., 2014[23]. For construction of a plasmid DNA library, the vector was digested with the type II restriction enzyme BsmBI and co-transformed with the library inserts into *S. cerevisiae* for homologous recombination (see 3.1.3).

2.3 Primers

AB_template_fwd
5'-GTCAGCCTGACCTGCCTGGTC-3'

CD_frag_back
5'-CACGGTGAGCTTGCTGTAGAGGAAGAAGGAGCCGTCGGA-3'

EF_template2_fwd
5'-TTCTCATGCTCCGTGATGCAT-3'

EF_template2_back_new
5'-GCAGGCAGCAAAATTAAAATTGTCTTAG-3'

LTKNQ_NNK
5'GTGTACACCCTGCCACCATCCCGGGATGAGNNKNNKNNKNNKNNKGTCAGCCTGACCTGCC TGGTC-3'

EF_CARA1
5'-TCCGACGGCTCCTTCTTCCTCTACAGCAAGCTCACCGTG—ELLA CARA1—
TTCTCATGCTCCGTGATGCAT-3'

EF_CARA2
5'-TCCGACGGCTCCTTCTTCCTCTACAGCAAGCTCACCGTG—ELLA CARA2—
TTCTCATGCTCCGTGATGCAT-3'

UP+0_RW_EF
5'TCCGACGGCTCCTTCTTCCTCTACAGCAAGCTCACCGTGNNKNNKNNKNNKNNKNNKNNKN NKNNKNNKTTCTCATGCTCCGTGATGCAT-3'

2.4 Chemicals

All chemicals were purchased standardized from following manufacturers and used unaltered:

Invitrogen

Life Technologies

Merck Chemicals

Roth

Sigma Aldrich

2.5 Media

Cultivation of *Escherichia coli*

LB medium	10 g/L peptone
sterilized by autoclaving	5 g/L yeast extract
	5 g/L NaCl
1000X Ampicillin	100 g/L

For agar plates, 1.5 % agar-agar was added to the media

Cultivation of *Saccharomyces cerevisiae*

SD-CAA medium	10 g/L bacto casamino acids
sterilized by autoclaving	20 g/L glucose
	0.1 M KH_2PO_4/K_2HPO_4 , pH 6
	3.4 g/L yeast nitrogen base
	10 g/L $(NH_4)_2SO_4$
	0.1 g/L L-leucine
SGR-CAA medium	10 g/L bacto casamino acids
sterilized by autoclaving	20 g/L galactose
	10 g/L raffinose
	0.1 M KH_2PO_4/K_2HPO_4 , pH 6
	3.4 g/L yeast nitrogen base
	10 g/L $(NH_4)_2SO_4$
	0.1 g/L L-leucine

YPA medium sterilized by autoclaving	20 g/L peptone 10 g/L yeast extract 40 mg/L adenine-hemisulfate
100X Penicillin/Streptomycin	6.05 g/L Penicillin G sodium salt 10 g/L Streptomycin sulfate salt
PBS/BSA sterilized by filtration	8 g/L NaCl 0.2 g/L KCl 1.15 g/L of Na_2HPO_4 (anhydrous) 0.2 g/L of KH_2PO_4 20 g/L bovine serum albumin

For agar plates, 1.5 % agar-agar was added to the media.

3 Methods

In the course of this work, all centrifugation steps for small volumes were carried out in a 1-15 Microfuge (Sigma, St. Louis, MO) using rotor 12124 and for larger volumes or microtiter plates in a Heraeus Multifuge 1S-R (Thermo Fisher Scientific, Waltham, MA) using rotor 75002000 and the buckets 2001 or 2014, respectively.

3.1 Generation of IgG1-Fc mutant libraries

3.1.1 Cultivation of *Escherichia coli*

Incubation

Bacterial cultivation steps were performed in Ecotron shaking incubators (INFORS HT, Bottmingen, CH).

3.1.2 Cultivation of *Saccharomyces cerevisiae*

Incubation

For all yeast cultivation steps in this work, incubation was performed in shaking incubators 3033 (GFL Gesellschaft für Labortechnik mbH, Burgwedel, DE).

Determination of cell density

Optical density of cultures (OD) was determined by spectrophotometrical measurement of the absorbance at 600 nm wavelength (Hitachi U-1100/Agilent 8453). Depending on the OD, cultures were diluted 1:10 or 1:100 in a disposable cuvette in order to be able to measure within the linear range of the instrument. Blanking was done using the respective growth media. An OD_{600} value of 1 corresponds to approx. 10^7 *S. cerevisiae* cells/ mL culture.

3.1.3 Site saturation mutagenesis for randomization of AB- and EF-loop

According to the library design, AB loop and EF loop were randomized on a codon level. Therefore, degenerate oligonuceliotides (Ella Biotech, Martinsried, DE) synthesized from trimer phosphoramidites were used, which allows circumvention of codon bias, generation of stop-codons and frameshift mutations as 20 distinct codons representing the 20 amino acids are used out of the 64 possible codons.

For the EF loops of EF-CARA1 and EF-CARA2, the desired amino acid distribution for each residue is achieved by mixing the trimers in the calculated molar ratios with respect to the different coupling reaction factors during primer synthesis. The codons encoding cysteines were excluded for both libraries, as randomly introduced cysteines increase the probability of incorrect cysteine pairing.
For the EF loop of UP+0_RW and the AB loops only degenerate NNK codons (N = A/C/G/T and K, either G or T) were used for primer synthesis, resulting in 32 possible codons encoding all 20 amino acids but only one stop codon. This offers a way of "full" randomization balancing between highest alteration possibility and minimal translation stops.

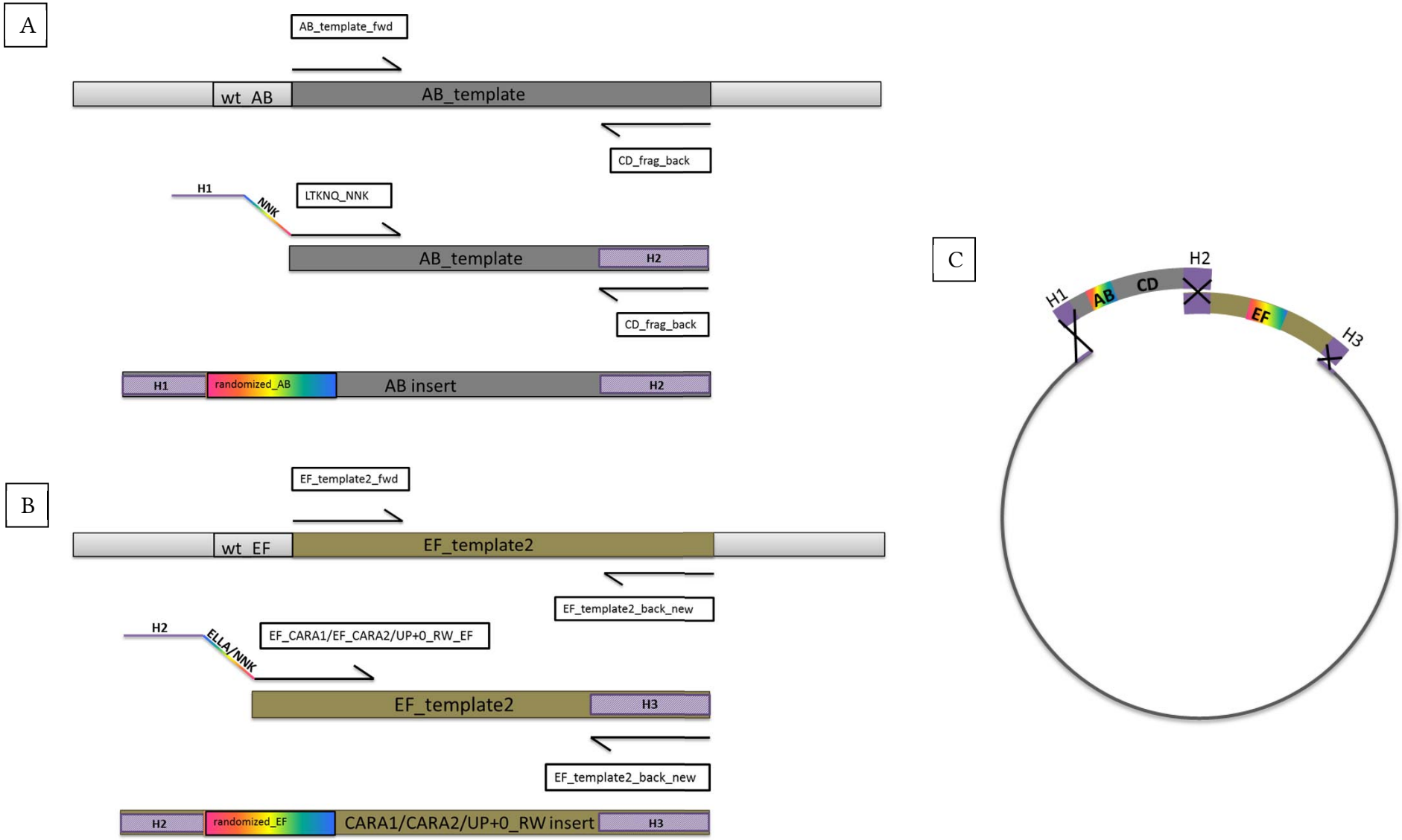


Figure 3-1 Generation of the IgG1-Fc libraries with randomized AB- and EF loops in a multi-step PCR – A: NNK randomization of the AB loop B: ELLA and NNK randomization of the EF loop, from left to right: CARA1, CARA2, UP+0_RW C: vector reconstitution in *S. cerevisiae* by gap-repair

3.1.3.1 Preparation of vector DNA

pYd1dH3s3Ba

The plasmid was amplified in *E. coli*. For this purpose 2x 200 mL LB-medium in 1 L flasks were inoculated with cells from a cryoculture and incubated overnight at 37 °C, shaking at 220 rpm. Plasmid purification was done using the Plasmid *Plus* Maxi Kit (Qiagen, Venlo, NL) according to the manufacturers' instruction. Digest of the vector backbone was done by adding 1 µL of the restriction enzyme BsmBI (New England Biolabs, Ipswich, MA) per 10 µg vector and the corresponding buffer (NEB 3) to yield a final concentration of 1X, followed by incubation for 4 h at 55 °C. Finally, the cut vector was purified by agarose gel electrophoresis (see 3.1.3.3).

3.1.3.2 Production of library inserts

For generation of the three IgG1-Fc libraries with randomized AB- and EF-loops, a cloning strategy with two subsequent polymerase chain reactions was pursued. In the first step a template sequence was generated, with forward primers designed to anneal just behind the region for the AB- or EF- wild type loop sequence, in order to provide the right starting point for the second PCR. Hence, for insert production, the degenerate forward primers could anneal to the template and provide the sequence for randomization (NNK or ELLA trimers) as well as a homology region at the 5' end (H1 for the AB insert and H2 for the EF inserts) for gap repair-driven vector reconstitution in *S. cerevisiae*. All reverse primers covered a region at the 3' end (H2 and H3, respectively).

Starting from pYD1dH3s carrying the gene coding for wild type IgG1-Fc, templates for each insert were synthesized using the primers AB-template_fwd and CD_frag_back for the generation of the AB template, and EF_template2_fwd and EF_template2_back_new for the generation of the EF template. In the second PCR setup, a mutagenic forward primer (LTKNQ_NNK for the AB insert, EF CARA1, EF_CARA2 and UP+0_RW_EF for the EF inserts) was used for insert generation (see Figure 3-1).
As specified in Table 3-1, a touch-down polymerase chain reaction setup was chosen in order to reach a balance between specificity of primer annealing and efficiency of amplification.

For a 20X PCR reaction, the following mastermix was prepared:

- 550 µL Q5® High-Fidelity 2X Master Mix (New England Biolabs, Ipswich, MA)
- 5.5 µL 100 µM forward primer
- 5.5 µL 100 µM reverse primer
- 539 µL PCR water (Sigma, St. Louis, MO)

After putting 50 µL of the mastermix aside as a negative control, 10.5 µL of template were added to the solution which was then split up to PCR tubes in 50 µL portions.
The samples were briefly spun down and put into the thermal cycler ARKTIK (Thermo Fisher Scientific, Waltham, MA). Samples that were processed on at a later stage were stored at -20 °C.

Table 3-1 Thermal cycle setup

Step	**Cycles**	***T* [°C]**	**time**
Initialization	1	98	30 s
Denaturation	10	98	10 s
Annealing		70	20 s
Extension		72	20 s
Denaturation	20	98	10 s
Annealing		70	20 s
Extension		72	20 s
Final Extension	1	72	10 min
Cooling	1	4	10 min

3.1.3.3 Purification and quantification of DNA

Agarose gel electrophoresis

PCR product and digested vector were purified from unspecific products, primer dimers etc. on a 2% agarose gel and subsequent gel extraction.

For this purpose, 6.8 g LE agarose (Biozym, Hessisch Oldendorf, DE) were supplemented with 6.8 mL 50X TAE buffer (242 g/L Tris, 57.1 mL/L glacial acetic acid, 100 mL 0.5 M EDTA, pH 8.3) filled up to 340 mL with ddH_2O and melted in the microwave.

After addition of 3.4 µL 10000X SYBR® Safe DNA Gel Stain (Invitrogen, Carlsbad, CA), the gel was poured into a gel carrier, a comb put in the liquid gel for formation of pockets, and let solidify.

6X DNA loading dye (Thermo Fisher Scientific, Waltham, MA) was added to each sample and loaded onto the gel in an electrophoresis chamber (Peqlab, Erlangen, DE) filled with 1X TAE buffer. Also, 15 µL of DNA ladder (MassRuler DNA Ladder Mix Ready-to-Use, Thermo Fisher Scientific, Waltham, MA for vector purification and Quick-Load® 100 bp DNA Ladder, New England Biolabs, Ipswich, MA for purification of inserts) was loaded onto the gel. A voltage of 100 V was applied for approximately 60 min.

The stained DNA was detected under UV light in the Gene Flash 2301 Macrodrive 1 (Syngene Bio Imaging, Cambridge, UK) and gel slices of the bands of desired sizes were cut out with a razor blade. Finally, the DNA was purified from the gel using the QIAquick Gel Extraction Kit (Qiagen, Venlo, NL) according to the manufacturers' instruction and a Vac-Man® Laboratory Vacuum Manifold (Promega, Fitchburg, WI).

Ethanol precipitation

For further purification of gel-extracted DNA fragments, 3 µL Pellet Paint® Co-Precipitant (69049, Merck Millipore, Billerica, MA) and 10% 3M NaAc pH 5.4 were added to the DNA solution and mixed briefly followed by addition of 3-4 V of 100% EtOH. The vigorously mixed solution was incubated for 5 min at room temperature and subsequently centrifuged for 10 min at 4 °C for and 14000 rpm. After visually checking for a lively pink colored pellet , the supernatant was removed and the pellet washed twice with 500 µL 100 % EtOH. The DNA pellet was air dried for 1-2 h and resuspended in the desired volume of sterile H_2O until the pink dye was completely dissolved.

illustra GFX PCR DNA and Gel Band Purification Kit

Rapid clean-up and concentration of DNA in solution was done using a glass fiber matrix based column kit (GE Healthcare, Waukesha, WI).

Quantification

3 µL of DNA samples were loaded onto a Nanodrop ND-1000 spectrophotometer (Thermo Fisher Scientific, Waltham, MA), which measures the absorbance of a sample in a range of wavelengths (220 nm - 750 nm). The value derived at 260 nm allows determination of the DNA (OD260 = 1 ≙ 50 µg/mL DNA) as well as important ratios (e.g. 260 nm/280 nm) that provide information about the purity of the solution.

3.1.4 Electroporation of *S. cerevisiae*

For library construction, *S. cerevisiae* EBY100 were transformed with the BsmBI cut pYD1dH3s3Ba together with the gel-purified inserts using an adapted protocol of Chen et al. 2013[33].

All centrifugation steps were performed for 3min at 4 °C and 3000 rpm.

Preparation of electrocompetent *S. cerevisiae* cells

An EBY100 colony from a two times single streak plate on YPD was inoculated in 20 mL YPD medium in a 100 mL shake flask and cultured overnight at 30 °C, shaking at 250 rpm.

For a 4-fold transformation with two controls, 150 mL of pre-warmed YPDA in a 1 L shake flask without baffles were each inoculated to yield an OD_{600} of 0.2, using the following formula:

$$V_{\text{YPD cultures}} = \frac{OD_{600,\ \text{required}}}{OD_{600,\ \text{YPD cultures}}} \times V_{\text{YPD, resuspension}} = \frac{0.2}{OD_{600,\ \text{YPD cultures}}} \times 150$$

The cultures were grown for approx. 5 h at 30 °C, shaking at 250 rpm, to an OD_{600} of 1.3 – 1.4. Cells were harvested by centrifugation in 50 mL conical tubes and the medium was removed.

