Universität für Bodenkultur Wien
University of Natural Resources
and Life Sciences, Vienna

Master Thesis

Development of an assay for reducing non-specific binding of engineered proteins

Submitted by

Alex Alton, BSc.

in the framework of the Master programme

Biotechnology

in partial fulfilment of the requirements for the academic degree

Diplom-Ingenieur

Vienna, November 2021

Supervisor:

Univ. Prof. Dr. Christian Obinger
Institute of Biochemistry
Department of Chemistry

Co-supervisor:

Dipl.-Ing. Dr. Michael Traxlmayr
Institute of Biochemistry
Department of Chemistry

Eidesstattliche Erklärung

Ich erkläre eidesstattlich, dass ich die vorliegende Arbeit selbständig angefertigt habe. Es wurden keine anderen als die angegebenen Hilfsmittel benutzt. Die aus fremden Quellen direkt oder indirekt übernommenen Formulierungen und Gedanken sind als solche kenntlich gemacht. Diese schriftliche Arbeit wurde noch an keiner Stelle vorgelegt.

Wien, Oktober, 2021 _____

Signatur

Abstract

Over the recent years, technological progress and the emergence of display technologies like yeast surface display (YSD) facilitated the generation of high affinity binders. However, next to high affinity a number of so called 'drug like properties' like solubility, expression in mammalian cells and little or no non-specificity were identified as crucial aspects for the clinical success of a therapeutic protein. Non-specific binding characteristics can be described as the relative propensity of a binder to interact non-specifically with other than its target antigen which also includes self-interaction. Such off-target effects were shown to promote aggregation *in vitro* and an increased clearance rate *in vivo*. A lack of specificity caused many promising lead candidates to fail in the clinical stages of therapeutic antibody discovery and rose the necessity of novel tools that could help in the prediction of such undesired behaviour.

Throughout this project, we attempted to establish a highly avid bead capture based assay that could assist in the reduction of non-specifically interacting protein variants already during the developmental process of engineered proteins. We immobilized proteins on the surface of the magnetic beads which showed the characteristics described in the literature as the main factors that fuel non-specific interactions, namely hydrophobicity and positive charge. We hypothesized that those protein coated beads would successfully capture non-specific binders when incubated with any given protein library. Subsequently, those undesired protein variants could be depleted from the rest of the library by magnetic selection. The set of proteins coupled with the beads consisted of two Sso7d based binders, one binder based on the 10th type III domain of human fibronectin (FN3) and one single chain variable fragment (scFv). All those binders were expressed in fusion with a stabilizing protein, either the small ubiquitin-like modifier (SUMO) or the maltose binding protein (MBP). The functionality of the bead selection assay was tested with a set of YSD samples that displayed alternative scaffold proteins with known characteristics. For all the samples, the fraction of displayers present before and after a selection with the protein coated beads was evaluated by flow cytometric measurement. Ideally, the assay should have led to a reduction of displayers for all the samples containing non-specifically interacting proteins while the other samples should not have been altered by the assay. In summary, we were not able to observe any reduction of non-specifically binding proteins caused by the bead selection assay.

Kurzfassung

Der technologische Fortschritt der letzten Jahre und die Entwicklung der Display-Technologien, z.B. yeast surface display (YSD), haben es ermöglicht Binder hoher Affinität gegen eine Vielzahl von Antigenen zu generieren. Dennoch wurden neben hoher Affinität auch hohe Bindungsspezifität, sowie auch andere sogenannte ,,Drug-like Properties" als entscheidend für den Erfolg von therapeutischen Proteinen in klinischen Studien erkannt. Binder reduzierter Spezifität weisen neben ihren spezifischen auch unspezifische Bindungseigenschaften auf. Unspezifische Bindungseigenschaften werden als die relative Neigung eines Binders, etwas anderes als sein eigentliches Zielprotein zu binden bezeichnet, was auch die Interaktion mit sich selbst miteinschließt. Eine derartige Neigung führt zu Aggregatbildung *in-vitro* und zu einer erhöhten Clearance-Rate *in vivo*. In der Forschung von therapeutischen Proteinen ist ein vielversprechender Antikörper schon oft in den klinischen Studien gescheitert, weil er unspezifische Bindungseigenschaften aufwies. Deswegen müssen gezielt neue Methoden entwickelt werden, welche diese ,,off-target" Effekte schon früh in der Binderentwicklung reduzieren können.

In diesem Projekt haben wir versucht ein solches Tool zu entwickeln, mittels einer Methode, die auf ,,Highly avid bead capture" basiert. Wir immobilisierten Proteine auf der Oberfläche der magnetischen Beads, die genau solche Eigenschaften zeigten, welche in der Literatur als Hauptfaktoren für unspezifische Bindungseigenschaften beschrieben werden, nämlich Hydrophobizität und positive Ladungen. Wir stellten die Hypothese auf, dass solche mit Proteinen besetzte Beads erfolgreich an nicht spezifische Binder in einer Protein Library anhaften sollten. Anschließend könnte man diese dann gemeinsam mit den Beads durch einen Magneten abtrennen. An die Beads wurden die folgenden vier Proteine gekoppelt: zwei Sso7d-basierte Proteine, ein Binder, der auf der dritten humanen Fibronektin Domäne (Fn3) basierte und schlussendlich auch ein single chain variable fragment (scFv). All diese Proteine wurden gemeinsam mit einem stabilisierenden Fusionsprotein exprimiert, nämlich dem small ubiquitin-like modifier (SUMO) Protein oder dem Maltose-bindenden Protein von *E. coli* (MBP). Die Wirksamkeit des Bead-Selektions Assays wurde mit einer Reihe von YSD Proben getestet, welche alternative Scaffold Proteine mit bekannten Eigenschaften beinhalteten. Von jeder Probe wurde die Anzahl an Displayern im Sample vor und nach der Bead-Selektion mittels Flow Zytometrie bestimmt. Wir erhofften uns, dass der Assay bei allen Proben die unspezifisch bindende Proteine enthielten zu einer Reduktion der Displayer Anzahl führt, während alle anderen Proben durch die Bead-Selektion nicht hätten beeinträchtigt werden sollen.

Bei der Auswertung unserer Daten konnten wir keine Effekte beobachten, die auf den Bead-Selektions Assay zurückgeführt werden konnten. Entgegen unserer Hypothese deuten diese Ergebnisse darauf hin, dass die von uns verwendete Methode oder die von uns auf den Beads getesteten Proteine nicht für die Reduktion von unspezifisch bindenden Proteinen in einer Protein Library geeignet ist.

Acknowledgements

Through the completion this project, I have gained a lot of knew knowledge and experience in the field of protein engineering and molecular biology. While I am very thankful for that I would like to point out even more the great deal of support that I have received.

First of all, I am very thankful for the exceptional supervision by Charlotte Zajc. Thank you Charlotte for always being very supportive and providing all the assistance in the laboratory, for continuously being available for all my questions and sharing your expertise.

Next I would like to thank my supervisor Michael Traxlmayr for giving me the opportunity to do research on such an interesting topic. Thank you for all your guidance throughout the project and all your valuable advices.

Of course I would also like to thank Christian Obinger for the possibility to do my master thesis project in his laboratory and being part of his research group.

Furthermore, I appreciated it a lot to have such great colleagues with the BioB research team by my side. Thank you for all your support and for creating a very enjoyable working environment.

Also, it was a great experience to be part of the CD laboratory for Next Generation CAR T cells. Thank you to Manfred Lehner and the whole research group from the CCRI for all the discussions and insights you shared with me.

And finally, I would like to thank my girlfriend, friends, my sister and especially my parents for their consistent and unconditional support over the course of my studies.

Acronyms

BSA bovine serum albumin. 14, 16, 33

BVP baculovirus particle. 43

CDR complementarity determining region. 6, 27, 28, 42

CIC cross-interaction chromatography. 43

dNTP's desoxyribonucleotide triphosphates. 16, 19

EGFR epidermal growth factor receptor. 3, 9, 26, 41

FACS fluorescence-activated cell sorting. 2, 3, 7, 8, 43, 44

FN3 10th type III domain of human fibronectin. ii, 3, 9, 27, 28, 40, 42

HA hemagglutinin. 1, 2

HPLC high-performance liquid chromatography. xi, 11, 16, 23, 33–35, 41

IMAC immobilized metal affinity chromatography. 22

LB lysogeny broth. 13

MALS multi angle light scattering. xi, 30–34, 41

MBP maltose binding protein. ii, 9, 22, 26, 28, 30, 31, 33, 40–42

MES 1,3,5-trimethylbenzene. 14

MW molecular weight. xi, 22, 31–34

PBS phosphate buffered saline. 14, 22, 23

PCR polymerase chain reaction. xi, 2, 19

PSR polyspecificity reagent. 5, 7, 43, 44

rcSso7d reduced charge Sso7d. 27

scFv single chain variable fragment. ii, iii, 3, 7, 9, 21, 23, 26–31, 40–42

SEC size exclusion chromatography. x, 27, 30–35, 41

SUMO small ubiquitin-like modifier. ii, iii, 9, 15, 17, 18, 21, 22, 26–30, 40–42

TAE tris-acetate-EDTA. 16, 20

TGS tris/glycine/SDS. 16

YSD yeast surface display. ii, iii, x, 1–3, 7–10, 26, 27, 36, 37, 42–44

Contents

Abstract ii

Kurzfassung iii

Acknowledgements v

Acronyms vi

1 Introduction 1
- 1.1 Yeast surface display . . . 1
  - 1.1.1 Advantages compared to other display technologies . . . 2
  - 1.1.2 Applications and recent impact of YSD . . . 3
- 1.2 Therapeutic binder's requirements beyond affinity: The concept of 'developability' . . . 3
  - 1.2.1 The importance of specificity . . . 4
- 1.3 Major factors that drive non-specific binding . . . 5
  - 1.3.1 Positive charge . . . 6
  - 1.3.2 Hydrophobicity . . . 6
- 1.4 Magnetic bead based selection as an efficient alternative to FACS for capturing weak protein-protein interactions . . . 7

2 Aim of the project 9

3 Materials and Methods 11
- 3.1 Materials . . . 11
  - 3.1.1 Laboratory equipment . . . 11
  - 3.1.2 Disposables . . . 12
  - 3.1.3 Kit systems . . . 12
  - 3.1.4 Buffers, solutions and media . . . 12
  - 3.1.5 Plasmids . . . 15
  - 3.1.6 Enzymes and antibodies . . . 15
  - 3.1.7 Yeast and bacterial strains . . . 15
  - 3.1.8 Software . . . 16
  - 3.1.9 Miscellaneous material . . . 16

3.1.10 Primer sequences . . . 17
3.1.10.1 Primers . . . 17
3.1.11 Amino acid sequences of expressed proteins . . . 17
3.2 Methods . . . 19
3.2.1 Molecular biology methods . . . 19
3.2.1.1 Polymerase chain reaction (PCR) . . . 19
3.2.1.2 Restriction digest . . . 19
3.2.1.3 HiFi DNA assembly . . . 20
3.2.1.4 Agarose gel electrophoresis . . . 20
3.2.1.5 Heat shock transformation . . . 20
3.2.1.6 Frozen yeast transformation . . . 21
3.2.1.7 SDS PAGE . . . 21
3.2.1.8 Cryo stocks . . . 21
3.2.2 Expression and purification . . . 21
3.2.2.1 Expression . . . 21
3.2.2.2 Talon gravity flow purification . . . 22
3.2.2.3 Purification with the Bio-Rad system . . . 22
3.2.3 Biotinylation of solubly expressed proteins . . . 23
3.2.4 Size exclusion chromatography (SEC) - multi angle light scattering (MALS) . . . 23
3.2.5 Yeast surface display methods . . . 23
3.2.5.1 Cultivation and induction of yeast cells . . . 23
3.2.5.2 Yeast cell harvest and staining . . . 24
3.2.5.3 Bead selection assay . . . 24
3.2.5.4 Evaluation by flow cytometry . . . 25

4 Results **26**
4.1 Selection of the proteins associated to the streptavidin beads . . . 26
4.1.1 Design of solubly expressed fusion proteins . . . 28
4.2 'Sticky' and 'non-sticky' binders for the evaluation of our bead selection assay performance . . . 29
4.3 Biophysical characterization of solubly expressed proteins . . . 29
4.3.1 SDS-PAGE . . . 30
4.3.2 SEC-MALS measurement . . . 31
4.4 Evaluation of biotinylation . . . 34
4.5 Definition of a suitable displayer to non-displayer ratio . . . 35
4.6 Bead selection assay . . . 36

5 Discussion **40**

List of Figures

1.1 A-agglutinin anchor system used for YSD. 2
1.2 Flag based system performance analysed by the company Adimab 4
1.3 Approved antibodies and clinical stage antibodies statistically differentiate in assays that evaluate their specificity 5
1.4 Mutlivalency of magnetic bead capture 8

4.1 Interacting beta strands and loops of expressed proteins 27
4.2 Design of solubly expressed proteins 28
4.3 SDS-PAGE of SUMO-R1 and MBP-S. 31
4.4 SEC data of solubly expressed proteins. 32
4.5 Peak tailing of E5. 33
4.6 SEC analysis of standard proteins and log(MW) vs retention time. 33
4.7 Verification of successful biotinylation 35
4.8 Titration experiment for the evaluation of a good displayers to non-displayers ratio. 36
4.9 Data from bead selection experiments. 39

List of Tables

3.1 Laboratory equipment 11
3.2 Disposables 12
3.3 Kit systems 12
3.4 Complete media and solutions for bacterial and yeast cell cultivation . . . 13
3.5 Buffer compositions 14
3.6 Plasmids 15
3.7 Enzymes 15
3.8 Antibodies for flow cytometry 15
3.9 Yeast and bacterial strains 15
3.10 Software programs 16
3.11 Miscellaneous material 16
3.12 Primers used for DNA amplification. All of them were ordered at the company Microsynth. The capitalized sequence part matches with the respective template, while the lower case part represents the overhang. . . 17
3.13 Amino acid sequences of expressed proteins 18
3.14 General setup of PCR reactions 19
3.15 PCR cycle conditions 19
3.16 General reaction setup for Hifi DNA assembly 20
3.17 Extinction coefficients and molecular weights of purified proteins 22

4.1 Set of binders with known properties displayed on yeast cells. 29
4.2 Amount of protein purified from soluble expressions. 30
4.3 Theoretical and MALS-determined molecular weights of standard proteins. 34
4.4 Theoretical MWs, MALS-determined MWs and MWs calculated from HPLC standard 34

Chapter 1

Introduction

1.1 Yeast surface display

Contributing to numerous success stories, yeast surface display YSD has become one of the most important methods in the space of protein engineering. Originally, the presentation of proteins on the surface of *Saccharomyces cerevisiae* was utilized for the immobilization of pathogens in vaccine development. There, it was demonstrated that the C-terminal half of the baker's yeast mating protein α-agglutinin contains all the information needed to anchor a desired protein of interest into the cell wall of the organism [1]. Next to this N-terminal fusion to α-agglutinin, plenty of other anchor proteins were discovered over the years. Many of them belong to the glycosylphosphatidylinositol (GPI) anchors like the Flo1p, Cwp1p and Cwp2p. They are embedded in the cell wall through a C-terminally linked glycophospholipid [2].

A-agglutinin that is related to α-agglutinin and is part of the GPI protein family offers the most popular way to link a protein of interest to the surface of yeast cells. Unlike α-agglutinin, a-agglutinin consists of two subunits, the Aga1p and the Aga2p. While the Aga1p functions as a covalent β-glucan linkage to the cell wall, the Aga2p is accessible on the surface and connected to the Aga1p via two disulfide bonds. The a-agglutinin anchor system was first described by Wittrup et. al. in 1997 [3]. There, the Aga2p is expressed in fusion with the protein of interest and is naturally linked to the Aga1p after secretion. C-terminally to the Aga2p the fusion protein usually comprises a nine amino acid hemagglutinin (HA) tag that can be used for detection, followed by any given protein of interest and an additional 10 residue c-myc tag [3, 4]. The Figure 1.1 represents a schematic illustration of the a-agglutinin system.

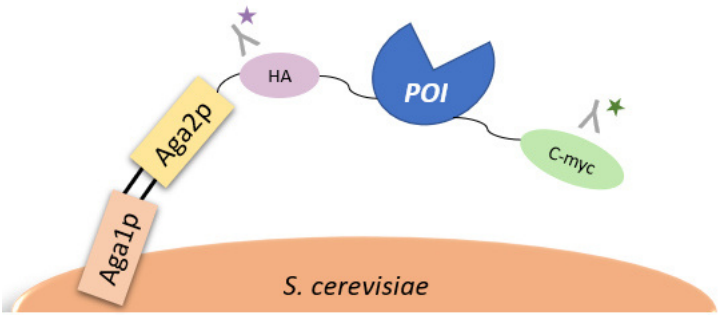


Figure 1.1: The a-agglutinin anchor system which is most often applied for yeast display. The Aga1p is embedded in the cell wall and linked to the Aga2p via two disulfide bonds. The Aga2p comes in fusion with the protein of interest (POI) sandwiched between two tags that are useful for detection, the HA and c-myc tag [3].

The most important feature of yeast display and other *in vitro* display methods like phage display [5] or ribosome display [6, 7] is that the phenotype of a desired protein is linked to its genotype present inside the cell or entity the protein is connected to. Thereby, when favorable protein variants are identified in a diversified library those can simply be reproduced and identified through their genotypes. Typically, methods such as error-prone PCR are applied for the generation of diversified libraries of a protein. Following randomization, the library is incubated with a desired antigen which is subsequently stained using fluorophore conjugated antibodies. Finally, the yeast cells are screened for antigen binding and sorted sequentially through fluorescence-activated cell sorting FACS with high throughput. In addition, staining of the mentioned detection tags can be performed which yields information about the fusion protein expression and full length display. As the phenotype-genotype linkage enables simple reproduction of desired variants, such randomization and selection steps can be performed for multiple rounds and result in an enrichment of antigen binding proteins [8, 9, 10].

1.1.1 Advantages compared to other display technologies

With a library size of up to 10^9 clones, YSD libraries are typically smaller compared to other methods like phage display, where up to 10^{14} variants can be screened [10]. Moreover, initial YSD screens with library diversities of over 10^8 should be performed using magnetic bead based selection systems which will be discussed in more detail below (see section 1.4) [11, 12]. Despite that, the yeast display platform stands out among other display technologies with some remarkable advantages. Most notably, as a eukaryotic expression system yeast cells are able to perform post-translational modifications carried out through the endoplasmic reticulum, e.g. the formation of disulfide bonds. Moreover, variations in glycosylation patterns between yeasts and mammalian cells have not prevented successful display of glycosylated human proteins. Using the detection tags one can quantify and relate the surface expression to a desired characteristic like antigen binding. Consequently, clones that show high expression in

addition to high antigen affinity for instance can be identified. Moreover, the evaluation through flow cytometry is much more reliable in distinguishing slight changes in affinities in comparison to the 'panning' method applied for phage display. Finally, important kinetics' values like the dissociation rate constants (k_{off}) or the affinity of binding can be determined by FACS measurement [3, 10, 13].

1.1.2 Applications and recent impact of YSD

In early applications, YSD was used for the affinity maturation of antibodies and single chain variable fragments (scFvs) [14]. More recently and with the emergence of alternative binding scaffolds also other human and non-human derived proteins were engineered to bind a range of different targets with high affinities. The 10th type III domain of human fibronectin (FN3) was engineered to bind lysozyme with a K_D in the picomolar range [15, 16]. Diversified libraries of the Sso7d scaffold, a protein that originates from the hyperthermophilic archaeon *Sulfolobus solfataricus* yielded binders in the nanomolar range that bind various targets including epidermal growth factor receptor (EGFR), K-Ras, streptavidin, chicken and mouse immunoglobulin and lysozyme [17, 18, 19]. When the affinity of a protein towards a specific target should be increased, the library based approach often poses a more valuable alternative compared to rational design as structure function relationships of proteins are elucidated only to a certain extend. Over the years yeast display was also used for other purposes next to the modulation of binding affinities. Some popular examples are epitope mapping, stability and specificity engineering or the generation of *de novo* binders from combinatorial libraries. Overall, yeast surface display has become an indispensable and versatile tool in the protein engineering space and is extensively applied for therapeutics development [10, 17].

1.2 Therapeutic binder's requirements beyond affinity: The concept of 'developability'

Over the last decade, protein based therapeutic binders have become a major player in drug industry. In particular, the emergence of antibodies as drugs had a huge impact on the therapeutics' realm with 8 of the top 10 bestselling drugs being biologics in 2018 [20]. Improvements of *in vivo* and *in vitro* discovery technologies like YSD fueled this development and turned the affinity maturation of binders towards a wide range of targets into a more routine task. Yet, it is still quite challenging to generate protein variants that show on top of high affinity also desirable biophysical characteristics which are crucial for clinical success. Such attributes are often termed 'drug-like' properties or as 'developability' of a binder and include high expression in mammalian cells, stability during storage, low aggregation propensity, high solubility and little or no non-specificity [21, 22].

In 2017, the company Adimab carried out a large study where they defined a dozen of assays with respective cutoff values that would eventually evaluate biophysical properties

necessary for drug-like behaviour [23]. Their goal was to provide drug developers a tool that could assist in the decision about which early stage lead candidate is most likely to be successful in the subsequent developmental steps and whether it should be subjected to further protein engineering or formulation before advancement of the program. Their dataset consisted of 137 human IgG1 isotype antibodies, where 48 were clinically approved and the remaining ones at least in phase 2 of their clinical trials. For every biophysical assay, the 10% worst performing candidates were assigned a 'red flag' [23]. To their surprise, approved drugs received consistently less flags compared to antibodies in earlier stages (Figure 1.2A). Consequently, a lower number of flags seemed to correlate with a higher probability of passing clinical trials. Moreover, antibodies engineered by phage display showed a significantly higher amount of flags compared to antibodies that derived from mammalian sources indicating worse biophysical characteristics (Figure 1.2B) [23].

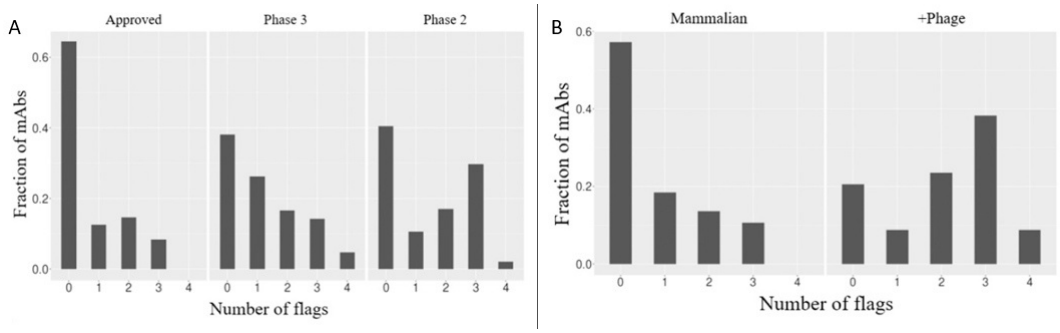


Figure 1.2: Flag based system performance analysed by the company Adimab. For every biophysical assay the 10% worst performing antibodies got a red flag. A - Comparison of flag number assigned to approved antibodies, phase 3 and phase 2 antibodies. B - Number of flags assigned to antibodies that were engineered by phage display opposed to mammalian source derived ones (modified from [23]).

1.2.1 The importance of specificity

Among the 'drug-like' biophysical properties analysed by Adimab specificity can be argued as being one of the most undervalued ones. Yet it is assumed to be a key factor for therapeutic success. Non-specificity can be described as the relative propensity of a binder to interact non-specifically with molecules other than their target antigen. It includes both, non-specific off-target binding and self-interaction which lead to poor performance *in vivo* because of rapid in vivo clearance and *in vitro* through aggregation, respectively. Such interactions are rather weak and re-dissociation occurs promptly [24, 23, 25]. Here, non-specific binders are also described as 'sticky', while others that interact only specifically with their actual target antigen are accordingly termed 'non-sticky' binders.