Each cell pellet was resuspended in 24.75 mL sterile 100 mM LiAc and 250 µL freshly prepared 1M DTT was added to each tube to yield a final concentration of 10mM. Cells were incubated for 10 min at 30 °C, shaking at 225 rpm in the conical tubes. Afterwards the

pellets were washed with 25 mL ice-cold sterile H_2O, centrifuged and resuspended in 250 µL chilled sterile H_2O. The total volume was approximately 500 µl of cells for each 50 ml original culture volume. Cells were stored and all further steps were performed on ice.

DNA preparation

4 µg (1.2 moles) purified vector were added to the purified AB- and the EF-inserts, corresponding to a 20-fold molar excess of vector over insert, which was the optimum determined by test transformations. Two negative controls were prepared, one without any DNA and one with vector only.

Electroporation

For a 1-fold electroporation, 250 µL cells were added to each DNA tube and transferred to a pre-chilled 2mm electrode gap electroporation cuvette (VWR, Radnor, PA).
The square wave protocol on the Bio-Rad Gene Pulser Xcell was chosen and a single pulse at 500 V with a 15 ms pulse duration was performed. Typical "Droop" readings reported by the machine were within 5-6%. Cells were rescued with 1mL YPD and transferred to a 14 mL round bottom tube. The cuvette was rinsed with 1 mL YPD and the tubes were incubated for 1h at 30 °C, without shaking.
After incubation, the cells were pooled and the tubes rinsed with a total of 2 mL YPD.
Total volume of the library was 10 mL and 2 mL for each control.

Size determination

Serial dilutions of the suspensions of transformed cells were prepared. 100 µl of 10^{-2}, 10^{-3}, 10^{-4} and 10^{-5} dilutions were plated onto selective SD-CAA plates and incubated at 30 °C for 2 - 3 days. Library sizes were calculated from the colony count.

Cells were harvested, resuspended in 200 mL SD-CAA medium per 1-fold transformation and grown overnight at 30 °C, shaking at 250 rpm.

3.1.5 Cryo conservation of libraries

For long-term storage of the naïve and sorted libraries, cryocultures of the *S. cerevisiae* libraries were prepared.
For characterization purposes (per analysis – p.a.), the library size had to be covered once, as only 10000 events were to be recorded in flow cytometric analysis (see 3.2). For this

purpose, 1 mL of yeast culture in SDCAA medium with an OD_{600} of about 20 (corresponding to 2 x 10^8 cells/ mL) was mixed with 200 µL sterile 100% glycerol. For flow cytometric sorting, SDCAA culture containing a number of cells corresponding to at least 10 times the library size, was collected by centrifugation and resuspended in 5 mL SDCAA containing 15% of sterile 100% glycerol.
Cryocultures were stored at – 80 °C.

3.1.6 Evaluation of diversity

Isolation of plasmid DNA from yeast

The quality and diversity of libraries was determined by sequencing. Library DNA was isolated from 1 mL *S. cerevisiae* overnight liquid culture with the Zymoprep™ Yeast Plasmid Miniprep II Kit (Zymo Research), using 9 µL Zymolyase enzyme per pellet and prolonged digestion times of up to 2 hours.

E. coli transformation

For amplification of plasmid DNA, chemically competent *E. coli* cells (New England Biolabs, Ipswich, MA) were mixed with approx. 100 ng DNA for each library and incubated on ice for 30 min. Transformation mixes were heat shocked for 30 seconds at 42 °C on a thermoblock and rescued in 950 µL SOC Outgrowth Medium (New England Biolabs, Ipswich, MA). After 60 min incubation at 37 °C, shaking at 250 rpm, aliquots of the cells were plated on LB-agar plates supplemented with 100 µg/ ml ampicillin. Plates were incubated over night at 37 °C.

Sequencing

96 clones of the transformation plates were picked and transferred to two identical plates (96 well plates filled with LB-amp agar, 150 µL/ well). One plate was sent to LGC genomics for Sanger sequencing, the other plate was stored at 4 °C for backup. The primer pYD1dHback (5'-GGGAAAACATGTTGTTTACGGAG-3') used for sequencing was added on site.

3.2 Characterization of IgG1-Fc libraries by FACS measurements

3.2.1 Yeast display cultures for characterization

Pre-culture

Per library, 19 mL of SDCAA medium in a 100 mL sterile shake flask were inoculated with 1 mL cryoculture p.a. and incubated over night at 30 °C, shaking at 180 rpm. In addition, an EBY100 culture carrying the pYD1 plasmid containing the gene coding for wild type IgG1-Fc was prepared as well as a control.

Induction of surface expression

In order to ensure, that the cultures are all in the same growth phase, OD_{600} was adjusted to 1 in SDCAA, followed by incubation for another 4 h at 30 °C, shaking at 180 rpm.

Surface expression of the Fcab molecules was induced by medium change to SGRCAA. Galactose in the medium activates expression of the Aga2p fusion protein under control of the GAL1 inducible promoter. Hence, after OD_{600} measurement, the number of cells needed to yield an OD_{600} of 1 in 20 mL SGRCAA medium was harvested by centrifugation at 3000 rpm, RT for 3 min. For minimization of pre-selection by the protein quality control system of the yeast[34] an induction temperature of 20 °C was chosen. The cultures were incubated overnight, shaking at 180 rpm.

3.2.2 Assessment of thermal stability

The OD_{600} of overnight cultures was adjusted to 1 in PBS/BSA and split into 700 µL aliquots. First, all the samples were incubated on ice for 10 min and then one aliquot per culture was heat shocked for 30 min at 50, 55, 60, 65, 70 or 75 °C on a thermoblock, shaking at 300 rpm. Samples were put on ice immediately after the heat incubation and 50 µL/ well were applied to a 96 well round bottom plate. The plate was incubated for 30 min at 4 °C, shaking at 300 rpm.

3.2.2.1 Staining

All plate centrifugation steps were carried out for 5 min and 4 °C at 2000 g.
Staining the displayed IgG1-Fc variants, with a fluorescence labelled antibody against the expression tag (Xpress) and a structural ligand was done simultaneously.
After sample incubation, the plate was centrifuged and the supernatant was discarded by briefly turning over the plate.
In the first step, the pellets were stained with 2.5 µg/mLanti Xpress antibody (Invitrogen, Carlsbad, MA) conjugated to APC using the LYNX Rapid APC Antibody Conjugation Kit (AdB Serotec, Kidlington, UK) and either incubated with 1 µg/mL His-tagged Fcγ-receptor I (FcγRI/CD64) (R&D Systems, Abingdon, UK), or stained with 4 µg/mL FITC labelled anti human IgG CH2-domain antibody (AdB Serotec, Kidlington, UK). For staining of the FcγRI-His, then pellet was resuspended in 2 µg/mL anti His AF 488 conjugate antibody (Qiagen, Venlo, NL).
All staining steps were carried out in PBS/BSA. Washing steps with 100 µL PBS/BSA were performed after each staining step.
50 µL/well of the FcγRI/CD64-His solution and 25 µL/well of the anti Xpress*APC, anti His*AF 488 and anti CH2*FITC staining solutions were applied to the round bottom plate with a multichannel pipette. In addition, unstained controls of each culture were prepared, through resuspension of the cell pellets in PBS/BSA without staining agent to ensure that all the cells undergo the same treatment.
After the last centrifugation, the pellets were resuspended in 100 µl ice-cold PBS/BSA and transferred to 200 µl PBS/ BSA in FACS tubes.

3.2.3 Determination of enrichment in antigen binding

The enrichment in antigen binding of the IgG1-Fc libraries in the course of the naïve selections (see 3.3) was also determined on the yeast surface.
OD_{600} of the cultures was adjusted to 1 in PBS/BSA and the samples were incubated on ice for 10 min. 100 µL/well were applied to a 96 well round bottom plate. The plate was incubated for 30 min at RT, shaking.

3.2.3.1 Staining

All plate centrifugation steps were carried out for 3 min at 4 °C and 3000 rpm.
The capability of yeast populations displaying IgG1-Fc, to bind HER3 at different stages of the sorting process, was determined by incubation with biotinylated HER3 (HER3*bio, see 1.3.1 and 3.2.3.2) which was subsequently stained with fluorescence labelled streptavidin. Simultaneously, the expression level was determined using a fluorescence labelled antibody against the expression tag (Xpress).
In the first step, the yeast cells were stained with 2.5 µg/mL anti Xpress antibody, conjugated with APC (see 3.2.2.1), and incubated with a 300 nM antigen. Biotinylated HER3 is stained with 10 µg/mL PE-conjugated streptavidin (SA-PE, Life Technologies, Carlsbad, CA) or 10 µg/mL PE-conjugated NeutrAvidin (NA-PE, Life Technologies, Carlsbad, CA).
All staining steps were carried out in PBS/BSA. Washing steps with 100 µL PBS/BSA were performed after each staining step.
100 µL/ well of the HER3*bio and 50 µL/ well of the anti Xpress*APC and SA/NA*PE staining solutions were applied to the round bottom plate with a multichannel pipette. In addition, unstained controls of each culture were prepared by resuspending cell pellets in PBS/BSA without staining agent to ensure that all the cells undergo the same treatment.
Following the final centrifugation step, the pellets were resuspended in 100 µl ice-cold PBS/BSA and transferred to 200 µl PBS/ BSA in FACS tubes.

3.2.3.2 Biotinylation of HER3

Lyophilized HER3 (Sino Biological, BDA Beijing, CN) was reconstituted with sterile H_2O for 20 min at 20 °C, shaking on a thermoblock. For biotinylation of protein in solution, 30-fold molar excess of the biotinylation reagent (EZ-Link® Sulfo-NHS-LC-LC-Biotin, Thermo Fisher Scientific, Waltham, MA) was added according to the manufacturers' instruction and incubated at room temperature for 30 min.

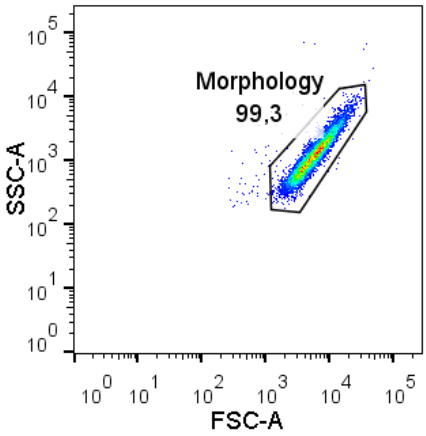
As a biotin quantitation was performed in order to exclude the possibility of using over- or under-biotinylated antigen, the protein solution was treated with Amicon® Ultra 0.5 mL centrifugal filter devices (30 K cutoff) (Merck Millipore, Billerica, MA) to remove non-reacted biotin.

The grade of biotinylation was determined with a HABA/Avidin test using the Pierce Biotin Quantitation Kit (Thermo Fisher Scientific, Waltham, MA) according to the manufacturers' instruction. Horeradish peroxidase (1 mg/mL) with 1 mole biotin/mole HRP was used for normalization. The number of biotin molecules per molecule HER3 should be in the range of 1-5.

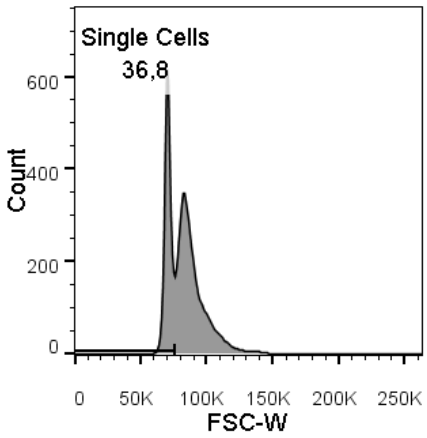
3.2.4 Flow cytometric analysis

Fluorescence stained *S. cerevisiae* cells were analyzed with a FACS CantoII (BD, Franklin Lakes, NJ). 10000 events were recorded and cells were gated for

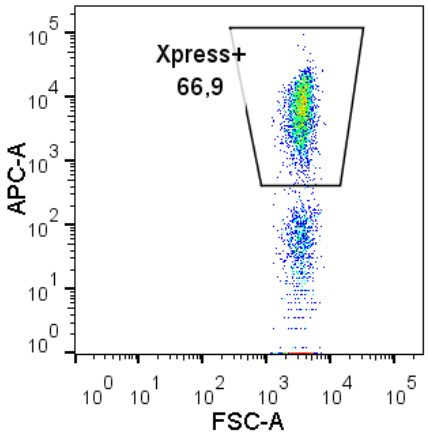
- desired morphology (forward scatter area vs. side scatter area)



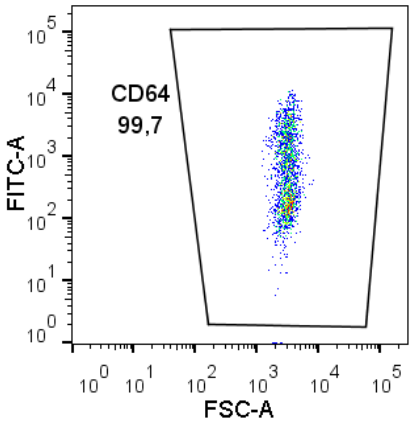
- presence as single cells (forward scatter width vs. histogram)



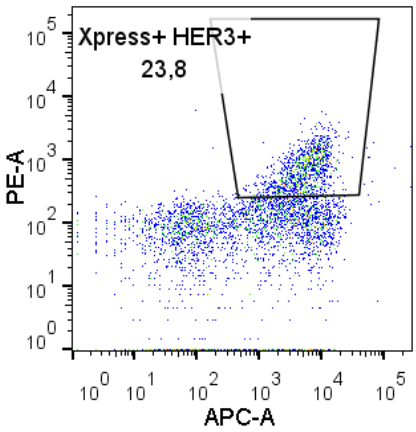
- display of the Xpress-tag (forward scatter area vs. APC fluorescence signal area)



- and fluorescence signal of the respective fluorescence labelled ligand bound to the displayed IgG1-Fc (forward scatter area/APC fluorescence signal area vs. PE/FITC fluorescence signal area)



Xpress+: **Geometric Mean** : *FITC-A*: 532



3.2.4.1 Data analysis for $T_{1/2}$ determination

Resulting data were analyzed using the FACSDiva software (BD, Franklin Lakes, NJ). To derive a measure of overall library stability from flow cytometric data, temperatures of half-maximal irreversible denaturation ($T_{1/2}$) were calculated for each distinct library as described by Hasenhindl et al. (2013)[35]. The mean fluorescence intensity (MFI) of anti-Xpress positive yeast cells displaying wild-type Fc on the cell surface, stained with fluorescence-labeled FcγRI and incubated at 0° C was set to 100% relative MFI, while the lowest MFI of a measurement series was set to 0%. Resulting data were fitted according to

$$y = \frac{a}{(1+e^{\frac{-(x-x_0)}{b}})}$$

where x corresponds to the respective incubation temperatures, y corresponds to the relative MFI after heat incubation, a and b correspond to the maximum and minimum y values as defined by the model and x_0 corresponds to $T_{1/2}$ [35].

3.2.4.2 Data analysis for enrichment in antigen binding

Resulting data from enrichment experiments were analyzed with the FlowJo software. Here, the gate for correct morphology and the subordinate gate for single cells were set based on one arbitrary unstained control for all samples. The Xpress+/HER3+ gate was then drawn for each library and sorting step on the basis of the respective double control (aXpress*APC and SA/NA*PE).

3.3 Naïve selections

For *de novo* sorting of the libraries, a combination of magnetic bead cell sorting (MACS) and fluorescence activated cell sorting (FACS) was performed.

MACS, with the ability to sort $\geq 10^{10}$ yeast cells[36], is a high-through-put method that can quickly deplete complete non-binders and incorretly folded clones from a naïve library. Its major advantage lies in the possibility to eliminate many non-functional mutants from libraries with very high complexity, but nevertheless include correctly processed molecules with low affinity.
This is due to utilization of avidity effects[36] during the magnetic bead selection process, where a single cell is displaying multiple copies of a weak binder is also retained in the magnetic column upon staining with streptavidin-coated microbeads. Thus, the multiplicity of cells processing functional proteins and hence the overall discovery efficiency can be increased. Since MACS is a method of low stringency, it is performed prior to FACS experiments.

For isolation of higher-affinity binders *fluorescence activated cell sorting* (FACS) has to be applied. This method is more precise as cells can be gated for correct morphology, singularity and presence of the expression tag.
During this phenotypic selection process, a narrow stream of liquid is divided into single drops at a fixed distance upon travel through a vibrated nozzle. Each droplet ideally contains one single cell, which is scanned by a laser. Thereby, two kinds of information are gained. On the one hand, the light that is scattered by the cells can be used to determine size and morphology; while on the other hand, subpopulations can be separated by excitation of fluorescent dyes that are bound to tagging antibodies or ligands. Sorting of cells with the desired signals is accomplished by electrical charge. As determined by the set selection gates, charge is applied to the stream during formation of the drop containing a distinct cell. A charged drop forms which is subsequently deflected by electrodes into a sample tube. Drops with no cell, double cells or cells lacking the desired tags are discarded into a waste tube.

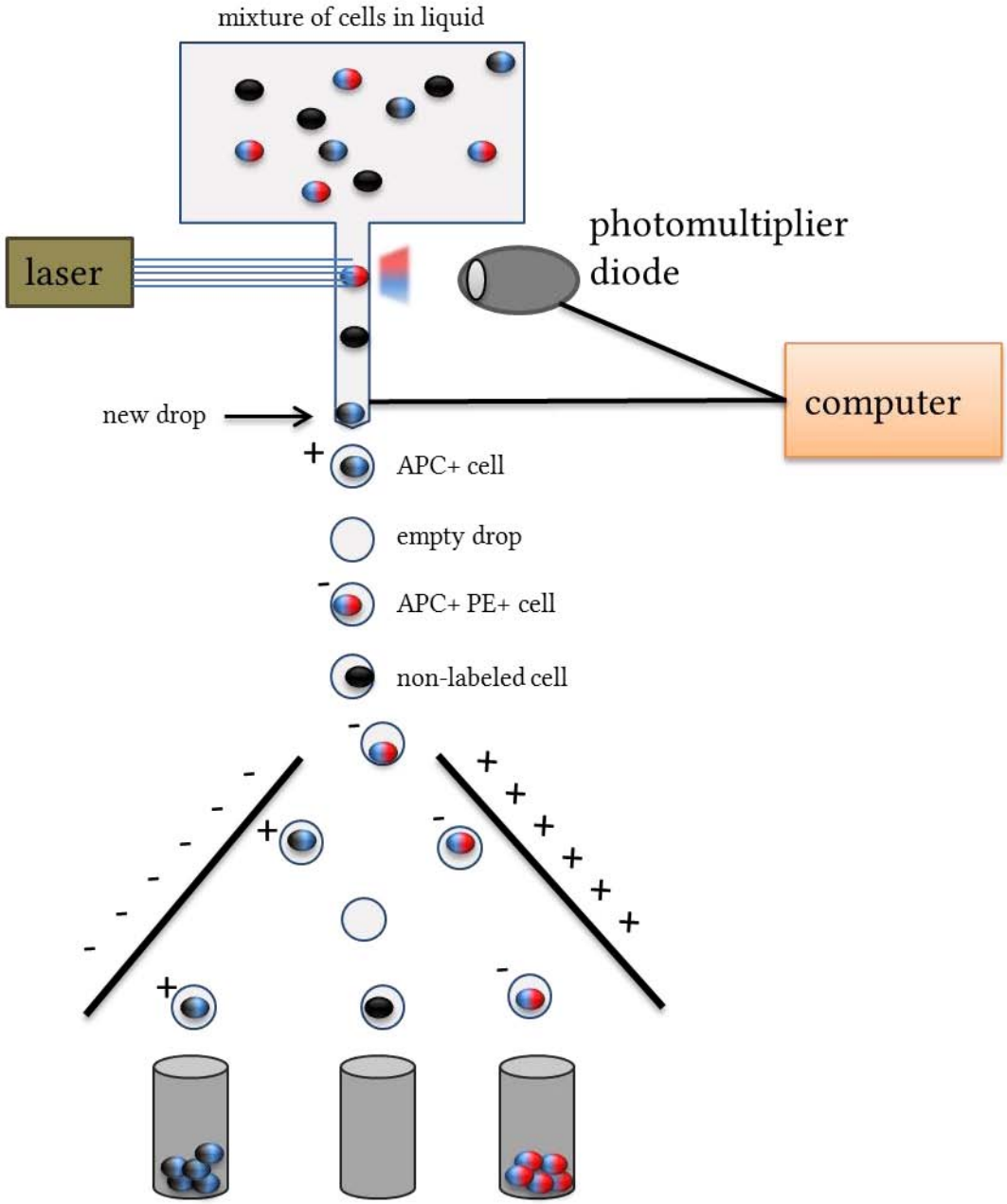


Figure 3-2 Fluorescence activated cell sorting (Adapted from Department of Biology, Davidson College, Davidson, NC 28036, 2001) Schematic representation of a fluorescence activated cell sorter. From all cells, fluorescently labeled with aXpress*APC and streptavidin/neutravidin*PE, displaying cells that bind the biotinylated antigen (APC+ PE+) can be separated from cells that display clones without antigen binding (APC+), and cells that do not express any protein on their surface.

3.3.1 Yeast display cultures for sorting

Pre-culture MACS

Per library, 250 mL of SDCAA medium in a 1 L sterile shake flask was inoculated with 1 mL dense cryoculture for sorting and incubated over night at 30 °C, shaking at180 rpm.

Pre culture FACS

Per library, 19 mL of SDCAA medium in a 100 mL sterile shake flask was inoculated with 1 mL dense cryoculture for sorting and incubated over night at 30 °C, shaking at 180 rpm.

Induction of surface expression

In order to remove glycerol from the cryoculture from the culture and ensure a similar growth phase, OD_{600} was adjusted to 1 in SDCAA, followed by incubation for another 4 h at 30 °C, shaking at 180 rpm.

As described above, the surface expression of the Fcab molecules was induced by adjusting the OD_{600} to 1 in SGRCAA. Again, the induction temperature was 20 °C. The cultures were incubated overnight, shaking at 180 rpm.

3.3.2 MACS - Magnetic bead cell sorting

All centrifugation steps were carried out for 5 minutes and 4 °C at 2000 g.

After OD_{600} measurement of the induced overnight cultures, the volume containing 5 x 10^9 cells was harvested by centrifugation and washed with 2.5 mL PBS/BSA. Next, the pellet of each library was stained in 2.5 mL staining solution (300 nM biotinylated HER3) and incubated room temperature and shaking for 45 min. Stained cells were incubated on ice for 15 min, washed in 10 mL PBS/BSA and resuspended in 3.5 mL PBS/BSA with 100 µL streptavidin microbeads (Miltenyi Biotec, Bergisch Gladbach, DE). During the following 15 min of incubation at 4 °C, LS columns (Miltenyi Biotec, Bergisch Gladbach, DE) were loaded into the magnet and pre-treated by gravity-flow of 3 mL cold PBS/BSA. Stained cells were then resuspended in 20 ml PBS/BSA and filtered through a cell strainer (40 µm pore size, Falcon, Corning, NY) before application to the LS column. The cell strainer was washed with additional 5 mL PBS/BSA. From the resulting 25 mL cell suspension, 100 µL were diluted serially and plated onto selective SD-CAA plates for calculation of the input titer.

Cells were loaded onto the LS columns in portions, with intermediate reorientation of the magnetic beads within the columns. The cells were washed once with 3 ml PBS/BSA and subsequently eluted by adding 7 mL PBS/BSA with use of a plunger to push all remaining cells into a 15ml falcon. Again, a 100 µL aliquot was taken out for each library, diluted serially and plated onto selective SD-CAA plates for calculation of the output titer.
Eluted cells were harvested by centrifugation and resuspended in 20 ml SD-CAA supplemented with 1:100 Penicillin/Streptomycin. The cultures were incubated overnight at 30 °C, shaking at 180 rpm.
On the following day, cryocultures of the sorted library cultures were prepared.

3.3.3 FACS - Fluorescence activated cell sorting

Cell sorting experiments were done using a FACSAria (BD, Franklin Lakes, NJ), with calibrated laser beam path, sensitivity and drop delay of the instrument. Data were collected in *log mode* and PBS/BSA was used as running buffer. During the sortings, the event rate was set to maximal 30000/s. Sortings were carried out in *enrich* or *purify* mode. *Enrich* means that in case two events, one positive and one negative, cannot be separated, both are collected, while in the *purify* mode, this would only be tolerated if both events are positive.
Sorted cell were collected into 5 mL round bottom tubes with 2 mL SDCAA medium complemented with Pen/Strep and grown overnight at 30 °C, shaking at 180 rpm. The following day, the cultures were either induced for further sorting rounds or cryopreserved.

3.3.4 Sequence analysis

As described in 3.1.6, plasmid DNA from yeast cultures of the different sorting steps were isolated and transformed into *E. coli.* The reverse three frames of the Sanger-sequenced clones were translated using the *EMBOSS Transseq* tool (http://www.ebi.ac.uk/Tools/emboss/). After that, the sequences containing the motif TYRVVS, whose presence indicates a correct IgG1-Fc sequence, were aligned with the *Jalview* software. The percentage of sequences without frameshifts or point mutations was calculated. For correct clones the amino acid sequences of the altered AB and EF loops were analyzed in *Excel.*

4 Results and Discussion

The following chapter describes the outcomes of the way towards isolation of novel binding molecules of the human epidermal growth factor receptor 3 (HER3).
First, three different variant libraries based on the Fc fragment of the human immunoglobulin G1 were constructed including the establishment of an optimized transformation protocol of the EBY100 strain of *S. cerevisiae*.
The functional presentation of the protein variants on the yeast surface was validated and the temperatures of half-maximal-irreversible denaturation ($T_{1/2}$) of each library were determined.
After validation and characterization, the libraries were screened for binders of the antigen HER3 employing magnetic bead cell sorting (MACS) and fluorescence activated cell sorting (FACS). Sequence analysis of last sorting steps repeatedly showed an enrichment of "sticky" clones with high abundance of the positively charged amino acids lysine and arginine in their antigen binding loops. This resulted in the proposal of distinct ideas for the improvement of strategies for *de novo* selections of yeast surface display libraries.

4.1 Generation of the libraries EF-CARA1, EF-CARA2 and UP+0_RW based on human IgG1-Fc

The yeast *S. cerevisiae,* used in this study, represents a valuable tool for directed evolution of antibody fragments because of its easy handling, cultivation and eukaryotic protein folding machinery. As described in the introduction (see 1.3), the presentation of antibodies on the yeast surface is a well-established and successfully employed method. The introduction of antigen binding sites into IgG1-Fc as well as different strategies for stability engineering of this molecule have been described previously.
Variant libraries with differing designs of the antigen binding EF loop were constructed in the present thesis. The amino acid distributions for the randomization is based on previous work. The aim was the evaluation of the applicability of a library (EF-CARA1) with a stability-biased amino acid distribution in the antigen binding loops for an efficient discovery of Fcabs.

4.1.1 Library design of EF-CARA1

For library generation, site-saturation mutagenesis of the loops via codon-based randomization was chosen. For EF-CARA1, the EF loop was designed by Hasenhindl et al. (not yet published) based on the information gained from several heat-shock sorting rounds against the conformational epitope of FcγRI/CD64 in combination with sequencing. Figure 4-1 shows the complete amino acid distribution for the EF loop, calculated from these pre-experiments. It becomes obvious that residues that have a stabilizing effect on a certain position show a higher percentage of occurence at that position, e.g. the wild type amino acid tryptophan at position 417 (W417) is enriched to a total of approx. 28% after two sorts. As presence of stop codons or cysteines (sites of disulfide bond formation) would increase the number of dysfunctional or incorrectly folded clones, respectively, both were excluded by design.

			EF loop									
	EF_CARA1 DESIGN		D413	K414	S415	R416	W417	Q418	Q419	G420	N421	V422
negatively charged	D	Asp	**19.3**	1.3	**26.8**	8.0	3.0	5.8	4.8	5.8	5.0	3.3
	E	Glu	9.5	0.9	7.9	8.5	2.8	5.1	7.3	6.6	3.9	6.3
positively charged	H	His	5.5	3.7	4.4	4.6	2.7	4.6	4.6	5.4	5.0	3.2
	K	Lys	3.3	15.0	3.9	5.6	2.0	5.4	5.2	4.8	9.9	4.2
	R	Arg	3.2	11.4	3.4	**11.4**	2.5	5.5	5.2	4.5	12.4	3.4
polar	N	Asn	9.1	**23.0**	7.1	7.7	3.8	8.0	8.2	13.3	**15.8**	**23.2**
	Q	Gln	6.2	6.0	4.0	7.0	2.6	**12.4**	**13.6**	4.7	9.4	4.0
	S	Ser	5.8	6.0	9.3	4.7	2.5	5.3	5.5	5.5	5.8	5.3
	T	Thr	3.6	4.4	4.8	4.3	3.4	6.2	5.6	4.8	7.6	10.0
hydrophobic	A	Ala	5.6	4.3	5.2	6.8	4.6	5.2	5.4	5.2	5.6	3.1
	F	Phe	1.8	1.8	1.7	2.7	5.7	3.9	4.8	2.3	0.7	2.2
	I	Ile	1.3	2.1	1.9	3.9	5.7	2.9	2.7	1.1	1.6	4.6
	L	Leu	2.7	2.7	2.0	3.9	10.2	5.8	4.1	2.4	1.7	3.1
	M	Met	2.5	3.6	2.5	4.8	5.8	5.7	4.2	3.7	3.3	3.2
	V	Val	1.6	2.4	2.4	4.6	4.9	3.8	2.8	2.1	2.7	15.7
	W	Trp	2.2	3.3	3.4	1.6	**27.8**	4.0	4.0	2.0	0.9	1.1
	Y	Tyr	2.2	3.4	2.2	2.9	5.6	3.7	3.3	2.8	1.0	1.2
special	C	Cys	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
	G	Gly	3.7	2.1	4.0	3.7	2.9	5.0	5.6	**19.6**	6.0	0.9
	P	Pro	10.7	2.2	3.0	3.3	1.4	1.9	2.9	3.4	1.7	2.1
	del		0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
	SUM		100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0

Figure 4-1 Library design EF CARA1 – in % total fold change. Amino acid distribution calculated from fold-change after two heat-shock selections with FcγRI/CD64.

			EF loop									
	EF_CARA1 naïve output		D413	K414	S415	R416	W417	Q418	Q419	G420	N421	V422
negatively charged	D	Asp	**18.6**	0.0	**40.0**	11.4	2.9	2.9	4.3	11.4	1.4	4.3
	E	Glu	10.0	1.4	7.1	12.9	4.3	10.0	8.6	11.4	4.3	10.0
positively charged	H	His	4.3	8.6	1.4	8.6	5.7	1.4	5.7	4.3	10.0	4.3
	K	Lys	2.9	5.7	1.4	7.1	1.4	0.0	2.9	2.9	7.1	0.0
	R	Arg	5.7	7.1	1.4	**8.6**	1.4	4.3	4.3	7.1	5.7	1.4
polar	N	Asn	5.7	**24.3**	4.3	5.7	5.7	7.1	10.0	8.6	**15.7**	**22.9**
	Q	Gln	10.0	4.3	7.1	2.9	2.9	**11.4**	**21.4**	2.9	10.0	5.7
	S	Ser	5.7	11.4	5.7	5.7	4.3	4.3	5.7	5.7	5.7	4.3
	T	Thr	1.4	2.9	4.3	2.9	1.4	10.0	4.3	4.3	8.6	7.1
hydrophobic	A	Ala	1.4	1.4	5.7	7.1	1.4	5.7	2.9	5.7	11.4	1.4
	F	Phe	1.4	2.9	2.9	5.7	5.7	4.3	4.3	5.7	0.0	1.4
	I	Ile	1.4	0.0	1.4	4.3	8.6	1.4	2.9	1.4	2.9	2.9
	L	Leu	4.3	4.3	1.4	1.4	7.1	1.4	1.4	2.9	0.0	2.9
	M	Met	0.0	7.1	2.9	1.4	8.6	1.4	4.3	1.4	2.9	1.4
	V	Val	1.4	1.4	2.9	4.3	4.3	10.0	4.3	2.9	4.3	24.3
	W	Trp	1.4	7.1	0.0	1.4	**21.4**	2.9	4.3	0.0	4.3	1.4
	Y	Tyr	5.7	5.7	2.9	4.3	7.1	2.9	0.0	0.0	0.0	1.4
special	C	Cys	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
	G	Gly	4.3	4.3	4.3	1.4	4.3	12.9	1.4	**18.6**	5.7	1.4
	P	Pro	14.3	0.0	2.9	2.9	0.0	1.4	2.9	0.0	0.0	1.4
	del		0.0	0.0	0.0	0.0	1.4	4.3	4.3	2.9	0.0	0.0
	SUM		100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0

Figure 4-2 Output of sequencing in % from 79 sequences. % occurrence of a distinct amino acid at each position of the EF loop

Comparing the library design to the naïve sequencing output, the desired amino acid distribution could be approximately achieved. The respective amino acid with the highest percentage at each position of the design was also the one with the highest percentage calculated from 79 sequences of the naïve sequencing output (Figure 4-2).