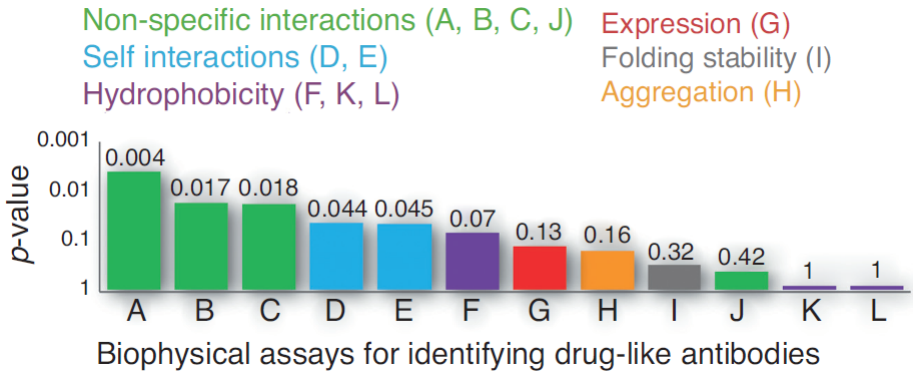


Figure 1.3: Jain et al. performed 12 biophysical property assays with 137 therapeutic antibodies [23]. Only assays that measure specificity could differentiate statistically between approved and clinical stage antibodies. Performed assays: (A) baculovirus particle (BVP), (B) ELISA (panel of six biomolecules), (C) polyspecificity reagent (PSR), (D) clone self-interaction by biolayer interferometry (CSI-BLI), (E) affinity-capture self-interaction nanoparticle spectroscopy (AC-SINS), (F) salt gradient affinity-capture self-interaction nanoparticle spectroscopy (SGAC-SINS), (G) expression titer in HEK293 cells, (H) aggregation (accelerated stability), (I) folding stability (melting temperature of Fab), (J) cross-interaction chromatography (CIC), (K) hydrophobic interaction chromatography (HIC), and (L) standup monolayer adsorption chromatography (SMAC) (modified from [24]).

Specificity remains a relative concept and highly depends on the assays used to evaluate it. The 12 biophysical property measurements mentioned before include two self-interaction and four non-specific cross interaction assays for the evaluation of antibody specificity [23]. Intriguingly, those were the only measurements that could statistically differentiate approved antibodies from those that are still in clinical studies (Figure 1.3). This finding leads to the assumption that approved antibodies were superior in their binding specificity compared to clinical stage candidates and underlines the importance of this biophysical characteristic for the successful advancement in clinical development [24].

1.3 Major factors that drive non-specific binding

Non-specific binding is believed to be fueled by both, electrostatic and hydrophobic interactions. The dynamics behind this type of interaction are quite complex and particularly hard to study because residues that are responsible for off-target binding are often critical for antigen recognition at the same time. However, recent efforts through experimental and computational approaches identified positive charges and hydrophobic patches as the potential main causes of reduction in specificity [26, 27, 28, 29, 30]. Such findings help to clarify the concept of non-specific binding and pave the way for the development of novel tools that could assist in avoiding sticky binders in the future [24].

1.3.1 Positive charge

Positively charged amino acids have been associated with non-specific binding in several studies of antibody complementarity determining region (CDR) loops' characteristics. In fact, elevated amounts of positive charges were found in antibodies produced by immature B-cells which potentially results in self-reactivity. While efficient checkpoint regulation leads to their removal from the antibody population during B cell development, *in vitro* technologies lack such control mechanisms [31]. Generally, an increase in the net positive charge of proteins was found to increase tissue retention and increased blood clearance *in vivo*. Possibly, electrostatic interaction occurs with the inherently anionic cell membranes whereby blood concentration and tissue disposition kinetics can be altered [32]. Especially, the positively charged amino acid arginine was repeatedly reported to be a main contributor to non-specific interactions. Both, as individual residue and within protein motifs. Thereby, its distributed positive charge as well as the pooly hydrated guanidinium group and its long hydrophobic side chain could be the crucial factors. Positive charge was also identified as possible driver of non-specificity in a study with limited diversity libraries. There, antibodies were enriched for non-specificity in their H3 loop and showed an increase in the occurrence of arginine and to a lesser extent lysine [26]. However, arginine was also found to be an important amino acid for proper antigen recognition [33, 27, 24]. Regarding positive charge and clearance rate *in vivo*, the isoelectric point of a protein can be a good measure for evaluation. Indeed, Igawa et al. demonstrated that within a small panel of antibodies differing only in their variable region sequence those with a lower isoelectric point have a longer half-life [27].

Unlike positive charge, negatively charged amino acids were mostly associated with specific antigen recognition in the past. Nevertheless, a recent homology modeling and molecular dynamics simulation study performed by Nichols et al. experimentally evaluated the impact of designed mutations on antibodies binding affinity next to other physicochemical properties. Notably, they found that a negatively charged patch increased the viscosity of the antibody [34].

1.3.2 Hydrophobicity

The second main characteristic that can eventually lead to non-specific interaction is hydrophobicity. Hydrophobic patches are usually buried inside the proteins' core where they contribute to the free energy of folding. In contrast, solvent exposed hydrophobic residues lead to a negative contribution. Interaction with another protein could compensate this negative free energy impact and therefore be fostered by those solvent accessible amino acids. Such hydrophobic clusters on the proteins' surface can be formed by consecutive amino acids in sequence, but also form through the three dimensional structure of a binder [26, 35, 24]. To investigate the influence of hydrophobicity to non-specific binding, an aggregation prone antibody was compromised in its aberrant behaviour through the replacement of amino acids in a hydrophobic hot spot. However, exchanging those residues with alanine simultaneously reduced its binding affinity which

underlines the importance of hydrophobic amino acids for effective antigen recognition [30].

Recently, a study with synthetic yeast display scFv libraries revealed some important insights regarding common features and motifs that can drive non-specific binding. A modified version of the PSR assay that will be discussed in more detail below (see chapter 5) was used to intentionally enrich non-specific binders in the constructed library. Next to the positively charged arginine and glycine which is likely contributing to non-specificity by providing flexibility the hydrophobic amino acids tryptophan and valine were predominantly present in those non-specifically acting libraries [26]. While initial screens already indicated that tryptophan could be a major driver of non-specificity the impact of valine was more surprising. The branched hydrophobic amino acid was found to be significantly enriched in the engineered library, whereby two or three consecutive Val as a motif seemed to have a considerable impact on non-specific interactions. Intriguingly, similar effects were not observed for the other aliphatic amino acids leucine and isoleucine [26]. The aromatic amino acid tryptophan is commonly found in antibody paratopes and as part of hydrophobic patches [36]. Interestingly, and in contrast to the other enriched amino acids it was only found to be predominantly present as a part of motifs but not as individual residue. Unlike tryptophan, its aromatic relative tyrosine has been primarily associated with specific recognitions in the past. Nevertheless, two consecutive tyrosines for instance were also identified as a non-specific interaction promoting motif by the investigators [26].

1.4 Magnetic bead based selection as an efficient alternative to FACS for capturing weak protein-protein interactions

As mentioned before, FACS based selection of desirable variants in a library is the state of the art method for protein engineering applications of YSD. It allows for normalization of surface display level and antigen binding resulting in selections without expression level bias, affinity differences can be distinguished and through the ability to quantify detected samples the kinetics of binding can be evaluated. Overall, FACS can be described as a high-throughput method with around 10^8 cells per hour and is well suited for affinity maturations [11, 3]. However, when it comes to the identification of de novo binding interactions, early sorting rounds and the detection of weak protein-protein interactions it shows some problematic drawbacks. First, the dependency on a fluorescent reagent for antigen detection can lead to selection of artifactual binders meaning interactions between the reagent itself and the yeast displayed proteins. Furthermore, weak binders on the yeast's surface require high concentrations to be administered, resulting in a considerable antigen demand. But most importantly, fast dissociation rates of weak interactions with a half-life in the range of seconds for micromolar affinities represent a major obstacle for the application of FACS given that another secondary incubation and several washing

steps need to be executed between the antigen incubation and FACS [11].

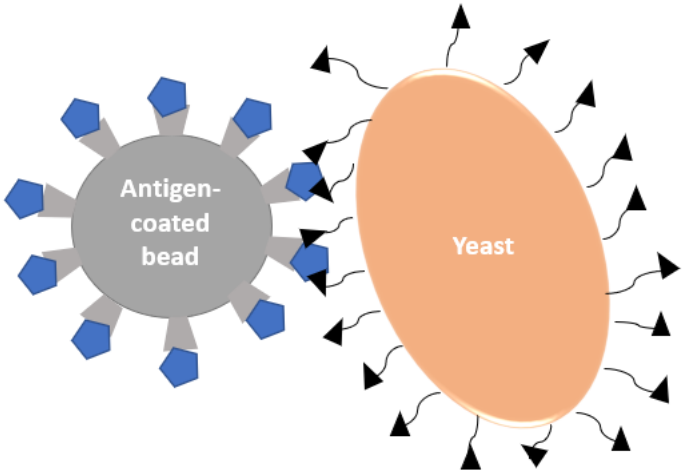


Figure 1.4: Multivalency of magnetic bead capture. Multiple interactions between displayed binders on the yeast cell (black triangles) and the antigen on the magnetic beads (blue pentagons) leads to considerable avidity effects in magnetic bead based selections.

In 2009, a novel method based on highly avid magnetic bead capture was published that enabled an elegant circumvention of the mentioned FACS' drawbacks regarding weak binders' selections [11]. Intriguingly, it still brings all the advantages which come with the protein presentation system of YSD. By antigen immobilization on magnetic (streptavidin) beads multivalent target presentation was coupled intentionally with the multivalency of yeast display. Thereby, avidity effects assure the selection of low-affinity binders (see figure Figure 1.4) [11]. The tremendous influence of polyvalency on the binding properties of antibodies was theoretically assessed by Crothers and Metzger [37]. While such estimations of bivalent interactions already demonstrate the tremendous effects of avidity in bivalent interactions, in multivalent setups such might become even more evident. As typical display levels of the protein to engineer are around 50,000 copies per yeast cell, this may lead to dissociation half-lifes of days instead of seconds [11]. Moreover, the method is inexpensive and billions of cells can be screened in roughly an hour. To sum up, highly avid bead selection represents a well-established alternative to FACS regarding the selection of weakly interacting binders and was furthermore applied for negative selection setups. A more detailed description of the approach can be found elsewhere [11].

Chapter 2

Aim of the project

The engineering of therapeutic binders (most notably monoclonal antibodies) using display technologies like YSD contributed to numerous success stories in the recent past. Over the years, advances in technology turned the affinity maturation of human and non-human based protein scaffolds to bind almost any given target into a more routine task. Despite this exciting development, the generation of therapeutic binders that show so called 'drug-like' properties next to high affinity binding remains an ongoing challenge. In 2017, the company Adimab has defined specificity or 'stickyness' among other 'drug-like' properties as a crucial factor for a therapeutic antibody's success in clinical trials [23]. A lack of specificity, i.e. a high susceptibility to interact non-specifically with other than the target antigen can fuel aggregation and cause rapid *in vivo* clearance [24]. The assessment and fixture of non-specific binding behaviour is typically addressed only after the binder discovery, thus limited to a few lead candidates and can be very cost intensive.

As specificity is suggested to be a key property to optimize already during early binder development, novel tools are required to fulfill this task. Within this project we aimed to establish a bead selection based assay that could assist in the reduction of such non-specific interactions in protein libraries. We hypothesised that proteins showing the characteristics suggested in the literature to be the main contributors to non-specific binding, namely hydrophobicity and positive charge, immobilized on the surface of magnetic beads would be able to provoke such interactions and thereby capture and successfully deplete non-specifically interacting variants from a differentiated library.

Consequently, we first aimed to express and characterize a set of proteins that exhibited non-specific binding properties. We planned to express two Sso7d based binders which were engineered to bind EGFR [18] & K-Ras [19], respectively, as well as one lysozyme-specific FN3 based binder [16] as SUMO fusion proteins. Additionally, we wanted to express an scFv based on the antibody sirukumab in fusion with the maltose binding protein (MBP) of *E. coli* [23, 38].

Next, we wanted to biotinylate and respectively immobilize those proteins on the surface of streptavidin-coated Dynabeads. According to our hypothesis, non-specifically interacting proteins in a library would associate to the sticky protein coated dynabeads and could subsequently be depleted via a negative selection, while the remainder of the

library would not be altered. Thereby, this bead selection based assay would represent a cost effective novel tool to reduce the non-specific binding content of a protein library that would be easy to be implemented in the developmental process of therapeutic binders.

To assess our hypothesis, we wanted to test every one of the expressed fusion proteins on the beads. To evaluate the performance of the assay, each protein on the beads should be evaluated with a selection of YSD samples that contained proteins with known characteristics. When incubated with sticky protein-displaying yeast cells the bead selection assay should cause a reduction of the fraction of displaying cells, while non-sticky protein samples should not be affected by the beads. The fraction of displayer before and after the selections should be evaluated using flow cytometry.

Chapter 3

Materials and Methods

3.1 Materials

3.1.1 Laboratory equipment

Table 3.1: Laboratory equipment

Equipment	**Supplier**
Arktik Thermal Cycler	Thermo Scientific
Gel Dox XR+ Gel Documentation System	Bio-Rad
Constant Voltage Power Supply (Model 1000/500)	Bio-Rad
Thermomixer comfort	Eppendorf
Centrifuge 5424	Eppendorf
Centrifuge 5415D	Eppendorf
Heraeus Multifuge 3S-R	Heraeus
Heraeus Multifuge X1 Centrifuge Series	Thermo Scientific
Sorvall RC6 centrifuge	Thermo Scientific
DeNovix DS-11 FX	DeNovix
HPLC Prominence LC20 System	Shimadzu
SPD-M20A UV/VIS Photodiode Array Detector	Shimadzu
RID-10A Differential Refracotmetric Detector	Shimadzu
MALS Dawn8+	Wyatt
Autoclave HiCLave HG-80	HCM Europe
Incubator and shaker 3033	GFL
Incubator Heracell 240i	Heraeus
Scale XS204	Mettler Toledo
Scale Entris	Sartorius
Vortex mixer Genie 2	Scientific Industries Inc.
Pipetboy Pipetus	Hirschmann Laborgeräte
CytoFLEX S2-laser device	Beckman Coulter

3.1.2 Disposables

Table 3.2: Disposables

Disposables	**Supplier**
Serological pipettes (2, 5, 10, 25 and 50 mL)	Greiner Bio-One
CELLSTAR Polypropylene Tubes, conical (15, 50 mL)	Greiner Bio-One
Pipette tips (200 µL, 1000 µL)	Gilson
Dual filter tips (10, 20, 100, 200, 1000 µL)	Biozyme & Starlab
Petri dishes	Thermo Scientific
Microtubes (1.5 mL, 2 mL)	Eppendorf
Cryo tubes (2 mL)	Thermo Scientific
96 Well plates	Greiner
Amicon Ultra (0.5 and 15 mL) Centrifugal Filter (50 kDa MWCO)	Merck Group
0.1 μm Ultrafree MC VV centrifugal filter	Merck Group
Millipore Express PLUS 0.22 µm PES	Greiner Bio-one

3.1.3 Kit systems

Table 3.3: Kit systems

Kit	**Supplier**
Monarch PCR & DNA cleanup kit	New England Biolabs
Monarch DNA gel extraction kit	New England Biolabs
Monarch plasmid miniprep kit	New England Biolabs
NucleoBond Xtra Midi/Maxi kit	Macherey Nagel
Frozen-EZ Yeast Transformation II kit	Zymo Research
EZ-Link Sulfo-NHS-LC-LC-Biotin kit	Thermo Scientific

3.1.4 Buffers, solutions and media

Table 3.4: Complete media and solutions for bacterial and yeast cell cultivation

Complete media and solutions	**Composition**
LB	10 g/L peptone 5 g/L yeast extract 200 mM NaCl
LB agar	10 g/L peptone 5 g/L yeast extract 200 mM NaCl 10 g/L agar
Antibiotics for selection	50 µg/mL Kanamycin Ampicillin 25 µg/mL Chloramphenicol
Glucose solution (10% (w/v))	Milli-Q water 100 g/L glucose
YPD medium	10 g/L tryptone 5 g/L yeast extract 10 g/L glucose
YPD agar	10 g/L tryptone 5 g/L yeast extract 15 g/L agar 10 g/L glucose
SD-CAA medium	20 g/L glucose 6.7 g/L yeast nitrogen base 5 g/L bacto casamino acids 10.83 g/L Tri-sodium citrate dihydrate 7.4 g/L citric acid monohydrate
SG-CAA medium	20 g/L galactose 2 g/L glucose 6.7 g/L yeast nitrogen base 5 g/L bacto casamino acids 10.2 g/L Disodium hydrogen phosphate heptahydrate 8.56 g/L Sodium dihydrogen phosphate hydrate
SD-CAA agar	20 g/L glucose 182 g/L sorbitol 6.7 g/L yeast nitrogen base 5 g/L bacto casamino acids 10.2 g/L Disodium hydrogen phosphate heptahydrate 8.56 g/L Sodium dihydrogen phosphate hydrate 15.0 g/L agar

Table 3.5: Buffer compositions

Buffers	**Composition**
His-purification TALON – Equilibration buffer (pH 8)	1.86 g/L sodium phosphate monobasic 9.25 g/L Disodium hydrogen phosphate heptahydrate 17.5 g/L NaCl pH adjustments with HCl and NaOH
His-purification TALON – Wash buffer 1	Equilibration buffer 5 mM imidazole
His-purification TALON – Wash buffer 2	Equilibration buffer 15 mM imidazole
His-purification TALON – Elution buffer	Equilibration buffer 250 mM imidazole
His-purification TALON – Regeneration buffer	3.94 g/L 1,3,5-trimethylbenzene 17.4 g/L NaCl
PBSA	PBS 0.1% BSA
SEC running buffer	PBS + additional 200 mM NaCl
His-purification Bio-Rad – Equilibration buffer	50 mM phosphate buffer pH 7.4 500 mM NaCl 10 mM Imidazole
His-purification Bio-Rad – Elution buffer	50 mM phosphate buffer pH 7.4 500 mM NaCl 500 mM Imidazole
His-purification Bio-Rad – Stripping buffer	PBS 50 mM EDTA
His-purification Bio-Rad – Regeneration solution	0,1 M $NiCl_2$

3.1.5 Plasmids

Table 3.6: Plasmids

Plasmid	**Resistance**	**Company**
pE-SUMO	Kanamycin	Life Sensors
pCTCon2	Ampicillin	Genscript
HMBP-3C-STAR.66-285	Ampicillin	Addgene (plasmid 100094)

3.1.6 Enzymes and antibodies

Table 3.7: Enzymes

Enzyme	**Company**
NotI-HF restriction enzyme	New England Biolabs
BamHI-HF restriction enzyme	New England Biolabs
HindIII restriction enzyme	New England Biolabs
Q5 Hot Start High-Fidelity DNA Polymerase	New England Biolabs
NEBuilder HiFi DNA Assembly Master Mix	New England Biolabs

Table 3.8: Antibodies for flow cytometry

Antibody	**Company**
Mouse anti-c-myc (clone 9E10)	Thermo Fisher Scientific
Goat anti-mouse IgG-Alexa Fluor 488 (polyclonal)	Thermo Fisher Scientific
Anti-HA-Alexa Fluor 488 (clone 16B12)	BioLegend

3.1.7 Yeast and bacterial strains

Table 3.9: Yeast and bacterial strains

Strain	**Company**
E.coli XL-10 competent cells	New England Biolabs
E.coli BL21 Rosetta (DE3) competent cells	Sigma-Aldrich
E.coli DH5α competent cells	Thermo Fisher Scientific
E.coli C41 (DE3)	Sigma-Aldrich
S. cerevisiae EBY100	ATCC

3.1.8 Software

Table 3.10: Software used for data analysis and illustrations

Software	Company
Astra	Cesbo Ltd.
LabSolutions	Shimadzu Corporation
ProtParam online tool	Expasy
FlowJo Version 10	FlowJo LLC.
PyMOL	Schrödinger
Microsoft Excel for Windows 2010	Microsoft

3.1.9 Miscellaneous material

Table 3.11: Miscellaneous material

Material	Supplier
Nuclease-free water	Thermo Scientific
dNTP's	Biozym
Agarose	Biozym
GeneRuler 1kb Plus DNA Ladder	Thermo Scientific
SYBR safe	Thermo Scientific
Dialysis tubing	Sigma-Aldrich
TALON metal affinity resin	Takara
Precision Plus Protein Unstained Protein Standards	Bio-Rad
10xtris/glycine/SDS (TGS) buffer	Bio-Rad
tris-acetate-EDTA (TAE) buffer	Thermo Scientific
Mini-PROTEAN TGX Precast Protein Gels	Bio-Rad
4x Laemmli sample buffer	Bio-Rad
CutSmart Buffer	New England Biolabs
high-performance liquid chromatography (HPLC) tubes	Bruckner Analysentechnik
Superdex 75 Increase 10/300 GL	Cytiva
Syringe filters (0.22 µm)	Thermo Scientific
BSA, ovalbumin, myoglobin, cytochrome c for HPLC standard	Sigma-Aldrich
PD-10 Desalting columns	GE Healthcare Life Sciences
Dynabeads Biotin binder (lot 00883190)	Thermo Fisher Scientic
Glass beads (diam. 5 mm)	Sigma Aldrich
Isopropyl b-D-1-thiogalactopyranoside (IPTG)	Sigma Aldrich

3.1.10 Primer sequences

3.1.10.1 Primers

Table 3.12: Primers used for DNA amplification. All of them were ordered at the company Microsynth. The capitalized sequence part matches with the respective template, while the lower case part represents the overhang.

Primer	**Sequence**
SUMO-R1	
fwd	5´-tggtaccggtctcactagagGCAACCGTGAAATTCACATAC
rev	5´-tggtggtggtgctcgagtgcTTATTGCTTTCCCAGCATCT
SUMO-L	
fwd	5´-tggtaccggtctcactagagGTGAGCGACGTGCCCAGG
rev	5´-tggtggtggtgctcgagtgcCTGGCTGGGCCTGTCGATC
MBP-sirukumab	
fwd	5´-tctgttccaggggcccggatccGAAATTGTTTTAACGCAGTC TCCTG
rev	5´-cggccagtgccaagcttttaGCTGCTGACAGTCACGGT
pCTCon2-bococizumab	
fwd	5´-gaggcggagggtcggctagcATGGAACTGGGGTTGAGTTG
rev	5´-ataagctttgttcggatccGTGGTGGTGGTGATGATG
Sequencing primers	
SUMO-5	5´-CCTTAAGATTCTTGTACGACG
HMBP-3C-STAR.66-285_seq	5´-AACAACCTCGGGATCGAG
Con2_seq	5´-CGTTTGTCAGTAATTGCGGTTCTC