4.1.2 Generation of libraries via homologous recombination

In the current work, gap repair-driven homologous recombination by *S. cerevisiae* was the basis of library construction. Regions of homology (28 - 39 bp) and the randomized loop sequence were added to template fragments by PCR (see 3.1.3) The resulting insert fragments carrying the partially altered IgG1-Fc gene then reconstituted the gapped vector, which was used for transformation of *S. cerevisiae.* Successfully transformed yeast were selected in a special cultivation medium over the auxothrophic marker gene *TRP1,* enabling transformants to perform *de novo* synthesis of the essential amino acid tryptophan, which is absent in the growth medium.

4.1.2.1 Production and purification of the pYD1 vector backbone and inserts

Plasmid DNA was amplified in *E. coli*, purified and digested with the restriction endonuclease BsmBI (see 3.1.3.1). Correct restriction digest was verified and purification was done in a preparative scale by agarose gel electrophoresis (see 3.1.3.3). Results are shown in Figure 4-3, A.

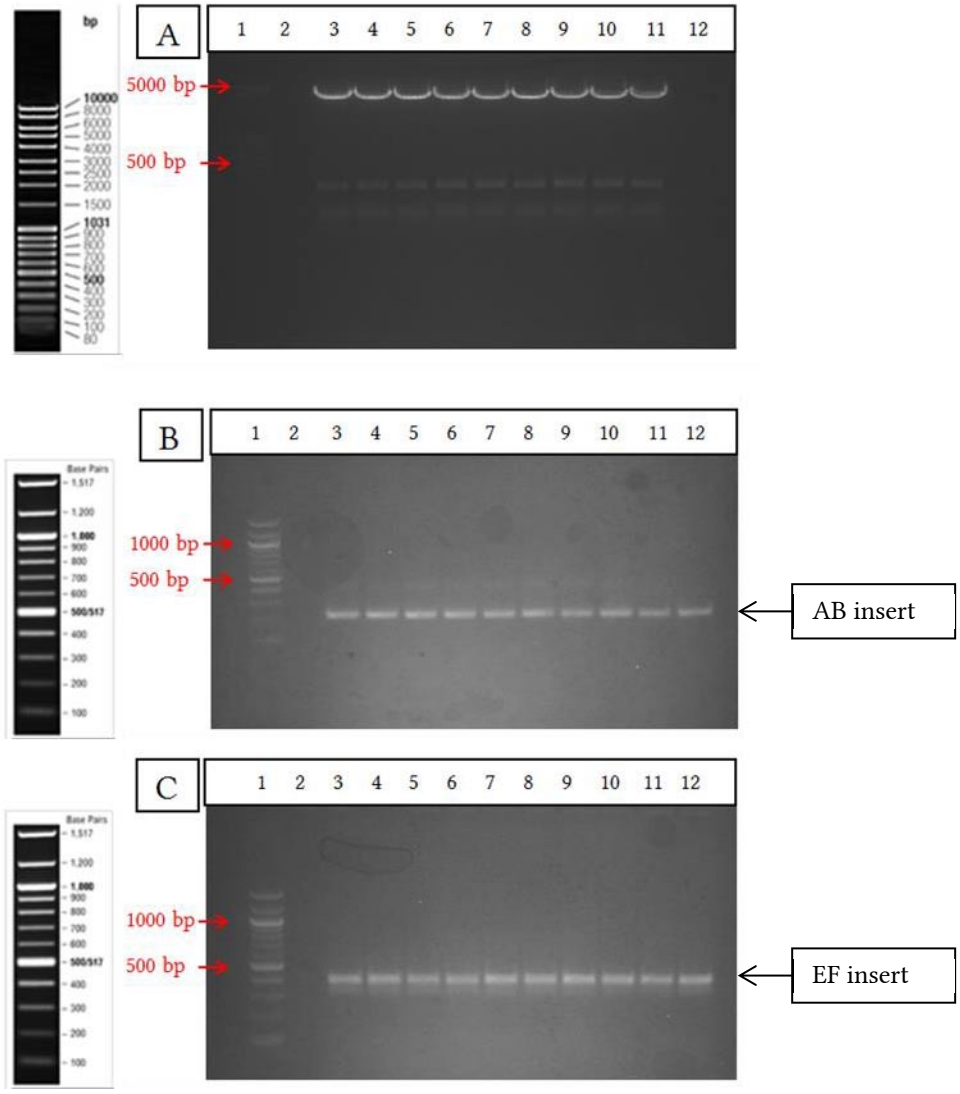


Figure 4-3 Agarose gel electrophoresis A: lane 1 = ladder, lanes 3-11 = vector digest, B: lane 1 = ladder, lanes 3-12 = AB insert PCR, C: lane 1 = ladder, lanes 3-12 = EF insert PCR.

According to the DNA ladder in lane 1, high-intensity bands of approx. 5000 bp corresponded to the vector backbone and were cut out for subsequent gel extraction, further purification and concentration by ethanol precipitation, and spectrophotometric quantification (see 3.1.3.3).

Similarly, PCR results for template and insert fragment generation were verified and purified via agarose gel electrophoresis. Figure 4-3, B and C show examples of the successful generation of the AB and EF insert fragments (195 bp and 379 bp). Following gel extraction, inserts were again purified and concentrated by ethanol precipitation, and quantified spectrophotometrically.

4.1.2.2 Transformation by electroporation

Various methods are applied for yeast transformation, e.g. the lithium method, spheroplast transformation or glass bead transformation[37]. In this thesis, *S. cerevisiae* EBY100 were transformed by electroporation (see 3.1.4).

A first attempt, performed according to the protocol of Benatuil et al. (2010)[38] resulted in unsatisfactory transformation efficiencies (= transformants/ µg gapped vector DNA). Next, different experimental setups with varying amounts of DNA were tested, resulting in the establishment of an improved transformation protocol. This is of high interest in the directed evolution protein engineering approach, as the size of a library is important for an efficient discovery process.

These 1-fold transformation (3.1.4) trials, were performed using the insert fragments for EF-CARA1. These cultures were pooled resulting in EF-CARA1. A 2-fold transformation with the determined optimal amounts of DNA (see 4.1.2.3) was performed for EF-CARA2 and UP+0_RW, each. Generated transformants were determined by plating dilutions on SD-CAA plates. The final size was 1 x 10^8 for all three libraries.

4.1.2.3 Optimization of the transformation efficiency

The initial situation and reason for the optimization of the electroporation process was the low transformation efficiency of the protocol described by Benatuil et al. (2010), at a very high DNA demand. The authors postulated the need of a weight ratio of insert:vector of 3:1 in order to gain 1.5 x 10^8 transformants/µg vector. No information was provided regarding vector reconstitution by integration of more than one insert fragment. For each library, 12 µg of each insert were needed for a 1-fold transformation using 4 µg gapped vector. These 1-fold transformations repeatedly resulted in only 1.25 x 10^6 transformants/µg vector DNA, which is 1:100 of what was suggested by Benatuil et al. As the inserts in the PCR reaction differed in the number of basepairs (AB insert, 195 bp and EF insert, 379 bp), the molar

concentration was accordingly different when using the same weight. Moreover, production and purification of these amounts of DNA are labourious both in terms of material and man-hours.

Thus, we hypothesized that for an efficient vector reconstitution the equimolar use of inserts at a certain molar excess over the vector would be sufficient. Also, we lowered the electroporation voltage from 2.5 kV to 0.5 kV and changed the pulse length to 15 ms (according to Chen et al. (2013)[33] and M. Traxlmayr, K. Dane Wittrup lab (MIT), personal communication).

So far, the only optimization of DNA input for *S. cerevisiae* was done for the lithium method[39]. Therefore, we performed a test series with varying insert: vector ratios (Table 4-1, Figure 4-4). It was shown that a 20-fold molar excess of insert over the gapped vector is an optimal DNA input. This resulted in an approx. 13-fold increased transformation efficiency with only 50% of the original amounts of EF insert and 25 % of AB insert needed.

Table 4-1 Transformation efficiency at varying DNA input

Insert/vector molar ratio	gapped vector [µg]	AB insert [µg]	EF insert [µg]	resulting transformants/ µg vector DNA
5:1	4	0.71	1.38	$2.98*10^6$
10:1	4	1.42	2.76	$11.18*10^6$
20:1	4	2.84	5.51	$\mathbf{16.90*10^6}$
50:1	4	7.09	13.78	$9.03*10^6$

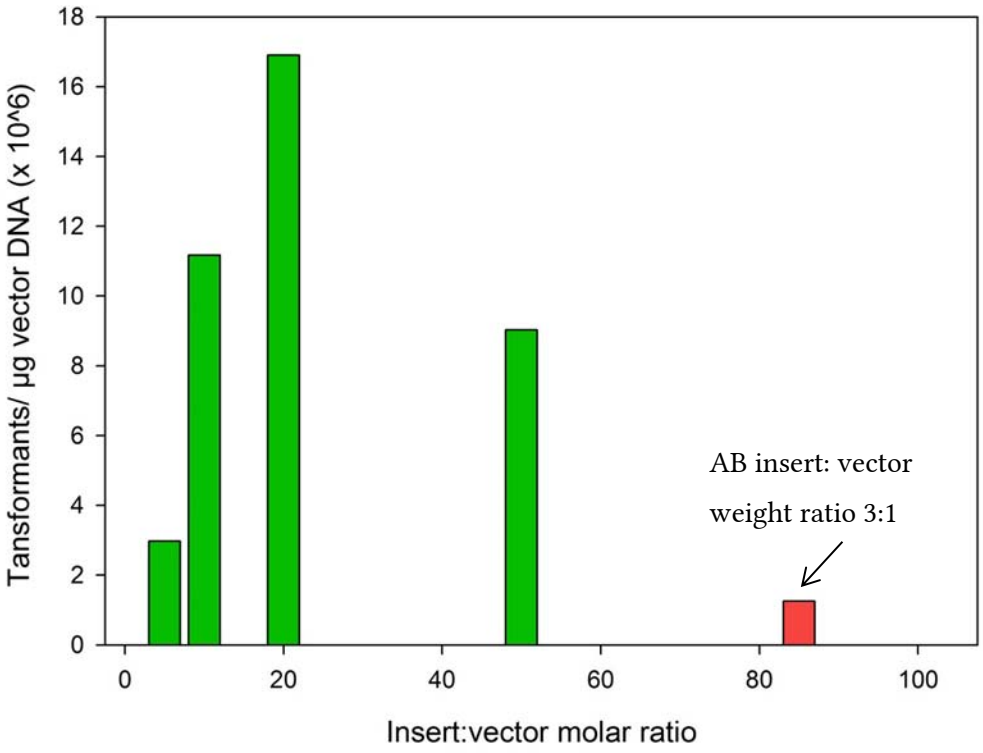


Figure 4-4 Differences in transformation efficiency at varying molar ratios insert: vector
Green bars show the results of the 1-fold test transformations, the red bar shows the resulting transformation efficiency when using a molar excess of AB insert conferring to a insert to vector weight ratio of 3:1 according to Benatuil et al.

4.2 Validation and characterization of the naïve libraries

Prior to the *de novo* selections of the libraries, the overall library quality and thermal stability was determined. For this purpose, pre-cultures of the naïve IgG1-Fc libraries in SD-CAA medium were diluted to an OD_{600} of 1 in SD-CAA, followed by a 4 h growth phase and subsequent surface induction through medium change to SGR-CAA (see 3.2.1).

Probing displayed IgG1-Fcs for their binding to conformational epitopes of FcγRI/CD64 and aC_H2 served as validation of structural integrity and correct folding on a library scale. Although both ligands bind to the C_H2 domain of IgG1-Fc, and not directly to C_H3, their binding is representative for the structural integrity of the C_H3 domain upon mutation, as there is a proven cross-talk between the two domains[24]. For this purpose, cultures were incubated with the ligands whose binding could be detected by flow cytometry (see 3.2.2.1).

Conduction of the same experiment with an additional heat-shock of the induced cultures (temperature range 50 -75 °C) gave information about the thermal stability and unfolding of the proteins in the form of *temperatures of half-maximal irreversible denaturation* ($T_{1/2}$).

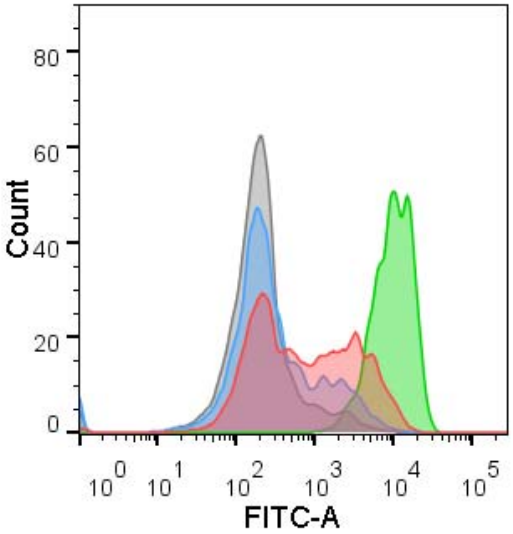
4.2.1 Binding to structurally specific ligands

Quality and correct processing of the yeast surface display libraries was assessed in comparison to non-heat shocked surface displayed wild type IgG1-Fc (Figure 4-5 and Figure 4-6).
Binding to aC_H2 and FcγRI/CD64 was detected by direct FITC-conjugation (aC_H2) or indirectly via an AF488-conjugated antibody against the His-Tag of FcγRI/CD64. Both the FITC and AF488 fluorescence signals are represented by FITC-A.

Experiments utilizing both aC_H2 and FcγRI/CD64 as structurally specific ligands showed that the overall quality of IgG1-Fc from EF-CARA1 is more wild type like than that of EF-CARA2, which again is superior to UP+0_RW (Figure 4-5: aC_H2, Figure 4-6: FcγRI/CD64). Upon analysis of the histograms and dot plots of a representative experiment, this improved quality can be assigned to two factors. First, the number of variants, that still bind to the structurally specific ligand is higher, meaning that less members of EF-CARA1 are impaired to a degree that prevents this ligand from binding. And secondly, the residual mean fluorescence intensity (i.e. percentage MFI with respect to the MFI determined for yeast

surface displayed wild type IgG1-Fc) of library members still binding to the structurally specific ligand is increased according to EF-CARA1 > EF-CARA2 > UP+0_RW, meaning that these respective library supposedly exhibit a more wild type-like structure (Dot plots Figure 4-5 and Figure 4-6).

These results give a first insight into the potential of the stability-biased amino acid distribution of the EF loop of the library EF-CARA1, whose quality is not only 3-5-fold better than that of library UP+0-RW (NNK-randomized EF loop), but also show an approx. 2-fold improvement over library EF-CARA2 (with a wild type bias in the EF loop).



	Sample Name	Count	Full Gating Path
	EF_CARA1_aCH2_ice.fcs	1923	Morphology/Single Cells/Xpress+
	EF_CARA2_aCH2_ice.fcs	1917	Morphology/Single Cells/Xpress+
	UP+0_RW_aCH2_ice.fcs	1978	Morphology/Single Cells/Xpress+
	Fc wt_aCH2_ice.fcs	*1554*	*Morphology/Single Cells/Xpress+*

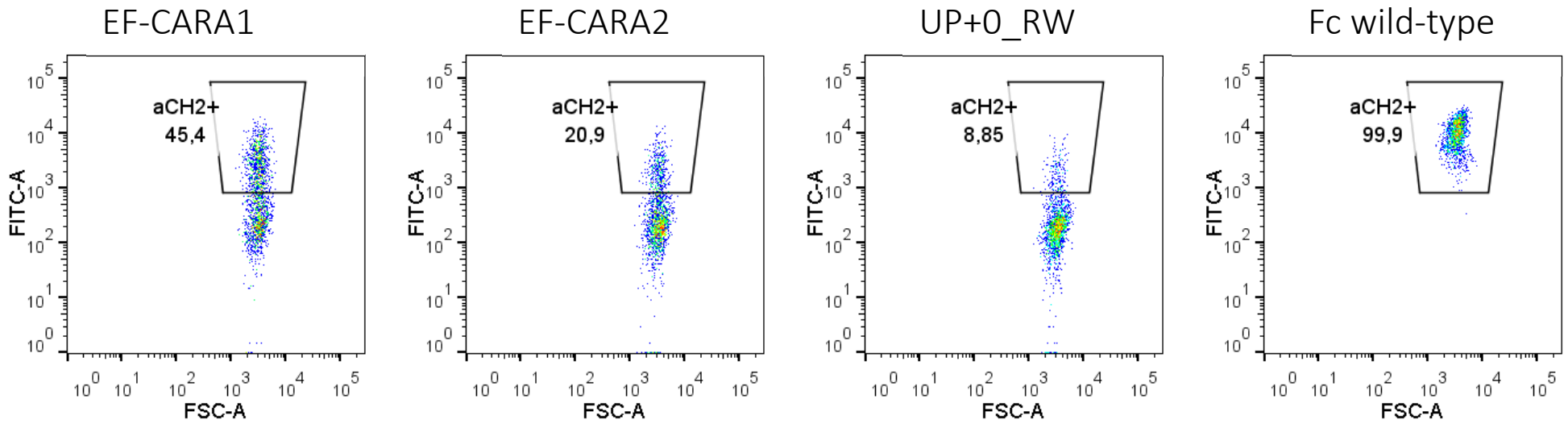
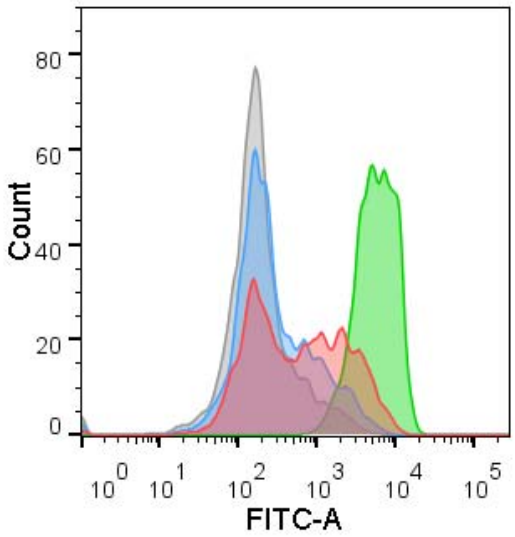


Figure 4-5 Binding to aC_H2 for the assessment validation of library quality. Upper figure: overlay of histograms; lower figure: dot plots. Data from one representative experiment are shown.