3.1.11 Amino acid sequences of expressed proteins

Table 3.13: Amino acid sequences of expressed proteins

Fusion protein	**Sequence**
SUMO-R1 [19]	
HIS-Tag	HHHHHH
SUMO	GSLQDSEVNQEAKPEVKPEVKPETHINLKVSDGSSEIFFKIKKTTPLRRLMEAFAKRQGKEMDSLRFLYDGIRIQADQAPEDLDME DNDIIEAH
Linker	REQIGG
R1	ATVKFTYQGEEKQVDISKIKWVIRWGQYIWFKYDEDGGAKGWGYVSEKDAPKELLQMLGKQ
SUMO-E5 [18]	
HIS-Tag	HHHHHH
SUMO	GSLQDSEVNQEAKPEVKPEVKPETHINLKVSDGSSEIFFKIKKTTPLRRLMEAFAKRQGKEMDSLRFLYDGIRIQADQAPEDLDME DNDIIEAH
Linker	REQIGG
E5	ATVKLTYQGEEKQVDISKITYVDRAGQFIWFEYDEGGGALGTGWVSEKDAPKELLQMLEKQ
SUMO-L [16]	
HIS-Tag	HHHHHH
SUMO	GSLQDSEVNQEAKPEVKPEVKPETHINLKVSDGSSEIFFKIKKTTPLRRLMEAFAKRQGKEMDSLRFLYDGIRIQADQAPEDLDME DNDIIEAH
Linker	REQIGGATAKFTYQGEEKQVDISKIKRVARYGQGIYFSYDEGGGAYGYGS
L	VSDVPRDLEVVAATPTSLLISWRGCPWAIYYGVTYGETGGSSLAQEFTMPGVTNATISGLEPGVDYTITVYAVTRVGRMLCAPGP ISINYRTEIDRPSQ
MBP-S [39, 23]	
HIS-Tag	HHHHHH
MBP	MKTEEGKLVIWINGDKGYNGLAEVGKKFEKDTGIKVTVEHPDKLEEKFPQVAATGDGPDIIFWAHDRFGGYAQSGLLAEITPD KAFQDKLYPFTWDAVRYNGKLIAYPIAVEALSLIYNKDLLPNPPKTWEEIPALDKELKAKGKSALMFNLQEPYFTWPLIAADG GYAFKYENGKYDIKDVGVDNAGAKAGLTFLVDLIKNKHMNADTDYSIAEAAFNKGETAMTINGPWAWSNIDTSKVNYGVTVLP TFKGQPSKPFVGVLSAGINAASPNKELAKEFLENYLLTDEGLEAVNKDKPLGAVALKSYEEELAKDPRIAATMENAQKGEIMP NIPQMSAFWYAVRTAVINAASGRQTVDEALKDAQ
Linker	TNSSSNNNNNNNNNNLGIEGRISEFLEVLFQGPGS
sirukumab	EIVLTQSPATLSLSPGERATLSCSASISVSYMYWYQQKPGQAPRLLIYDMSNLASGIPARFSGSGSGTDFTLTISSLEPEDFA VYYCMQWSGYPYTFGGGTKVEIKGSTSGSGKPGSGEGSTKGEVQLVESGGGLVQPGGSLRLSCAASGFTFSPFAMSWVRQAPG KGLEWVAKISPGGSWTYYSDTVTGRFTISRDNAKNSLYLQMNSLRAEDTAVYYCARQLWGYYALDIWGQGTTVTVSS

3.2 Methods

3.2.1 Molecular biology methods

3.2.1.1 Polymerase chain reaction (PCR)

The PCR reactions were performed using the Q5 Hot Start High-Fidelity DNA Polyerase (New England Biolabs). They were prepared according to the manufacturer's manual in a reaction volume of 50 µL (Table 3.14). The PCR product was analyzed with agarose gel electrophoresis and purified using the Monarch PCR & DNA cleanup kit.

Table 3.14: General setup of PCR reactions

Component	**Volume**
5x Q5 Reaction Buffer	10 µL
dNTP's [10 mM]	1 µL
Template DNA [10 ng/µL]	1 µL
Forward primer [10µM]	2.5 µL
Reverse primer [10µM]	2.5 µL
Q5 High-Fidelity DNA Polymerase	0.5 µL
Nuclease-free water to	50 µL
Total Volume	50 µL

Table 3.15: PCR cycle conditions

Step	**Temperature**	**Time**
Initial denaturation	98 °C	0:30
Denaturation	98 °C	0:10
Annealing	60-72* °C	1:00
Extension	72 °C	1:00
Final extension	72 °C	10:00
End	4 °C	∞
* = annealing temperature of respective primers		
For the PCR reactions, 25-35 cycles were applied		

3.2.1.2 Restriction digest

For restriction digests of the used vectors, 1-2 µg plasmid and 10 U of the respective restriction enzyme were used. 5 µL of cut smart buffer from New England Biolabs were added and the reaction volume was adjusted to 50 µL with ultrapure water. Digestions were performed for 1-2 h at 37°C. After the digestion, a column purification was performed using the Monarch PCR & DNA cleanup kit to remove the restriction enzymes.

3.2.1.3 HiFi DNA assembly

The assembly of DNA fragments was performed with the NEBuilder HiFi DNA Assembly Master Mix. The reactions were prepared according to the manufacturer's recommendations. Accordingly, molar ratios of vector to insert were 1:2. The used amount of vector and insert depended on their concentration & length. Conversion of mass to mole for vector and insert was done with the New England Biolabs online calculator. A general HiFi DNA assembly reaction setup is given in Table 3.16.

Table 3.16: General reaction setup for Hifi DNA assembly

Component	**Volume**
Vector and insert [molar ratio = 1:2]	X µL
Deonized water	5-X µL
Hifi DNA Assembly Master Mix	5 µL
Total Volume	10 µL
X = 1-5 µL of vector and insert mix depending on their respective concentration [0.03-0.2 pmol]	

3.2.1.4 Agarose gel electrophoresis

For the preparation of the agarose gels, 1 % (w/v) agarose were weighed in and the according amount of TAE buffer was added. The suspension was heated in the microwave at 800 Watt for around 2 minutes to dissolve the agarose properly. SYBR safe DNA stain gel (10000x) was added to the solution and after solidification the agarose gel was transferred to a gel electrophoresis chamber, where it was fully immersed into TAE buffer. As a reference, the GeneRuler 1kb Plus DNA Ladder from ThermoScientific was loaded next to the investigated samples. For analytical gels, premixed samples containing loading dye had a volume of 12 µL, while for preparative gels it was 60 µL. To run the electrophoresis, a voltage of 120 V was applied for approximately 45 minutes. Afterwards, the gels were imaged with the Gel Dox XR+ Gel Documentation System from Bio-Rad.

3.2.1.5 Heat shock transformation

To perform a heat shock transformation of chemically competent cells, 50 µL aliquots of the desired *E. coli* strain were thawed on ice. Next they were mixed with 2-10 ng plasmid DNA. HiFi DNA assembled constructs were diluted 1:4 and thereof 2 µL were used for the transformation. The mixture was placed on ice for 20 more minutes, followed by a heat shock at 42°C for 45 s and a recovery period of 2 min again on ice. Subsequently, 500 µL antibiotic free LB medium was added and the cells were incubated for around 40 minutes at 37°C while shaking (300 rpm). After the incubation, 100 µL of the transformed cells were plated on LB-plates using sterile glass beads. Alternatively, the cells were spun down, the supernatant was poured off, the cell pellet was resuspended in the remaining

medium in the tube and all of the cell suspension was plated. The LB plates contained the antibiotics corresponding to the resistances in the plasmids and resistances of the strains and were incubated overnight at 37°C to enable colonies to grow.

3.2.1.6 Frozen yeast transformation

For the transformation of DNA into the *S. cerevisiae* strain EBY100, the Frozen-EZ Yeast Transformation II kit from Zymo Research was used. According to the manufacturers' manual, for the preparation of transformation competent cells a yeast culture was grown at 30°C in YPD broth until mid log phase. Thereafter, the cells were centrifuged (5000 g, 4 minutes, room temperature), washed once with EZ 1 solution and finally resuspended in EZ 2 solution and were ready for transformation.

50 µL aliquots of competent cells were mixed with 0.2-1 µg DNA. 500 µL EZ 3 solution was added and mixed thoroughly with the cells before they were incubated at 30°C for 45 minutes. Subsequently, 100 µL of the transformation mix were spread over an SD-CAA plate and grown at 30°C for 2-4 days. The solutions EZ 1 to 3 were provided by the manufacturer Zymo research.

3.2.1.7 SDS PAGE

All the SDS PAGEs were performed using stain-free precast gels. 15 µL sample were mixed with 5 µL Laemmli buffer (4x) and placed on a heat block for 5 minutes at 95°C. In addition to the samples, 8 µL unstained protein standard was loaded on the gel. The samples contained 4-20 µg of protein each. By applying 180 V for around 25 minutes the gel was run and was imaged with the Gel Dox XR+ Gel Documentation System from Bio-Rad.

3.2.1.8 Cryo stocks

For the creation of cryo stocks, the bacterial or yeast culture was mixed in a 1:1 ratio with a solution containing 30 % glycerol. The resulting glycerol concentration was 15 %. Cryo stocks were placed at -80°C for storage immediately after mixture.

3.2.2 Expression and purification

3.2.2.1 Expression

The soluble expression of the proteins to be coupled to the magnetic streptavidin-coated Dynabeads (Thermo Fisher Scientific) was performed in *E. coli* strains. The scFv sirukumab fused to the maltose-binding protein of *E. coli* was expressed in C41, while the expression strain used for the Sso7d binders fused to the small ubiquitin-like modifier (SUMO) protein was the Rosetta (DE3) strain. The corresponding expression strains were transformed with the sequence-verified plasmids. After inoculation, the cultures were grown overnight at 37°C in LB medium containing the antibiotics corresponding to the resistances in the plasmids and the expression strains used. On the next day, cultures

were diluted in LB medium to an OD600 of 0.1-0.2, again containing the antibiotics and further cultivated until an OD600 of 1-2 was reached. Then, the protein expression was induced by addition of 1 mM isopropyl beta-D-1-thiogalactopyranoside (IPTG) and continued overnight at 20°C while shaking (180 rpm).

Bacterial cells were harvested by centrifugation at 5000 g for 20 min, 4°C. After resuspension in sonication buffer, cells were lysed by using ultrasound sonication (2 x 3 min, 50 % duty cycle, amplitude set to 95) on ice. Subsequently, the cell debris were spun down at a high speed (20000 g, 30 min, 4°C). The supernatant containing the protein of interest was filtered with a 0.45 µM filter and was supplemented with 10 mM imidazol to prevent unspecific binding to the purification resin.

3.2.2.2 Talon gravity flow purification

All the HIS-tagged proteins except MBP-sirukumab (MBP-S) were purified via immobilized metal affinity chromatography (IMAC) using gravity flow glass columns filled with TALON metal affinity resin. For each purification 1.5-2 mL resin were used. It was washed with RO water (2x20 mL), then equilibrated with 20 mL Equilibration buffer. The sample containing 10 mM imidazol to avoid unspecific binding was loaded onto the column twice. After the sample application the resin was washed with washing buffer 1 (5 mM imidazol) and 2 (15 mM imidazol), 20 mL each. The elution was performed by using 2x5 mL elution buffer (250 mM imidazol). The resin was cleaned by application of 3x20 mL regeneration buffer, followed by 2x20 mL RO water and was stored in 20 % ethanol at 4°C. To perform a buffer exchange, the purified proteins were filled in dialysis tubings (MWCO 10000, Sigma-Aldrich) and dialysed overnight in 5 L PBS at 4°C. Protein concentrations before and after buffer exchange were determined with a DeNovix spectrophotometer. Extinction coefficents and molecular weights of the proteins used for the measurements were determined with the online tool ProtParam and are listed in the following Table 3.17. The purified proteins were stored at -80°C.

Table 3.17: Extinction coefficients and molecular weights of solubly expressed proteins calculated with the online tool ProtParam.

Fusion protein	**Exctinction coefficient [M^{-1} cm^{-1}]**	**MW [Da]**
SUMO-R1	29450	19376
SUMO-E5	16960	19100
SUMO-L	30495	27863
MBP-S	127450	70888

3.2.2.3 Purification with the Bio-Rad system

For the fusion protein MBP-sirukumab, the purification was performed using the Bio-Rad purification system. The HIS column used for the purification was equilibrated with equilibration buffer containing 10 mM imidazole. Subsequently, the sample was applied to

the column and the MBP-S protein could bind to the column via its HIS-tag. Then, the column was washed with equilibration buffer until the UV 280 signal was approaching the baseline. The elution of the sample was performed by applying a linear gradient to the HIS column, where the concentration of elution buffer containing 500 mM imidazol was gradually increased. The eluted fractions were collected by a fraction collector. During the whole purification the UV signal at 280 nm was measured. The protein concentration was measured as described above (see section 3.2.2.2) and buffer exchange to PBS was performed with Amicon centrifugal filters.

3.2.3 Biotinylation of solubly expressed proteins

Biotinylation of the proteins coupled to the magnetic streptavidin-coated Dynabeads (Thermo Fisher Scientific) was performed by using the biotinylation kit EZ-Link Sulfo-NHS-LC-LC-Biotin of Thermo Scientific. N - Hydroxysuccinimide (NHS) esters of biotin associate to the proteins' N-terminus and primary amine containing amino acids (Lys-residues). Immediately before usage, a solution of 10 mM biotin reagent was prepared by dissolving 1 mg of the biotin reagent in 300 µL ultrapure water. A 5 times molar excess of biotin was mixed with the respective amount of protein solution and incubated at room temperature for 30 minutes. To bind all the excessive free esters a final concentration of 3 mM TRIS-HCl buffer was added to the mixture of protein and biotin reagent. Desalting and removal of the biotin was performed via a PD10 desalting column. The column was equilibrated via application of 5 x 4 mL of PBS followed by the application of the sample. Finally, the biotinylated protein was eluted with 3.5 mL PBS. The concentration of the biotinylated proteins were determined with a DeNovix spectrophotometer and the proteins were stored at -80°C.

3.2.4 Size exclusion chromatography (SEC) - multi angle light scattering (MALS)

For SEC-MALS measurements, the investigated protein sample was thawed and centrifuged to spin down potential precipitates (12000 g, 10 minutes, 4°C). Samples were diluted with PBS containing additional 200 mM NaCl. Finally, between 25 and 100 µg of protein were applied to a Superdex 75 column (10 mm x 300 mm) which was connected to an HPLC Prominence LC20 system. Protein elution was performed with a flow rate of 0.75 mL/min (25°C). Detection followed by a UV/VIS Photodiode Array Detector (Shimadzu) and a MALS Dawn 8+ detector (Wyatt).

3.2.5 Yeast surface display methods

3.2.5.1 Cultivation and induction of yeast cells

For the surface display of the Sso7d binders as well as the scFvs, the corresponding genes were cloned into the pCTCON2 vector using the NheI and BamHI restriction sites. Then, the *S. cerevisiae* strain EBY100 was transformed with the vector by Frozen yeast transformation as described above (see section 3.2.1.6). The transformed EBY100

cells were cultivated at 30°C overnight in SD-CAA medium while shaking, subsequently diluted to an OD600 of approximately 0.1 and further grown until an OD600 of 1-2 was reached. Then, the cell surface expression was induced by spinning down the yeast cells and resuspending them in SG-CAA medium. The cells were further cultivated overnight at 20°C and could then be used for yeast surface display experiments.

3.2.5.2 Yeast cell harvest and staining

To harvest the appropriate amount of yeast cells, first the cell density of the culture was determined by OD600 measurement. The corresponding volume of the cell culture was centrifuged (2000 g, 4 minutes, 4°C) and the cells were resuspended in PBSA, washed once with 1 mL PBSA and finally taken up in PBSA again. A ratio of 10 % of displaying yeast cells were spiked into a non-displaying yeast population. After the bead selection, the cells were seeded into a 96 well plate, where the cells were spun down (2000 g, 4 minutes, 4°C) and thereafter stained by the addition of an antibody specific to the HA tag or the C-myc tag on the yeast cell surface. The HA staining was performed by adding a final concentration of 2 µg/mL of the anti HA antibody (Alexa Fluor 488 clone 16B12) from BioLegend. Staining of the c-myc tag was performed with a primary staining step using a mouse anti-c-myc antibody (clone 9E10), followed by secondary staining with 1 µg/mL of a goat anti-mouse IgG antibody (conjugated to Alexa Fluor 488). For all the staining steps, stained samples were incubated for 20 minutes at 4°C while shaking. After the staining, the yeast cells were washed once again and were ready for the evaluation at the CytoFLEX.

3.2.5.3 Bead selection assay

For the bead selection assay, 36 pmol of biotinylated protein were mixed with 10 µL magnetic streptavidin-coated Dynabeads in 2 mL reaction tubes. The reaction volume was adjusted to 100 µL with PBSA. Then, the mix was incubated for at least 1 h at 4°C while rotating in order to coat the beads with the biotinylated protein. After the incubation, a washing step with PBSA was performed to remove any free antigen followed by a resuspension of the beads in 10 µL PBSA. To rule out any effects of the beads themself, in parallel all the steps were performed with bare beads too without adding any biotinylated protein. The reaction volumes were scaled up in accordance with the amount of conditions tested. For washing stepts and separations of the beads from the supernatant in general, tubes were placed on a magnetic rack for 2-5 minutes.

The yeast cells were harvested as described above (see section 3.2.5.2). Then, for every yeast clone that was tested, 10^7 & 10^6 cells containing a ratio of 10% displayer were taken up in a volume of 1 mL PBSA and respectively mixed with 10 µL protein coated beads. For comparison, a non-selected condition was also tested. There, no beads where added to the yeast cells. For all the conditions, the reaction mix was incubated for at least 2 h at 4°C with constant rotation. After the incubation, a negative selection was performed by applying the suspension containing the beads and the yeast cells to the

magnetic rack (2 minutes, 4°C). Thereby, the magnetic beads were pulled towards the magnet and the supernatant was taken off. Thereafter, an HA staining of the supernatant was performed as described above.

The frequency of displaying cells present after the negative selection with bare beads, protein coated beads and in the non-selected condition were evaluated by flow cytometry.

3.2.5.4 Evaluation by flow cytometry

All the flow cytometry measurements were performed with the CytoFLEX S2 - laser device in the plate sample mode. The data analysis was done with the FlowJo Version 10 software.

Chapter 4

Results

4.1 Selection of the proteins associated to the streptavidin beads

Next to high affinity recognition of a target antigen, the absence of non-specific binding is of crucial importance for therapeutic binders to be successful in clinical development. The aim of this project was to establish a bead selection based assay that could assist in the reduction of such non-specifically interacting protein variants present in a diversified library. We hypothesised that those undesired clones could be captured and negatively selected by sticky proteins immobilized on the surface of streptavidin coated Dynabeads. Accordingly, we selected a set of 4 proteins that together cover all the main characteristics associated with non-specific interaction, respectively, coupled them to the magnetic beads and tested their ability to capture non-specific binders in a YSD sample. Given that we expected some of those proteins to be unstable and prone to aggregation, they were expressed as fusion proteins. While three of them were expressed in fusion with the small ubiquitin-like modifier (SUMO) protein, the sirukumab-based scFv was fused to the maltose binding protein (MBP) of *E. coli* as it was shown before to enhance cytoplasmic scFv expression regardless of the redox state of the bacterial cytoplasm [38].

The selection included two reduced charge variants of the Sso7d protein from the hyperthermophilic arachaeon *Sulfolobus solfataricus*. This small alternative scaffold (7 kDa) is known to be very stable and exhibits a rigid beta sheet structure which acts as binding interface (see Figure 4.1) [17]. As mentioned above it was previously engineered to bind a variety of targets in the micromolar to nanomolar range [10]. Using YSD, Traxlmayr et al. engineered both the Sso7d based proteins expressed for this project [18, 19]. First, the wild-type Sso7d protein was charge neutralized. Subsequently, by mutating nine positions within three adjacent beta strands, two YSD libraries were constructed and enriched for EGFR binding [18].

One of them yielded E18.4.5 (named E5 within this project), an EGFR binding variant rich in hydrophobic amino acid content [18]. Hydophobic amino acids as single residues but more importantly as part of motifs are among the main factors associated

with non-specific binding [24]. By enrichment of the two mentioned YSD libraries for binding to K-Ras-G12D, the second Sso7d based protein of our data set R1 (originally named R11.1) was obtained [19]. Its binding interface is highly hydrophobic. In fact, 6 of the 9 interacting residues are aromatic amino acids. In addition, the binding site contains two positively charged lysin residues. Intriguingly, when measured by SEC in previous projects, the K-Ras binding protein interacted with the inert column material in such a strong manner that no elution could be observed. This behaviour underlines the susceptibility of R1 to non-specific binding and led us to the assumption that the protein could be a well suited candidate to deplete sticky binders from a library [19]. E5 as well as R1 were expressed in fusion with SUMO.

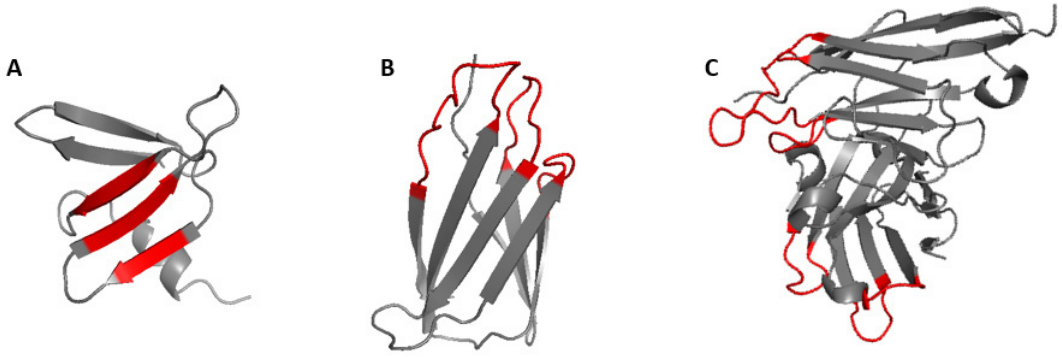


Figure 4.1: Exemplary scaffold structures of expressed proteins. Interacting beta strands of (A) rcSso7d (K-Ras specific mutant shown, PDB download: 5UFQ [19]) and loops of the (B) FN3 (wild-type FN3 shown, PDB download: 1TTG [40]) & the (C) scFv scaffold (ErbB2 specific scFv A21 shown, PDB download: 2GJJ [41] are highlighted in red. The structures were generated using Pymol Molecular Graphics System [42].

Additionally, the 10th type III domain of human fibronectin (FN3) based binder L8.5.2 that was engineered to recognize lysozyme was expressed in fusion with the SUMO protein and accordingly named SUMO-L [16]. The FN3 scaffold is a small (10 kDa) beta sandwich structure. In contrast to Sso7d binders, fibronectin based proteins interact with their binding partners via solvent exposed loops similar to an antibody complementarity determining region (CDR) as shown in Figure 4.1. The FN3 based binder L was engineered by Hackel et al. in 2008 [16]. Analytical SEC data of the group indicate that the protein was partially in a oligomeric state (80 % monomeric). Additionally, its delayed elution suggests non-specific interaction with the column matrix [16].

Finally, we included the single chain variable fragment (scFv) based on sirukumab. ScFvs are the variable domains of an antibody's heavy chain and light chain expressed as fusion protein and connected via a flexible linker. Figure 4.1 shows an exemplary scFv

scaffold. Interaction with potential bindig partners occurs via its flexible CDR loops shown in red. As mentioned above, the sirukumab-based scFv was expressed in fusion with the maltose binding protein of *E. coli* [39]. In 2017, the company Adimab performed a study with an extensive dataset of 137 antibodies that were either approved or in advanced clinical stage. Exclusively expressed as human IgG1 isotypes and therefore only different in their variable region, their performance was analysed in 12 biophysical property assays. The goal of the study was to define a practical guide for future antibody drug candidates regarding drug-like behaviour, which is recently gaining attention as an additional critical feature for a therapeutic binders' clinical success next to antigen recognition [23].

The dataset analysed by Adimab included a sirukumab-based IgG1 scFv which performed poorly in all of the biophysical property assays that test for specificity and aggregation behaviour. Therefore, this scFv was an ideal candidate to complete our dataset. Furthermore, its increased size (27 kDa vs. 7/10 kDa, for Sso7d and FN3-based binders, respectively) compared to the other proteins tested on the streptavidin beads would give as a hint about potential influences of the protein size on the function of our assay.