	Sample Name	Count	Full Gating Path
	EF_CARA1_CD64_ice.fcs	1992	Morphology/Single Cells/Xpress+
	EF_CARA2_CD64_ice.fcs	2264	Morphology/Single Cells/Xpress+
	UP+0_RW_CD64_ice.fcs	2288	Morphology/Single Cells/Xpress+
	Fc wt_CD64_ice.fcs	*1908*	*Morphology/Single Cells/Xpress+*

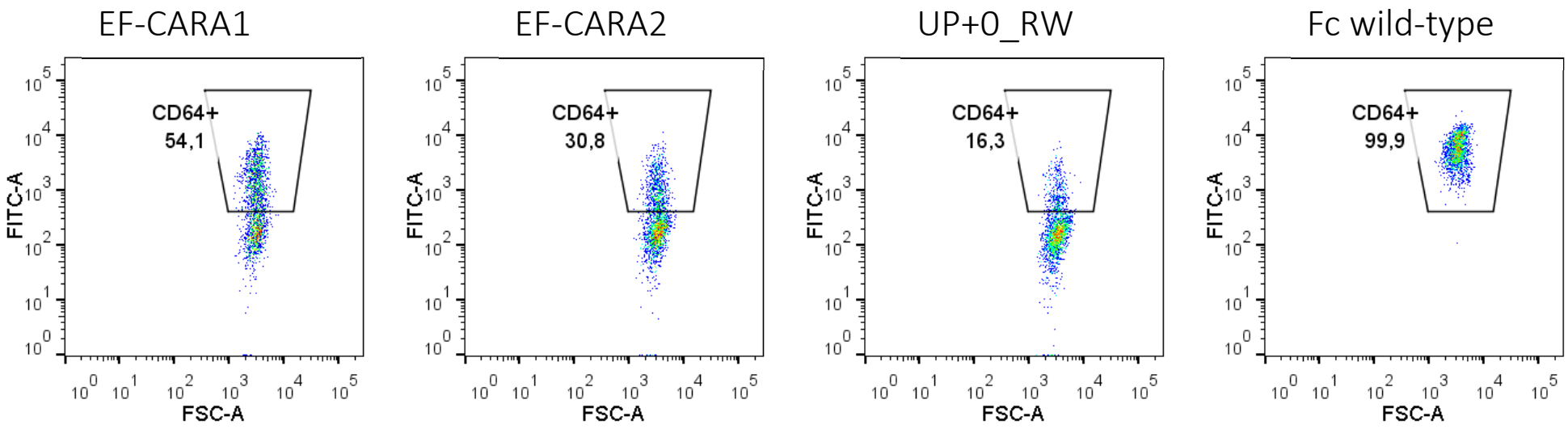


Figure 4-6 Binding to FcγRI/CD64 for assessment of ibrary quality. Upper figure: overlay of histograms; lower figure: dot plots. Data from one representative experiment are shown.

4.2.2 Thermal stability on a library scale – $T_{1/2}$ determination

The need for a quick way to assess the overall quality of a library recently led to the development of a new method for fast evaluation of the impact of a distinct loop design on the correct protein fold on a library scale[35,22]. Here, yeast cultures, induced for surface expression, are heat-shocked at different temperatures with subsequent testing of the displayed IgG1-Fc variants for residual binding to FcγRI/CD64 and aC_H2 (see Figure 4-7).

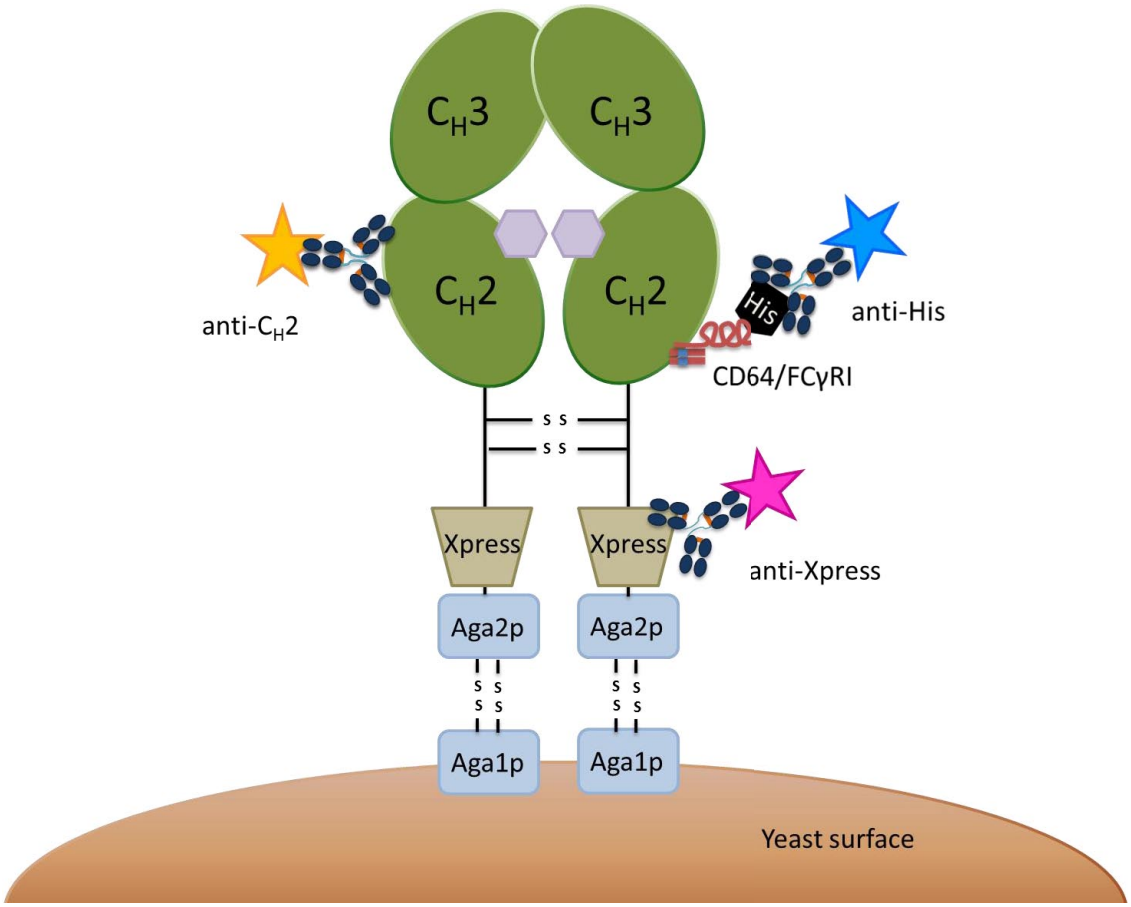


Figure 4-7 Yeast surface display setup for determination of thermal stability

The engineering scaffold IgG1-Fc (green) is fused to the Xpress tag (olive) for normalization. The N-terminal Aga2p as well as the cell-wall-linked Aga1p are depicted in light blue. Glycan structures on the C_H2 domain are shown in violet. The correct conformation of the scaffold is determined by staining either aC_H2*FITC (yellow) or the FcγRI/CD64*His and an aHis*AF488 (blue). Displaying cells are stained e.g. by an anti-Xpress*APC antibody (pink).

Data are analyzed and fitted as described in 3.2.4.1. The derived temperatures of half-maximal irreversible denaturation ($T_{1/2}$) of tested libraries have been shown to correlate strongly with DSC-derived mid-points of denaturation (T_m)[35] and can thus provide reliable information on the thermal stabilities of an entire library.

Table 4-2 shows the calculated $T_{1/2}$ values from the sigmoidal fits (Figure 4-8 and Figure 4-9) of the data points at each incubation temperature derived from the geometric mean of the fluorescence signal, corresponding to the binding of the structural markers aC_H2 and FcγRI/CD64 (see 3.2.4.1). The experiments were conducted in triplicates. However, residual binding to aC_H2 was calculated only from the first experiment, as too many outlier values were determined in the following two experiments.
The resulting $T_{1/2}$ values show an approx. decrease in thermal stability of 8-9 °C for EF-CARA1 and EF-CARA2 while $T_{1/2}$ of UP+0_RW is decreased by about 6 – 7.4 °C in comparison to the wild type IgG1-Fc.

Table 4-2 $T_{1/2}$ and Δ $T_{1/2}$ values (with respect to wild type IgG1-Fc)

IgG1-Fc library	aC_H2		FcγRI/CD64	
	$T_{1/2}$ [°C]	Δ $T_{1/2}$ [°C]	$T_{1/2}$ [°C]	Δ $T_{1/2}$ [°C]
EF-CARA1	54.99	-9.01	56.71	-8.28
EF-CARA2	54.96	-9.04	56.83	-8.16
UP+0_RW	57.95	-6.05	57.64	-7.35
wild type IgG1-Fc	64.00	/	64.99	/

Comparing these results to the validation of the libraries on ice we conclude that the application of a calculated (stability-biased) amino acid distribution does improve the overall library quality, but not the overall library stability. Nevertheless, calculated amino acid distributions are thought to enhance the discovery of stable clones with high antigen affinity, as the pool of well folded clones to select from is extended considerably.

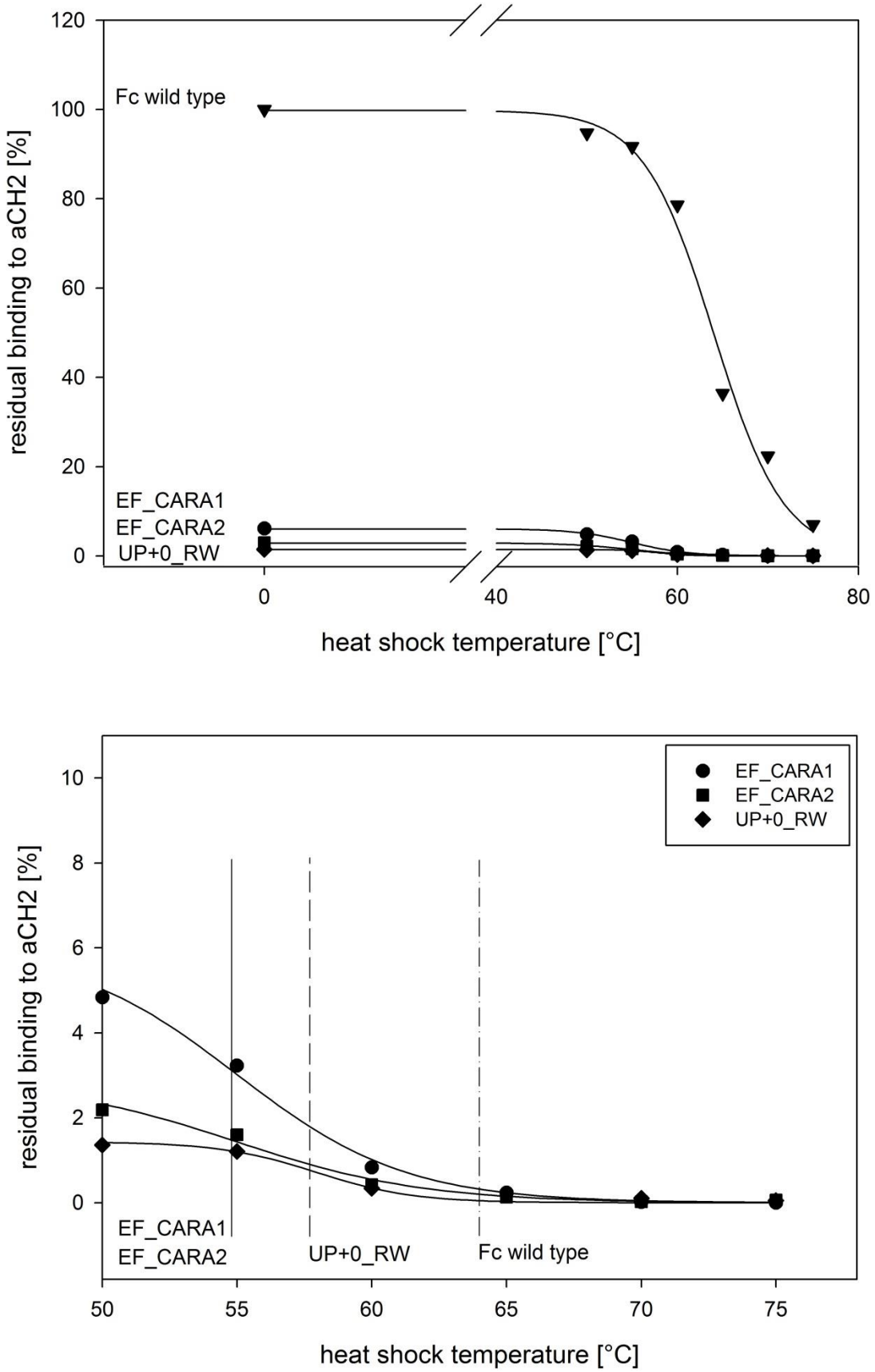


Figure 4-8 Thermal stabilities from yeast surface display of naïve libraries assessed with aC_H2. Residual binding to the structural marker at heat-shock temperatures ranging from 0 to 75 °C in comparison to surface displayed wild type IgG1-Fc

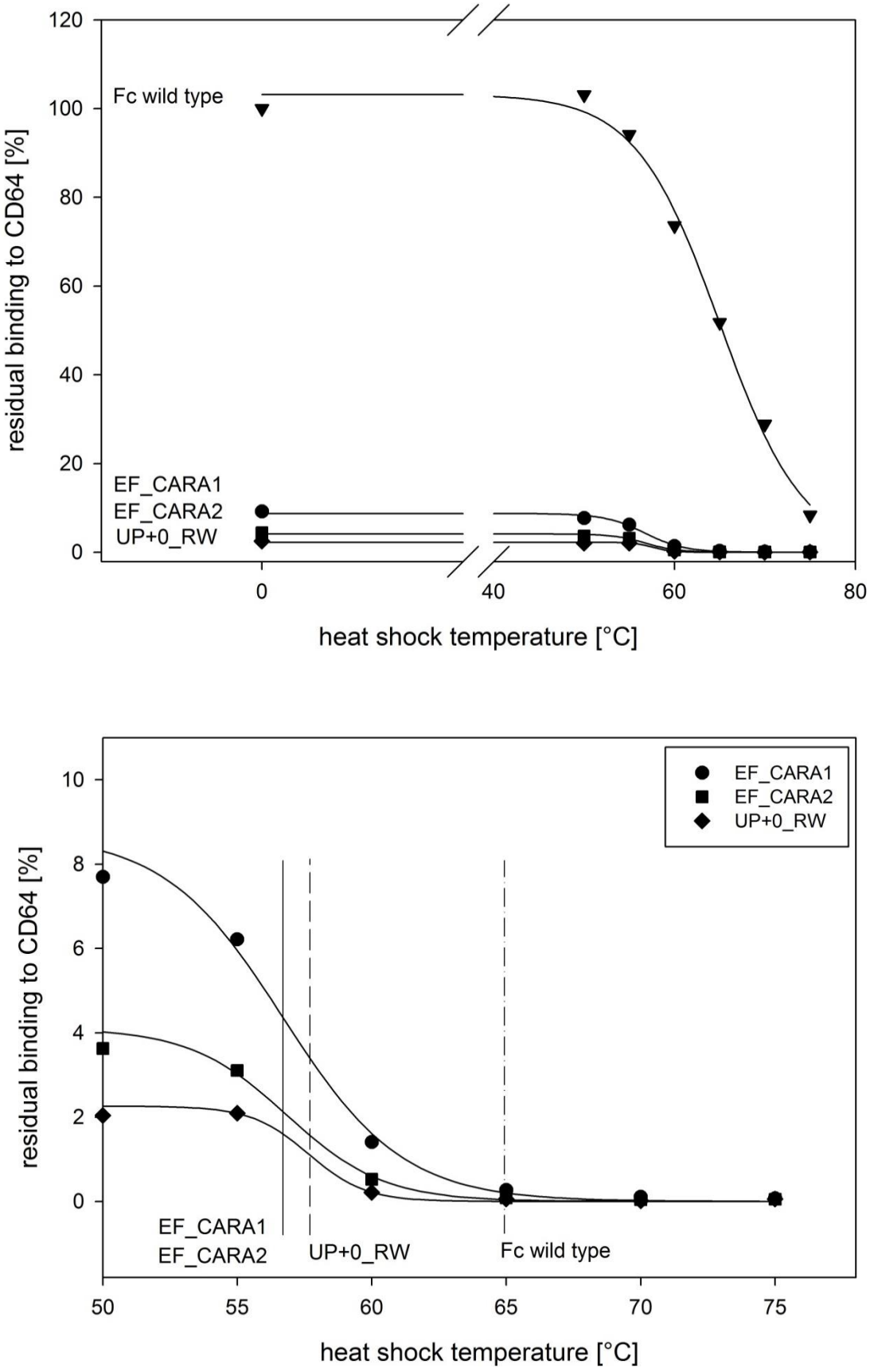


Figure 4-9 Thermal stabilities from yeast surface display of naïve libraries assessed with FcγRI/CD64. Residual binding to the structural marker at heat-shock temperatures ranging from 0 to 75 °C in comparison to surface displayed wild type IgG1-Fc.