4.1.1 Design of solubly expressed fusion proteins

All proteins that were tested on the streptavidin beads were expressed together with an N-terminal fusion partner that should enhance expression levels and stability. As mentioned above, in the case of the Sso7d & FN3 based proteins this was the SUMO protein, while the sirukumab-based scFv was expressed in fusion with MBP. In between the fusion partners is a stretch of amino acids referred to as linker which has a different length depending on the restriction sites used to clone the respective binder into the SUMO or MBP vector. The fusion proteins comprised an N-terminal hexahistidine tag (6xHIS) used for purification (Figure 4.2). For the full amino acid sequences of the respective fusion proteins, see Table 3.13.

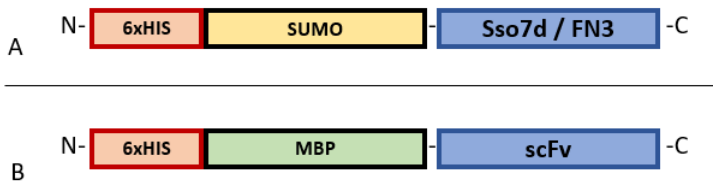


Figure 4.2: Design of solubly expressed proteins. A - Design of Sso7d & FN3 based proteins consisting of an N-terminal hexahistidine tag (6xHIS) followed by SUMO, a linker and the Sso7d/FN3 protein. B - Design of scFv fusion protein consisting of an N-terminal hexahidine tag (6xHIS), MBP, a linker and the sirukumab-based scFv. Full amino acid sequences can be found in Table 3.13.

4.2 'Sticky' and 'non-sticky' binders for the evaluation of our bead selection assay performance

To evaluate, if our bead selection assay was suitable for the reduction of non-specific binders present in any given library, we started off using a set of either sticky or non-sticky proteins with known characteristics. The hypothesis was that a negative selection with our assay would result in a reduction of yeast clones present in a sample of non-specifically interacting (sticky) binders. However, a sample of non-sticky displayers should not be altered by an exposure to the protein coated beads.

Accordingly, we cloned the DNA sequences of 5 different binders into the pCTCON2 vector as described in subsubsection 3.2.5.1. By transforming yeast cells with the vector, the respective binders were displayed on the yeast cell surfaces sandwiched between an HA- and a C-myc tag. Then, the selection assay was tested with every combination of binders on the yeast cells and protein on the beads. Four of the mentioned binders with known properties on the yeasts were Sso7d based. Two of them were R1 and E5, the Sso7d proteins also coupled to the beads as mentioned before (see section 4.1). Thereby, potential effects of self-interaction could be evaluated. Moreover, two non-sticky EGFR-binders were tested [18]. Finally, also the sticky bococizumab-based scFv was tested as a displayer on yeast cells. Just like the sirukumab-based scFv, this decision was based on the Adimab study mentioned above [23]. The Table 4.1 below summarizes all the binders tested on yeast cells and their biophysical properties regarding non-specific binding.

Table 4.1: Set of binders with known properties displayed on yeast cells.

Displayed binders	**Scaffold**	**Property**	**Original name & source**
E2	Sso7d	not sticky	E11.4.1 [18]
E4		not sticky	E18.1 [18]
E5*		intermediate sticky	E18.4.5 [18]
R1*	Sso7d	very sticky, positively charged	R11.1 [19]
Bococizumab	scFv	sticky	Bococizumab [23]

* The Sso7d binders R1 and E5 were also coupled to the streptavidin beads (section 4.1) as SUMO fusion proteins

4.3 Biophysical characterization of solubly expressed proteins

After expression and purification, we wanted to evaluate the biophysical characteristics of our soluble proteins. First of all, this was done to assess if our purification had

been successful. Furthermore, we wanted to be sure that our proteins exhibited the characteristics we hypothesised to be beneficial for the capturing of non-specific binders. Through the performance of SDS PAGEs we got a first impression of what was present in the fractions of the purification runs. For a more detailed characterization of the proteins, the analysis of their aggregation behaviour and their molecular weights SEC-MALS measurements were performed. The following Table 4.2 includes protein concentrations of our purified samples and the total amount gained from the expressions.

Table 4.2: Amount of protein purified from soluble expressions.

Fusion protein	**Protein conc. [mg/mL]**	**Purified protein in total [mg]**	***E. coli* culture volume [L]**
SUMO-R1	2.8	31	0.75
SUMO-E5	4.3	48	3
SUMO-L	0.7	8	0.75
MBP-S	2.5	7	6

4.3.1 SDS-PAGE

In the following Figure 4.3, the SDS-PAGEs of SUMO-R1 and MBP-S are shown. The gel on the left corresponds to the fractions of the SUMO-R1 (A). In the load (lanes 2 - 4) we saw an overexpressed band at the expected height ($\sim$ 20 kDa). Unfortunately, some protein was lost in the flow through (lanes 5 - 7). Nevertheless, we were able to gain a significant amount of pure protein. Overall, the expressions and TALON purifications of the SUMO fusion proteins worked very well.

For the expression & purification of the scFv on the other hand, we encountered some issues. The protein was purified with a HIS column using the Bio-Rad purification system. The fractions of the run are depicted on the right SDS-PAGE below (Figure 4.3B). While we saw a band at a height that could correspond to our fusion protein ($\sim$ 70 kDa), several other bands were visible in the eluted fraction too (lane 6 and 7 with 7.5 µg and 15 µg respectively applied). However, we decided to proceed with the eluted sample and test it in the negative selection assay. Accordingly, one has to consider that the amount of purified protein listed for MBP-S in Table 4.2 corresponds to the fractions shown on the SDS-PAGE above (Figure 4.3) rather than the pure fusion protein alone.

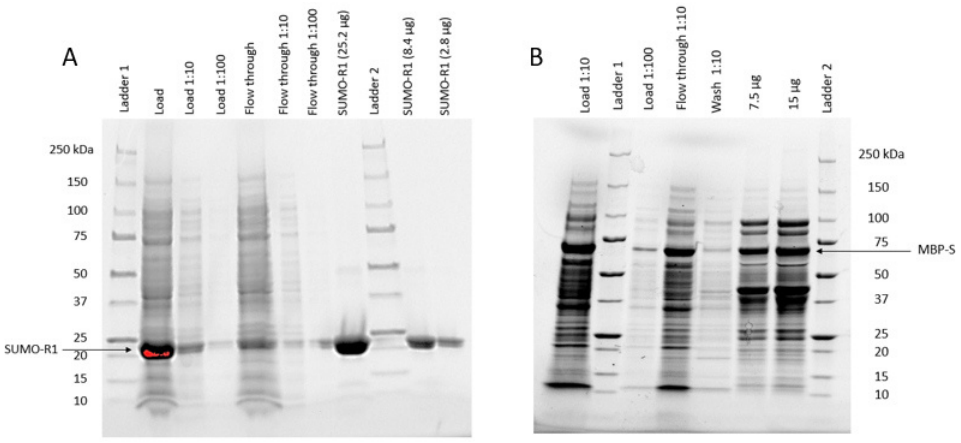


Figure 4.3: Fractions of SUMO-R1 and MBP-S purification applied on a SDS-PAGE. A - Fractions of the TALON purification run of SUMO-R1 are shown. B - Load, flow through, wash and eluted fraction of MBP-S purification with the Bio-Rad purification system.

4.3.2 SEC-MALS measurement

SEC-MALS measurement gives information about the aggregation behaviour of a protein. Additionally, protein binders that interact non-specifically with the column material are expected to show delayed elution, indicating sticky characteristics of the binders. Previously, it was shown for instance that the binder L eluted later from the inert SEC column than expected for its molecular weight due to non-specific interaction with the column material [16]. Similar effects have been observed with R1, which associated to the matrix in such a strong fashion that no elution could be measured (oral communication [18]). This observation underlines the hydrophobicity of R1 and its suitability for our application. However, it also kept us from measuring the protein again with SEC. The other 3 proteins were analysed at least twice. SUMO-E5 was even measured three times (see Figure 4.4).

Through the measured MALS data we could assign the eluted peaks to the aggregation states of our proteins. Table 4.4 summarizes the theoretical molecular weights of our proteins and the MALS determined masses of the measured peaks. Next to the determination with the MALS detector we also determined the molecular weights through a protein standard. As the MALS detector did not yield accurate data for MBP-S, we had to rely on the calculation via our standard for the sirukumab-based scFv. The chromatogram of the standard proteins is shown in Figure 4.6 (A). Their theoretical MWs as well as the MALS determined masses are listed in Table 4.3. Using the standard's retention times, a trendline was constructed (see Figure 4.6 B) which allowed us to calculate the molecular weights of our solubly expressed proteins in addition to the MALS determined masses (included in Table 4.4). The calculation over the trendline of the

standard allowed us to assign also the peaks of the measurements where the MALS detector did not function properly.

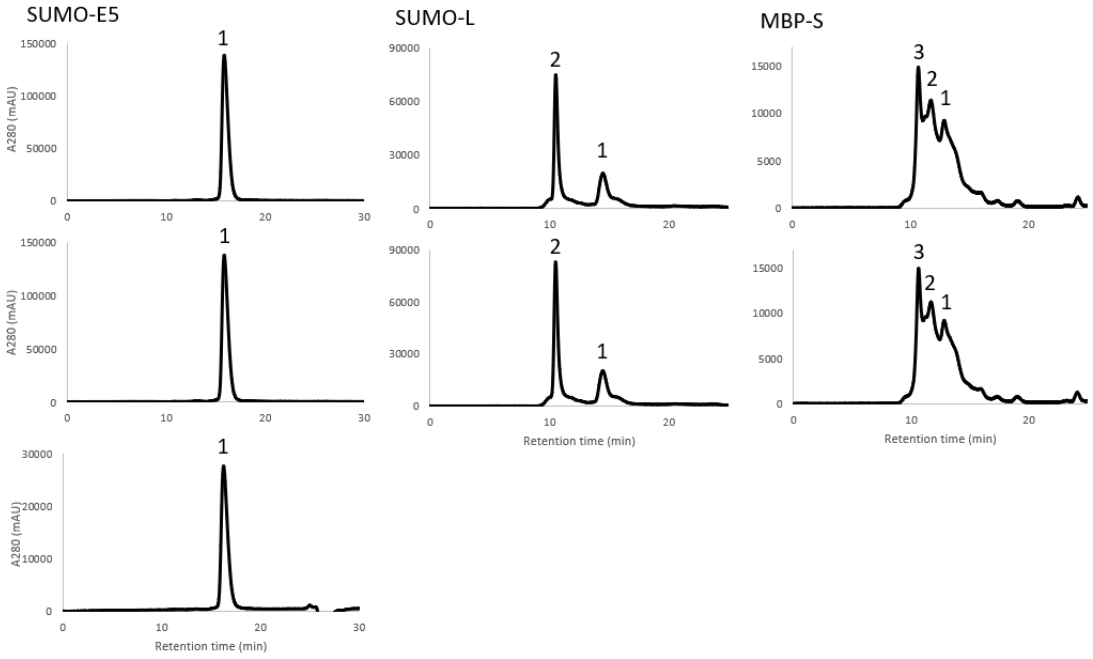


Figure 4.4: SEC data of solubly expressed proteins. Shown is the UV-signal at 280 nm. Retention times and molecular weights are listed in Table 4.4. SUMO-E5 was measured three times and 100 µg protein was respectively loaded. The other proteins were measured twice. Of SUMO-L and MBP-S 70 µg and 26 µg protein was loaded, respectively.

In all the three measurements, the Sso7d based protein E5 eluted as one clear peak (see Figure 4.4). The MW determined by MALS measurement as well as the molecular weight calculated through the standard correspond to the protein in a monomeric form. When analysed in more detail, the monomeric peak revealed an additional, typical feature of sticky proteins. The peak exhibited an asymmetric form which can be appreciated in the plot below (Figure 4.5). This phenomenon is called peak tailing and caused by non-specific interaction with the column material.