4.3 Naïve selections

In order to prove the concept of stability-biased amino acid distributions in the design of IgG1-Fc libraries, *de novo* selections against the therapeutically relevant antigen HER3 (see 1.3.1) were performed using magnetic- and fluorescence activated cell sorting (see 3.3). Magnetic bead cell sorting relied on the ability of antigen binders of any affinity to stick to a magnetic column (for subsequent elution) upon staining with biotinylated HER3 (3.2.3.2) and streptavidin-coated microbeads (3.3.2), while incorrectly folded clones were washed away. Fluorescence activated cell sorting is a more precise method, where sorted cells were gated for correct morphology, singularity and the presence of the Xpress expression tag (see 3.2.4).

Currently, the *de novo* selections are ongoing due to several problems in the course of the project.
Sequencing results revealed that so called "sticky" binders were enriched after the third round of FACS in the course of the first campaign. Such IgG1-Fc molecules show a high abundance of the positively charged amino acids lysine and arginine in the mutated loops. As sequences of MACS and the first round of FACS seemed to be free of "sticky" clone enrichment, we started a second campaign from the populations sorted in the first round of FACS. Again, after the third round of campaign 2, the majority of sequenced clones showed enrichment of lysines and arginines in the EF loop. As a consequence, a third campaign, starting from the MACS-sorted populations, is currently ongoing (Figure 4-10).

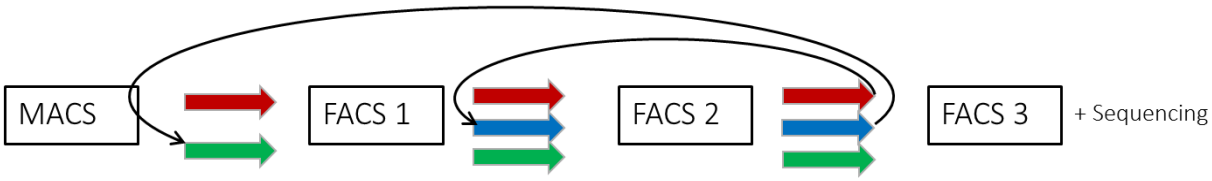


Figure 4-10 Flow scheme of the perfomed selection rounds and campaigns. Red: Campaign 1, blue: campaign 2, green: campaign 3, FACS1/FACS2/FACS3 = first/second/third round of FACS in the course of a campaign

However, the various experiences during these many selection rounds resulted in a number of hypotheses and postulation of strategies of how such naïve selection procedures might be improved in the future.

4.3.1 Enrichment in antigen binding and sequence analysis

Cryocultures from all sorting rounds of a campaign were tested for their antigen binding ability. For this purpose, cultures induced for surface display were incubated with 300 nM biotinylated HER3, stained with SA-PE and 10000 events analyzed on a flow cytometer (see 3.2.3). Exemplary, Figure 4-11, Figure 4-13 and Figure 4-15 show these enrichment measurements for the campaign 2. The percentage of enrichment in HER 3-binding ability is determined through comparison with a double control. For this control the same cell population was stained only with aXpress*APC and SA-PE and hence unspecific binding can be normalized. For all three libraries, an enrichment in antigen binding could be observed.

Sequencing of clones from the last sorting step was achieved by isolation of plasmid DNA from the respective yeast cultures via Zymoprep, followed by transformation of *E. coli* with the isolated plasmid DNA and picking of a certain number of clones to a microtiter plate which were sent for Sanger sequencing (see 3.1.6).
Figure 4-12, Figure 4-14 and Figure 4-16 show the results from cultures of the third round of FACS of campaign 2.
Both the NNK randomized AB loops as well as the EF loops with different amino acid distributions according to the library design, carry more than 3 lysines or arginines for the majority of the clones, which indicates the enrichment of undesired low-affinity IgG1-Fc variants (see 4.3.2.1). We hypothesize that this enrichment is due to the non-stringent sorting strategy (see Figure 4-20).

4.3.1.1 EF-CARA1

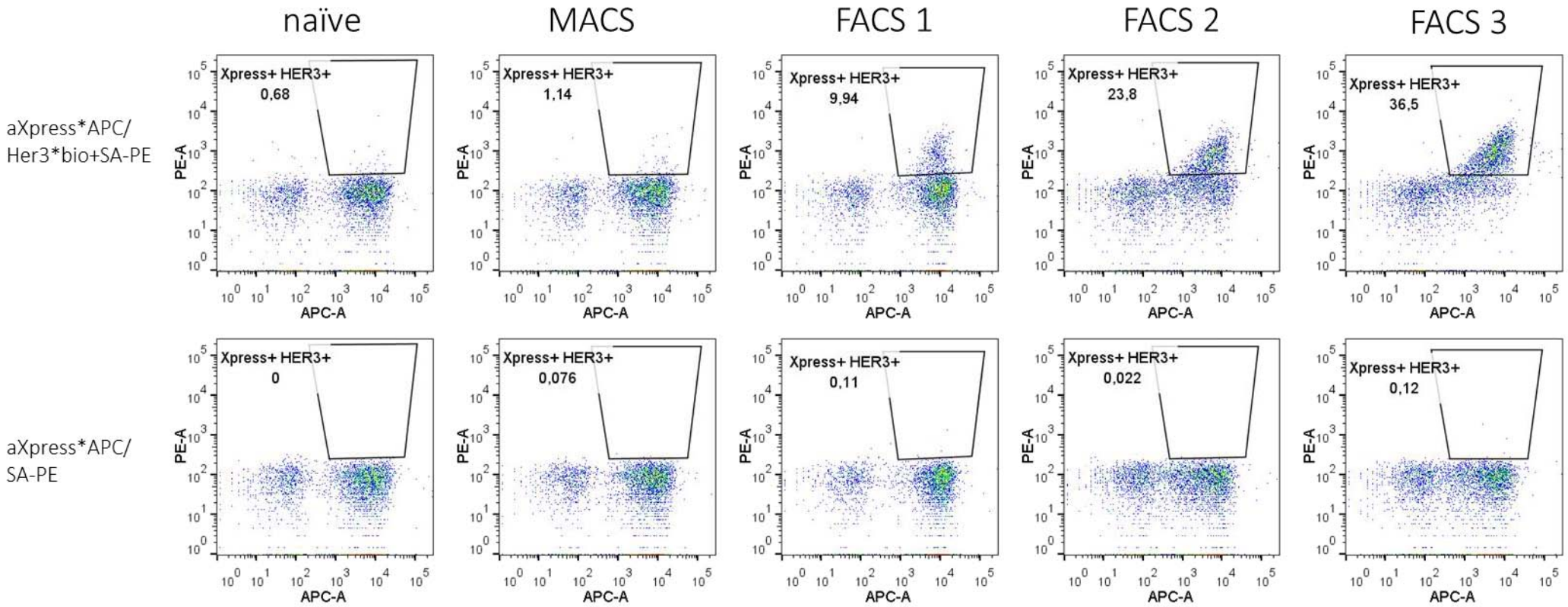


Figure 4-11 HER3 binding enrichment in the course of all sorting rounds of campaign 2 for EF-CARA1. Upper row: Samples stained with aXpress*APC and biotinylated HER3+SA-PE, lower row: double control stained with aXpress*APC and SA-PE for normalization of unspecific binding to the secondary stain (PE).

All plots show populations gated for correct morphology and presence as single cells.

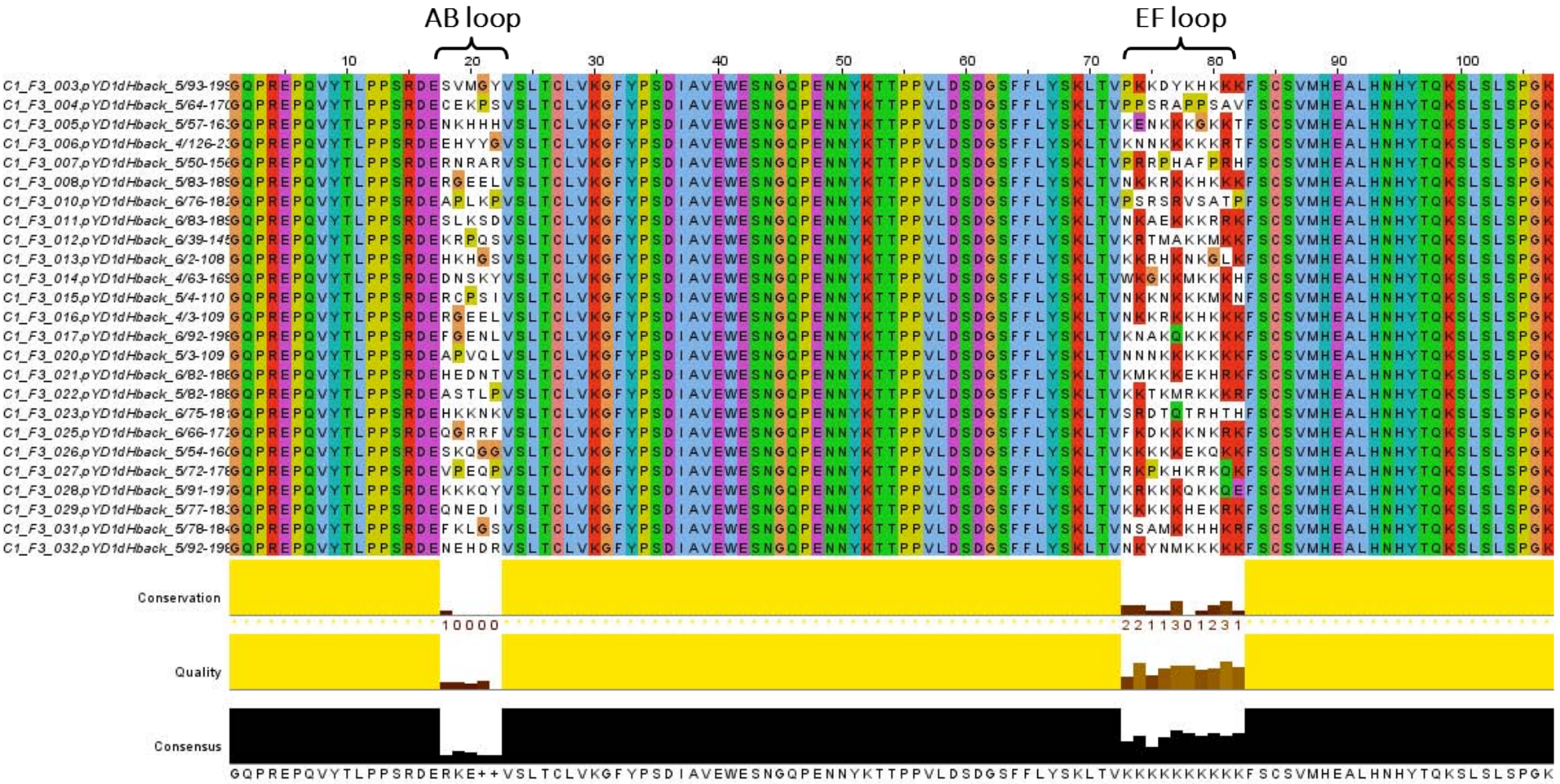


Figure 4-12 Sequence analysis of clones from cultures of the third round of FACS of campaign 2 for EF-CARA1. Translated nucleotide sequences that contain the motif TYRVVS and aligned using the software *Jalview*. Only sequences without frameshift or point mutations are shown, from the start of the C_H3 domain (GQPREP) to the C-terminus of IgG1-Fc (SLSLSPGK). Both AB loop and EF loop show enrichment in K = lysine and R = arginine.

4.3.1.2 EF-CARA2

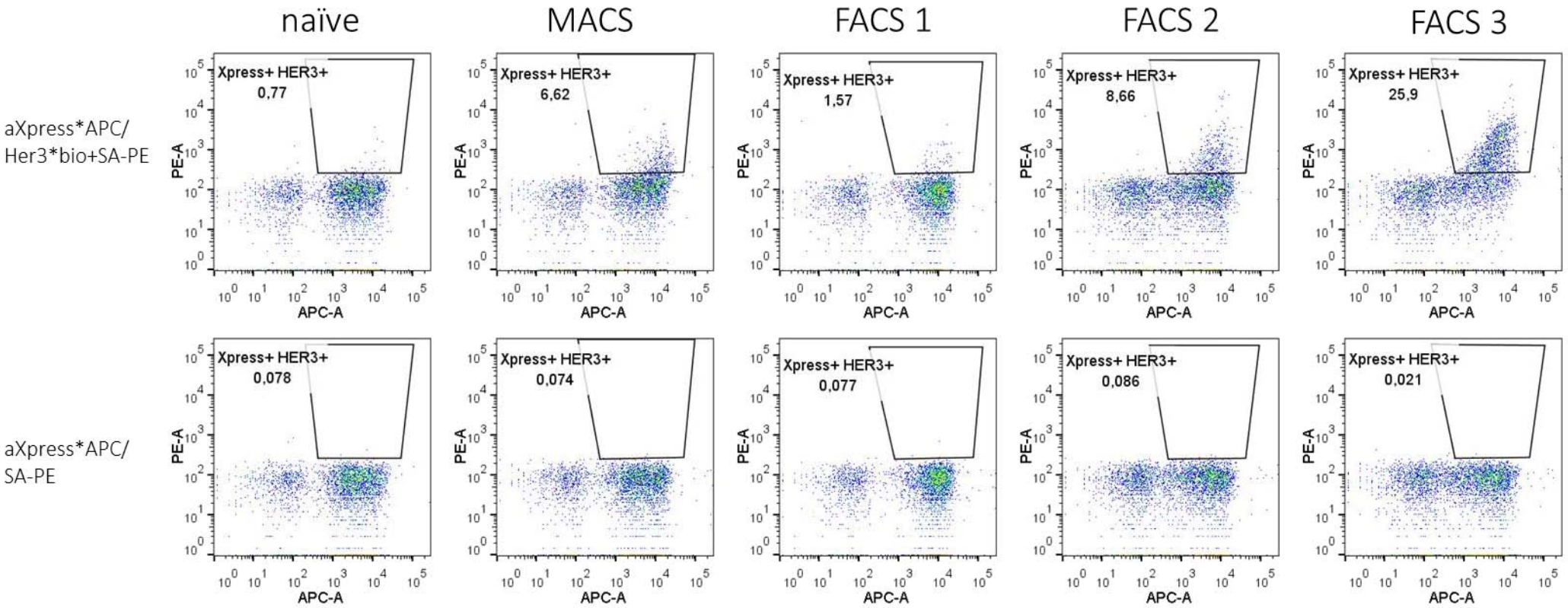


Figure 4-13 HER3 binding enrichment in the course of all sorting rounds of campaign 2 for EF-CARA2. Upper row: Samples stained with aXpress*APC and biotinylated HER3+SA-PE, lower row: double control stained with aXpress*APC and SA-PE for normalization of unspecific binding to the secondary stain (PE).

All plots show populations gated for correct morphology and presence as single cells.

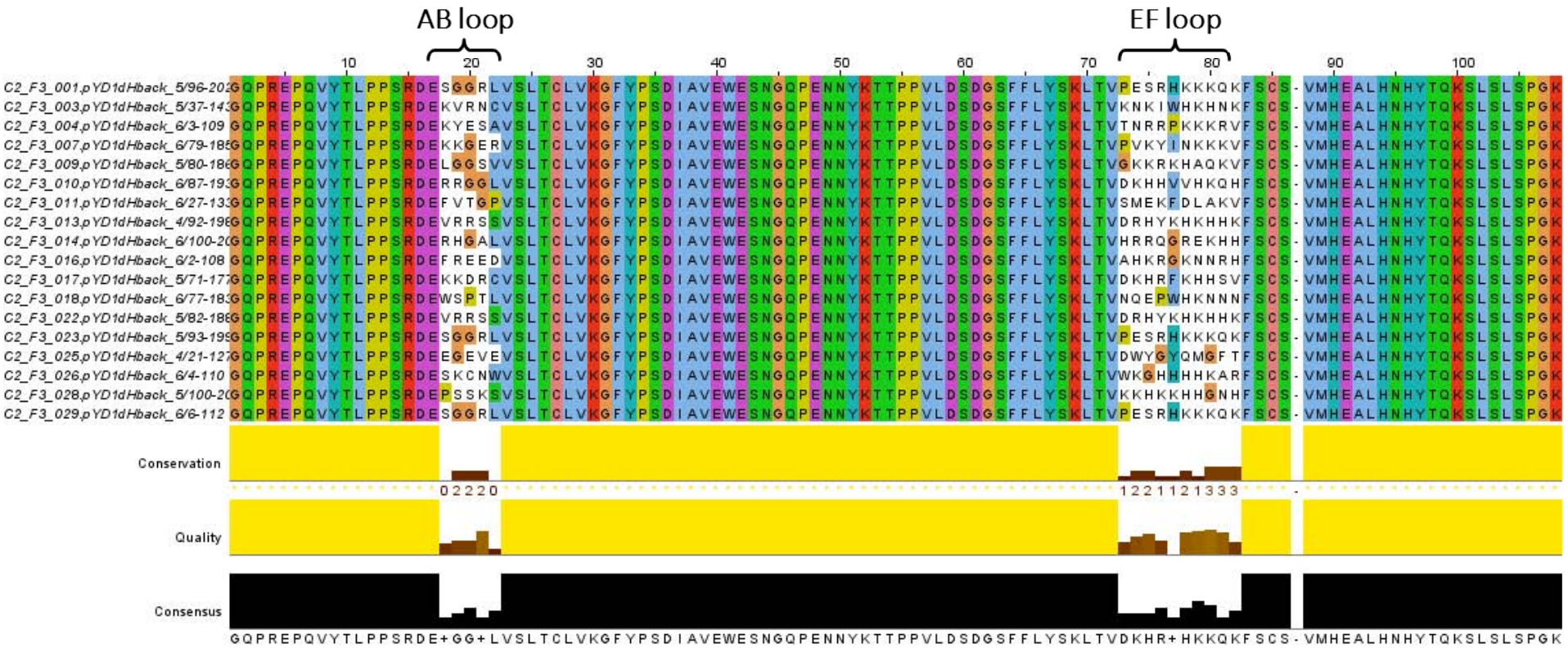


Figure 4-14 Sequence analysis of clones from cultures of the third round of FACS of campaign 2 for EF-CARA2. Translated nucleotide sequences that contain the motif TYRVVS and aligned using the software *Jalview*. Only sequences without frameshift or point mutations are shown, from the start of the C_H3 domain (GQPREP) to the C-terminus of IgG1-Fc (SLSLSPGK). Both AB loop and EF loop show enrichment in K = lysine and R = arginine.

4.3.1.3 UP+0_RW

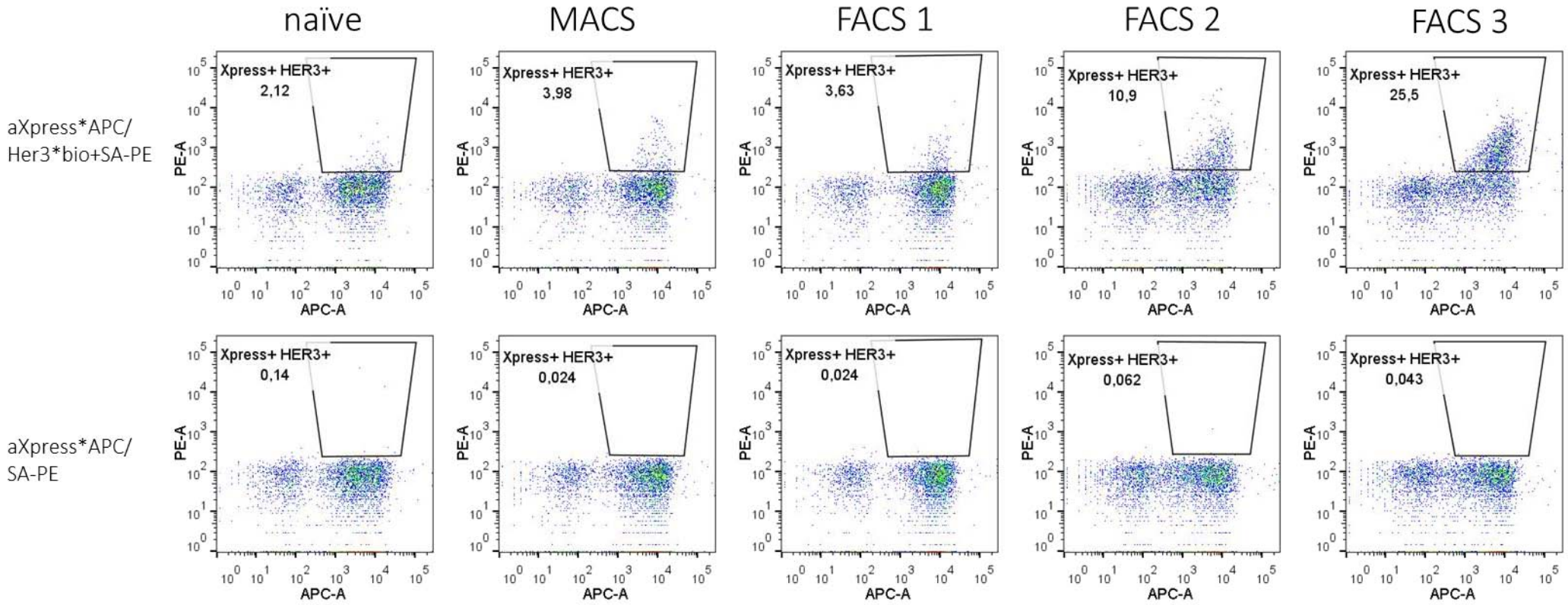


Figure 4-15 HER3 binding enrichment in the course of all sorting rounds of campaign 2 for UP+0_RW. Upper row: Samples stained with aXpress*APC and biotinylated HER3+SA-PE, lower row: double control stained with aXpress*APC and SA-PE for normalization of unspecific binding to the secondary stain (PE).

All plots show populations gated for correct morphology and presence as single cells.

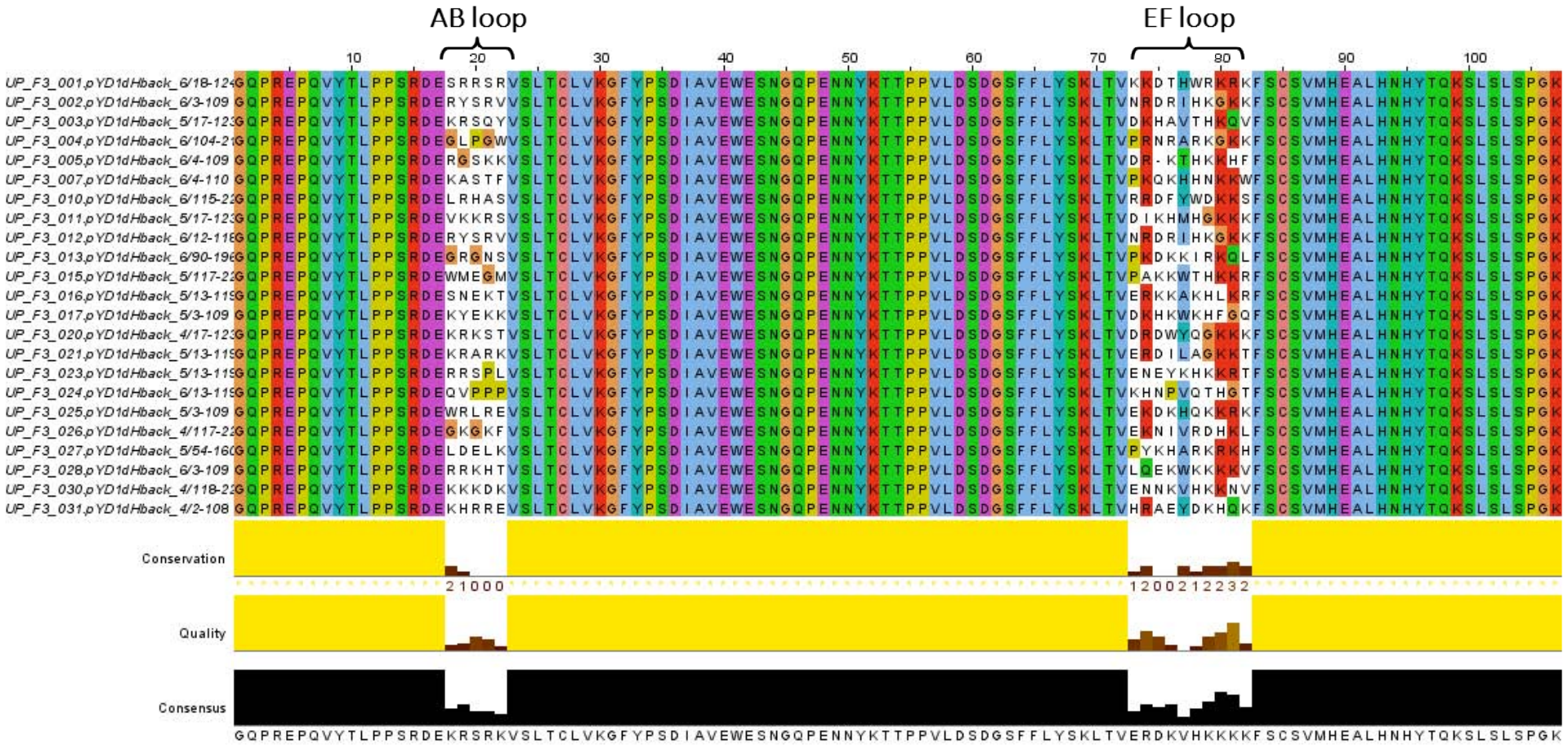


Figure 4-16 Sequence analysis of clones from cultures of the third round of FACS of campaign 2 for UP+0_RW Translated nucleotide sequences that contain the motif TYRVVS and aligned using the software *Jalview*. Only sequences without frameshift or point mutations are shown, from the start of the C_H3 domain (GQPREP) to the C-terminus of IgG1-Fc (SLSLSPGK). Both AB loop and EF loop show enrichment in K = lysine and R = arginine.

4.3.2 Proposed improvements of the naïve selection procedure

As selections repeatedly resulted in the enrichment of "sticky" binders, the information gathered in the course of the experiments was integrated in order to yield different measures to be taken in the future.

4.3.2.1 Gating strategy

The first idea concerns gating during the flow cytometric sorts. A sorting window can be dissected into 4 quadrants, where in the upper left quadrant (UL), no cells should be present unless they bind unspecifically to the fluorescent dye, in the lower left quadrant (LL) there are negative cells that are neither displaying nor binding. The lower right quadrant (LR) contains displaying cells without antigen binding which can be seen in the double control (Figure 4-17). For this control the same cell population was stained only with aXpress*APC and SA-PE/NA-PE in order to normalize unspecific binding to the secondary stain.

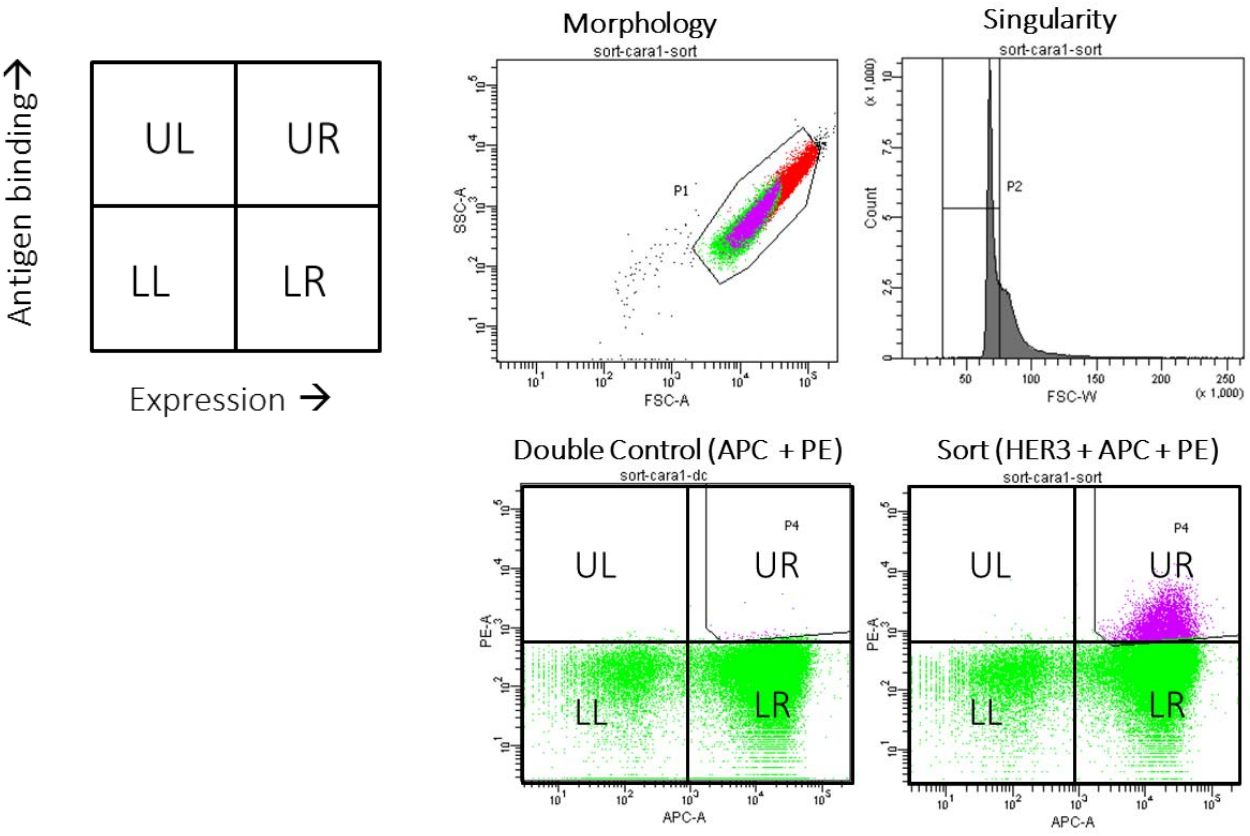


Figure 4-17 Dissection of a FACS sorting window into quadrants. Left: general quadrant designation and explanation of the meaning of X- and Y axis. Right: exemplary application to the second round of FACS in campaign 2 of EF-CARA1

Target cells, i.e. cells that display molecules with affinity to the antigen, are part of in the upper right quadrant (UR) (Figure 4-17).

Taking a closer look at this upper right quadrant (Figure 4-18) we hypothesize, that a "critical diagonal" can be drawn. Cells below this line are thought to bind to the antigen, but a large contribution to this binding can be accounted to the increased display level. On the other hand, cells above this critical line are thought to bind to the antigen with a high affinity. For desired cells, the binding is not proportional but over-proportional to the expression.

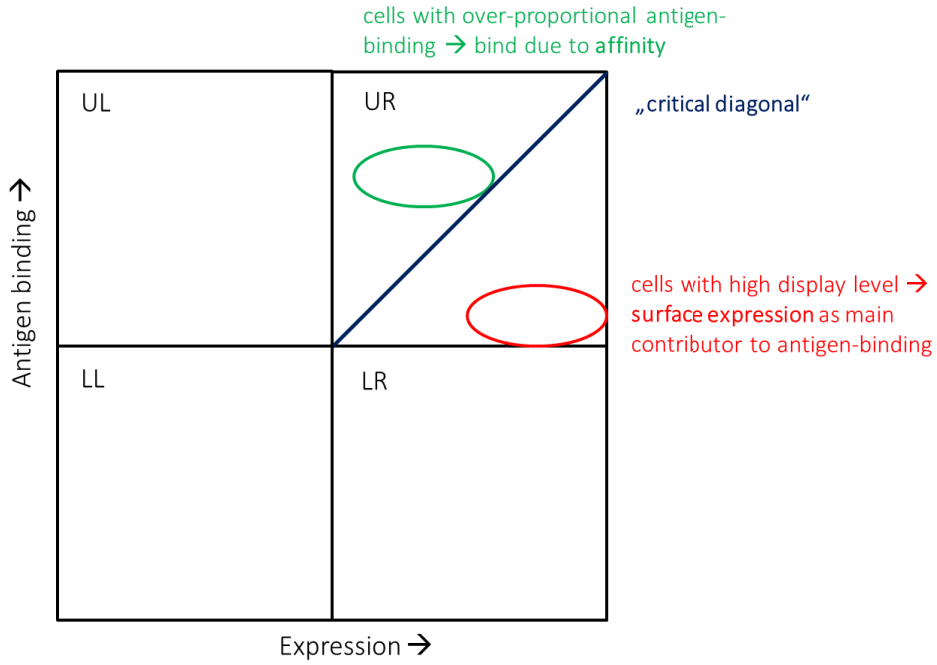


Figure 4-18 Interpretation of regions in the upper right quadrant of a FACS sorting window. The hypothesized "critical diagonal" is despicted in blue. Cells in the region of the red circle are thought to bind mainly due to the increased display level while the binding of cells in the region of the green circle is over-proportinal to their expression/display level and thus are thought to bind due to their high affinity to the antigen.

In Figure 4-19 this hypothesis is applied to the enrichment measurement of EF-CARA1, campaign 2. The overlay of the quadrant classification and particularly the critical line in the upper right quadrant of the target cells shows that in the course of all FACS rounds, the positive cell population stays below that critical line. This indicates, that the strong binding of the enriched clones was mainly caused by proportionally stronger expression levels, rather than over-proportionally high affinity to the antigen.

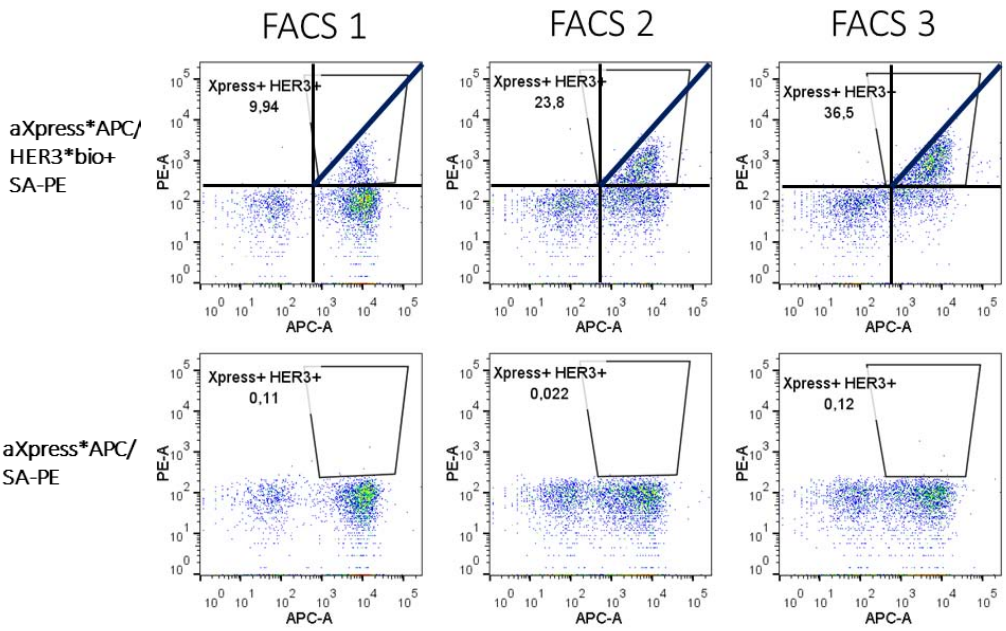


Figure 4-19 Application of the hypothesized sorting problem to the enrichment measurement of EF-CARA1, campaign 2. Upper row: Samples stained with aXpress*APC and biotinylated HER3+SA-PE, → for all consecutive FACS rounds, the cell population stays below the critical line (depicted in blue). Lower row: double control stained with aXpress*APC and SA-PE for normalization of unspecific binding to the secondary stain (PE).

In order to exclude these binders, one possible improvement could be to adjust the sorting gate for each step towards the critical line as depicted in Figure 4-20.

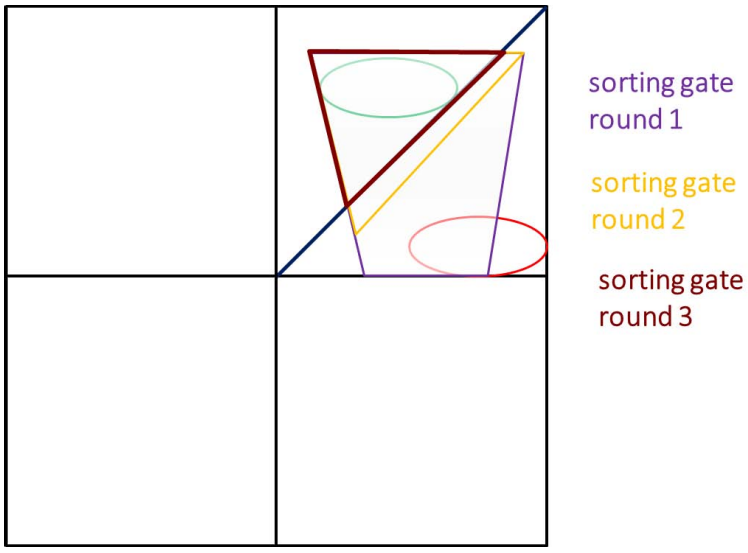


Figure 4-20 Proposed adaption of the sorting gate for elimination of undesired clones

4.3.2.2 Blocking reagent

There are different agents for blocking unspecific background staining for immunological procedures like ELISA, immunohistochemistry or flow cytometry. In the current thesis, bovine serum albumin (BSA) is used for saturating unoccupied sites on the yeast cell surface. BSA is a commonly and successfully used protein blocking reagent. However, some BSA preparations are very homogenous and thus have a disadisadvantage when it comes to blocking of some covalent surfaces (featuring hydrophobic or ionic characteristics)[40]. An alternative would be the use of nonfat dry milk[41], which has a higher diversity. This might be a determinant for the sensitive experimental set up in naïve selections.

4.3.2.3 Alteration of tags and dyes

Another proposed improvement of the FACS sortings is the strict alteration between both streptavidin and neutravidin for detection of the HER3 biotin tag and the fluorescent labels PE and APC. This is thought to prevent isolation of binders of the tags or the dyes which also facilitates the discovery of specific binding molecules.

4.3.2.4 Stringency

As the major challenge in naïve selections is the discrimination between weak binders and non-binders, another possibility is to tune the stringency of the sorts. For MACS selection, this could be achieved by varying the number washing steps and their duration[36]. For FACS selection steps, we think that lowering the antigen concentration in the later rounds, and thus an increased selection stringency, can be beneficial for successful affinity discrimination.

5 Conclusion and Outlook

This work offers first insights in how a distinct library design can influence the selection process of novel binding molecules to therapeutically interesting targets in directed evolution protein engineering. More precisely, the utilization of stability-biased amino acid distributions is promising for making the discovery process more efficient, as molecules with high affinity to the antigen can potentially be selected from a larger pool of functional and biophysically stable clones. This was shown by construction of yeast surface display variant libraries of the Fc fragment of immunoglobulin G1 and the assessment of the overall library qualities and thermal stabilities ($T_{1/2}$ values).

In the course of an number of naïve selection rounds (MACS/FACS), various problems were encountered. Consequently, a final proof-of-concept by comparison of solubly expressed single clones remains to be done. One highly interesting question that remains to be answered is, whether there is a trade-off between target-affinity and thermal stability. Also the potentially higher efficiency of the discovery of the EF-CARA1 library based on the present findings (see 4.2.1) would be interesting to prove. However, progress was made towards the efficient generation of large (10^9) yeast surface display libraries, and improvement of the *de novo* selection strategies.

6 References

1. Hooks, M. A., Wade, C. S. & Millikan, W. J. Muromonab CD-3: A Review of Its Pharmacology, Pharmacokinetics, and Clinical Use in Transplantation. *Pharmacotherapy: The Journal of Human Pharmacology and Drug Therapy* **11,** 26–37 (1991).
2. Nelson, A. L., Dhimolea, E. & Reichert, J. M. Development trends for human monoclonal antibody therapeutics. *Nature Reviews Drug Discovery* **9,** 767–774 (2010).
3. Monoclonal Antibodies Approved by the EMA and FDA for Therapeutic Use. *Animal Cell Technology Industrial Platform* (2013). at <http://www.actip.org/pages/monoclonal_antibodiestable.html>
4. Kubota, T. *et al.* Engineered therapeutic antibodies with improved effector functions. *Cancer Science* **100,** 1566–1572 (2009).
5. Scott, A. M., Wolchok, J. D. & Old, L. J. Antibody therapy of cancer. *Nature Reviews Cancer* **12,** 278–287 (2012).
6. Bang, Y.-J. *et al.* Trastuzumab in combination with chemotherapy versus chemotherapy alone for treatment of HER2-positive advanced gastric or gastro-oesophageal junction cancer (ToGA): a phase 3, open-label, randomised controlled trial. *The Lancet* **376,** 687–697 (2010).
7. Janeway Jr, C. A. *et al. Immunobiology*. (Garland Science, 2001).
8. Rubin-Pitel, S. B., Cho, C. M.-H., Chen, W. & Zhao, H. in *Bioprocessing for Value-Added Products from Renewable Resources* (ed. Yang, S.-T.) 49–72 (Elsevier, 2007).
9. Berg, J. M. *et al. Biochemistry*. (W H Freeman, 2002).
10. Cervenak, J. & Kacskovics, I. The neonatal Fc receptor plays a crucial role in the metabolism of IgG in livestock animals. *Veterinary Immunology and Immunopathology* **128,** 171–177 (2009).

11. Alberts, B. *et al. Molecular Biology of the Cell.* (Garland Science, 2002).
12. Halaby, D. M., Poupon, A. & Mornon, J.-P. The immunoglobulin fold family: sequence analysis and 3D structure comparisons. *Protein Engineering* **12,** 563–571 (1999).
13. De Marco, A. Biotechnological applications of recombinant single-domain antibody fragments. *Microbial Cell Factories* **10,** 44 (2011).
14. Harmsen, M. M. & De Haard, H. J. Properties, production, and applications of camelid single-domain antibody fragments. *Applied Microbiology and Biotechnology* **77,** 13–22 (2007).
15. Boder, E. T. & Wittrup, K. D. Yeast surface display for screening combinatorial polypeptide libraries. *Nature Biotechnology* **15,** 553–557 (1997).
16. Rüker, F. & Wozniak-Knopp, G. in *Recombinant Antibodies for Immunotherapy* **VI,** 231–24 (Cambridge University Press, 2009).
17. Ruiz, M. & Lefranc, M.-P. IMGT gene identification and Colliers de Perles of human immunoglobulins with known 3D structures. *Immunogenetics* **53,** 857–883 (2002).
18. Wozniak-Knopp, G. *et al.* Introducing antigen-binding sites in structural loops of immunoglobulin constant domains: Fc fragments with engineered HER2/neu-binding sites and antibody properties. *Protein Engineering Design and Selection* **23,** 289–297 (2010).
19. Traxlmayr, M. W. *et al.* Integrin binding human antibody constant domains—Probing the C-terminal structural loops for grafting the RGD motif. *Journal of Biotechnology* **155,** 193–202 (2011).
20. Wozniak-Knopp, G. & Rüker, F. A C-terminal interdomain disulfide bond significantly stabilizes the Fc fragment of IgG. *Archives of Biochemistry and Biophysics* **526,** 181–187 (2012).
21. Wozniak-Knopp, G., Stadlmann, J. & Rüker, F. Stabilisation of the Fc Fragment of Human IgG1 by Engineered Intradomain Disulfide Bonds. *PLoS ONE* **7,** e30083 (2012).

22. Lai, B., Hasenhindl, C., Obinger, C. & Oostenbrink, C. Molecular Dynamics Simulation of the Crystallizable Fragment of IgG1—Insights for the Design of Fcabs. *International Journal of Molecular Sciences* **15,** 438–455 (2014).
23. Hasenhindl, C. *et al.* Creating stable stem regions for loop elongation in Fcabs — Insights from combining yeast surface display, in silico loop reconstruction and molecular dynamics simulations. *Biochimica et Biophysica Acta (BBA) - Proteins and Proteomics* **1844,** 1530–1540 (2014).
24. Traxlmayr, M. W. *et al.* Construction of a Stability Landscape of the CH3 Domain of Human IgG1 by Combining Directed Evolution with High Throughput Sequencing. *Journal of Molecular Biology* **423,** 397–412 (2012).
25. Chao, G. *et al.* Isolating and engineering human antibodies using yeast surface display. *Nature Protocols* **1,** 755–768 (2006).
26. Kontermann, R. E. & Dübel, S. *Antibody Engineering*. (Springer Science & Business Media, 2010).
27. Lipovsek, D. *et al.* Evolution of an interloop disulfide bond in high-affinity antibody mimics based on fibronectin type III domain and selected by yeast surface display: molecular convergence with single-domain camelid and shark antibodies. *Journal of Molecular Biology* **368,** 1024–1041 (2007).
28. Gera, N., Hussain, M. & Rao, B. M. Protein selection using yeast surface display. *Methods* **60,** 15–26 (2013).
29. Prenzel, N., Fischer, O. M., Streit, S., Hart, S. & Ullrich, A. The epidermal growth factor receptor family as a central element for cellular signal transduction and diversification. *Endocrine-Related Cancer* **8,** 11–31 (2001).
30. Tzahar, E. *et al.* A hierarchical network of interreceptor interactions determines signal transduction by Neu differentiation factor/neuregulin and epidermal growth factor. *Molecular and Cellular Biology* **16,** 5276–5287 (1996).

31. Hsieh, A. C. & Moasser, M. M. Targeting HER proteins in cancer therapy and the role of the non-target HER3. *British Journal of Cancer* **97,** 453–457 (2007).
32. Sergina, N. V. *et al.* Escape from HER-family tyrosine kinase inhibitor therapy by the kinase-inactive HER3. *Nature* **445,** 437–441 (2007).
33. Chen, T. F., de Picciotto, S., Hackel, B. J. & Wittrup, K. D. in *Methods in Enzymology* (ed. Keating, A. E.) **523,** 303–326 (Academic Press, 2013).
34. Traxlmayr, M. W. & Obinger, C. Directed evolution of proteins for increased stability and expression using yeast display. *Archives of Biochemistry and Biophysics* **526,** 174–180 (2012).
35. Hasenhindl, C. *et al.* Stability assessment on a library scale: a rapid method for the evaluation of the commutability and insertion of residues in C-terminal loops of the CH3 domains of IgG1-Fc. *Protein Engineering Design and Selection* **26,** 675–682 (2013).
36. Ackerman, M. *et al.* Highly avid magnetic bead capture: an efficient selection method for de novo protein engineering utilizing yeast surface display. *Biotechnology Progress* **25,** 774–783 (2009).
37. Kawai, S., Hashimoto, W. & Murata, K. Transformation of Saccharomyces cerevisiae and other fungi. *Bioengineered Bugs* **1,** 395–403 (2010).
38. Benatuil, L., Perez, J. M., Belk, J. & Hsieh, C.-M. An improved yeast transformation method for the generation of very large human antibody libraries. *Protein Engineering, Design and Selection* **23,** 155–159 (2010).
39. Eckert-Boulet, N., Pedersen, M. L., Krogh, B. O. & Lisby, M. Optimization of ordered plasmid assembly by gap repair in Saccharomyces cerevisiae. *Yeast* **29,** 323–334 (2012).
40. Effective Blocking Procedures - ELISA Technical Bulletin - No. 3 - elisa3.pdf. at <http://csmedia2.corning.com/LifeSciences/media/pdf/elisa3.pdf>

41. Duhamel, R. C. & Johnson, D. A. Use of nonfat dry milk to block nonspecific nuclear and membrane staining by avidin conjugates. *The Journal of Histochemistry and Cytochemistry: Official Journal of the Histochemistry Society* **33,** 711–714 (1985).