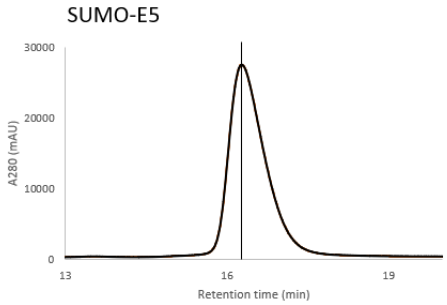


Figure 4.5: Zoomed HPLC plot of E5 showing the asymmetric form of the monomeric peak.

In the SEC-MALS analysis of SUMO-L two peaks were measured. As its MALS determined masses came with such a high uncertainty, we decided to calculate the molecular weights also using the protein standard. Nevertheless, the results were quite similar and identify peak 1 as the protein in a monomeric state, while the second peak can be interpreted as a trimer or as aggregates Figure 4.4.

Unfortunately, the strong aggregation and the fact that we could not gain a lot of purified protein made it almost impossible for us to interprete the plots of MBP-S correctly. Still, the MALS data and the masses calculated with the standard suggest that peak 2 could correspond to the fusion protein in a monomeric form. Moreover, peak 3 is probably composed of aggregates. Yet, we were not able to identify the protein that eluted in peak 1.

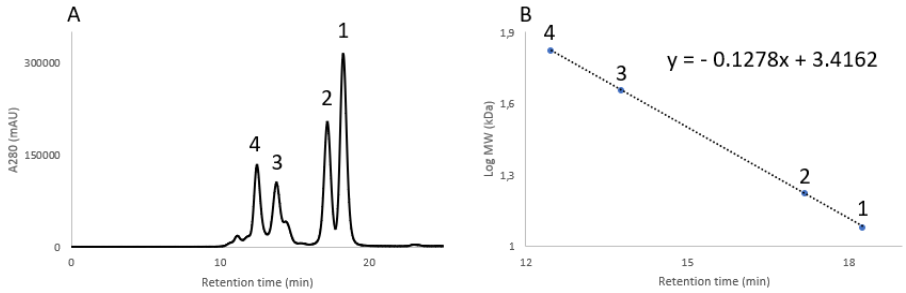


Figure 4.6: (A) SEC analysis of standard proteins (UV-signal at 280 nm) and (B) log(MW) plot against the retention time for the calculation of molecular weights. Standard proteins: (1) Cytochrome-C, (2) myoglobin, (3) ovalbumin, (4) BSA;

Table 4.3: Theoretical and MALS-determined molecular weights of standard proteins.

Standard	**Theoretical MW (kDa)**	**MALS determined MW (kDa)**
BSA (4)	66.5	70.2 (±1%)
Ovalbumin (3)	45	45.7 (±2%)
Myoglobin (2)	16.7	23 (±2%)
Cytochrome c (1)	12	15.1 (±2%)

Table 4.4: Theoretical MWs, MALS-determined MWs and MWs calculated through the calibration curve of the HPLC standard proteins. Peaks correspond to numbering in Figure 4.4. Representative results from one measurement are shown.

Protein	**Theoretical MW (kDa)**	**MALS MW (kDa)**	**Calculated MW from standard (kDa)**
SUMO-E5	19.1		
– Peak 1		24 (±8%)	25
SUMO-L	27.9		
– Peak 1		30 (±29%)	37
– Peak 2		109 (±49%)	117
MBP-S	70.9		
– Peak 1			59
– Peak 2		/	83
– Peak 3			113

4.4 Evaluation of biotinylation

To couple the expressed proteins to the beads for the assay, we biotinylated them with the biotinylation kit from Thermo Scientific (Table 3.3). Thereafter, we verified with SEC if the biotin was successfully conjugated. Samples of the biotinylated protein, a streptavidin Alexa Fluor 647 conjugate only and a mix of both (3x molar excess of streptavidin alexa fluor 647 conjugate) were analysed. The generated data is given below in Figure 4.7. First of all, the biotinylated protein only chromatograms look very similar to the SEC-MALS data of the non-biotinylated versions (Figure 4.4). As shown by the streptavidin plots in the middle, the protein fluorescent conjugate absorbs also at a wave length of 647 nm (dotted line) next to the usual protein absorption at 280 nm (continuous line). In the mixed samples on the right, streptavidin can associate in a very strong manner to all the biotin eventually present on the protein. This association increases the protein size and thereby leads to an earlier elution of the biotinylated protein. The disappearance of the original peaks indicates that an association took place. Furthermore, all the eluted fractions show also a 647 nm signal. Therefore the measurements of SUMO-E5 (A) and SUMO-L (B) included in Figure 4.7 suggest a successful biotinylation of our

fusion proteins. For the fusion protein MBP-S, similar effects were observed (data not shown). Due to the fact, that R1 could not be eluted from the column when measured with SEC-HPLC in previous projects (orally communicated [19]), the biotinylation of SUMO-R1 was not evaluated. Nevertheless, the evaluation of all the other proteins indicate a successful biotinylation reaction.

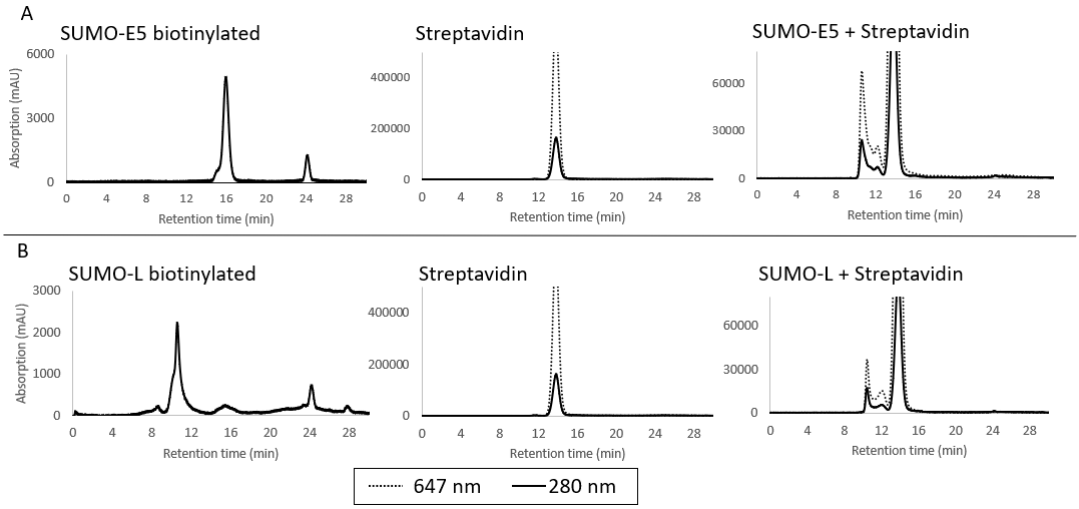


Figure 4.7: Verification of successful biotinylation. A - SEC-HPLC runs of SUMO-E5 biotinylated (5 µg), streptavidin (41 µg) and SUMO-E5 biotinylated (5 µg) + streptavidin (41 µg) from left to right. B - SEC-HPLC runs of SUMO-L biotinylated (5 µg) on the left, followed by streptavidin (28 µg) and SUMO-E5 biotinylated (5 µg) + streptavidin (28 µg).

4.5 Definition of a suitable displayer to non-displayer ratio

In a library, the undesired non-specifically interacting fraction is usually very small. Thus, also in our samples of yeast cells subjected to the bead selection assay, only a small ratio should actually display our selected binders as they represent those non-specific binders in a library. We sought to identify a suitable frequency of displayers through the performance of a titration experiment. In this flow cytometric evaluation, a dilution series from 'displayers only' down to 1:300 displayers in non-displayers was generated, where the appropriate concentration of displaying yeast cells was respectively spiked into a non-displaying population. With decreasing frequencies of displayers present, also the displaying population present in the plots was diminished (see displayers gates from left to right Figure 4.8). In the plots, the side scatter is shown on the x-axis, while on the y-axis the display level is plotted that was measured through c-myc detection. From top to bottom, you can see the dilution series of E4 (A), E5 (B) and R1 (C) and evidently, they all show the same trend. The dilutions 1:100 and 1:300 seem to be too close to the detection limit and were therefore interpreted as noise signals, as no further reduction

of displayers could be measured. Obviously, 1:3 ratio of displayers in non-displayers would not represent very well the ratio of non-specific binders present in a library as their frequency in a library is very low. On the other hand, the 1:30 displayers in non-displayers condition seemed to be fairly close to the detection limit where possible effects caused by the bead selection assay could be less evident and more difficult to quantify. Thus, we concluded from this expriment that a ratio of 10 % would suit best our purposes and performed all the bead selection assays discussed in the following section with that fraction of displayers spiked into non-displayers.

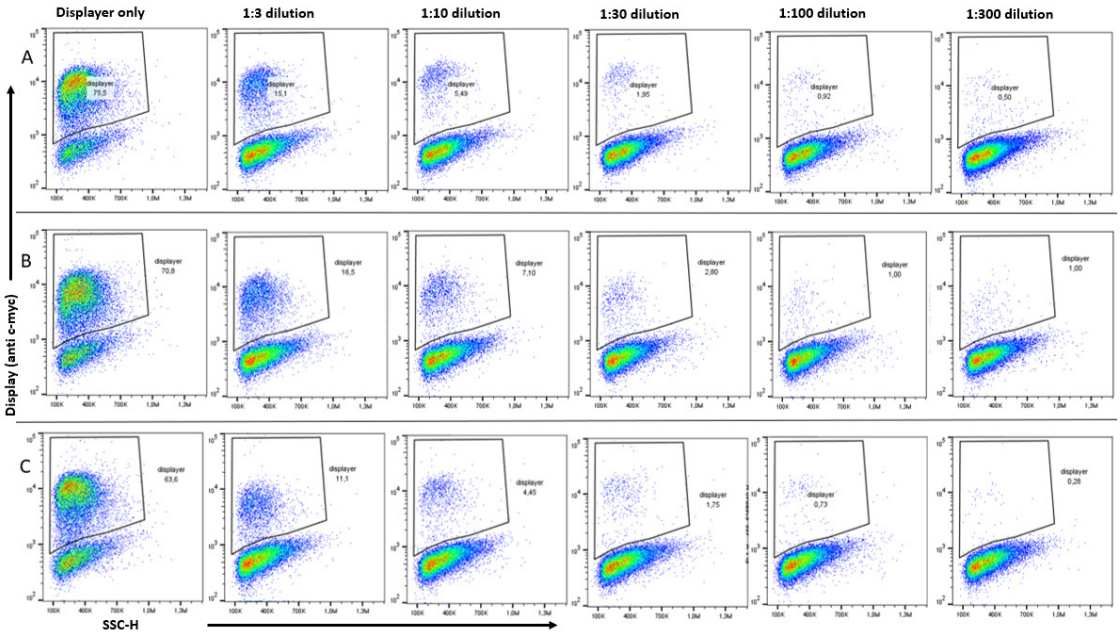


Figure 4.8: Titration experiment for the evaluation of a good displayers to non-displayers ratio. Above, a geometric series up to a dilution of 1:300 of displayers in a non-displaying population are shown. Measurements were performed for E4 (A), E5 (B) and R1 (C) displayed by yeast cells.

4.6 Bead selection assay

Our bead selection assay was hypothesised to reduce non-specifically interacting binders in a YSD sample. To evaluate its performance, it was tested with every possible combination of fusion-protein on the beads and binders on the yeast cells within our dataset (see Figure 4.9). The binder-displaying yeast cells were prepared by respectively spiking the previously determined ratio of displayers (10%) in a non-displaying population. We saw some fluctuations in this ratio which is especially evident in Figure 4.9D and may be caused by pipetting errors during the adjustment of the displayer ratio and seeding of the yeast cells into 96 well plates. Nevertheless, the samples could be evaluated regarding possible reductions in the frequency of displayer through the bead selection assay. For

every combination of proteins on the yeasts and beads, three different conditions were evaluated by flow cytometry.

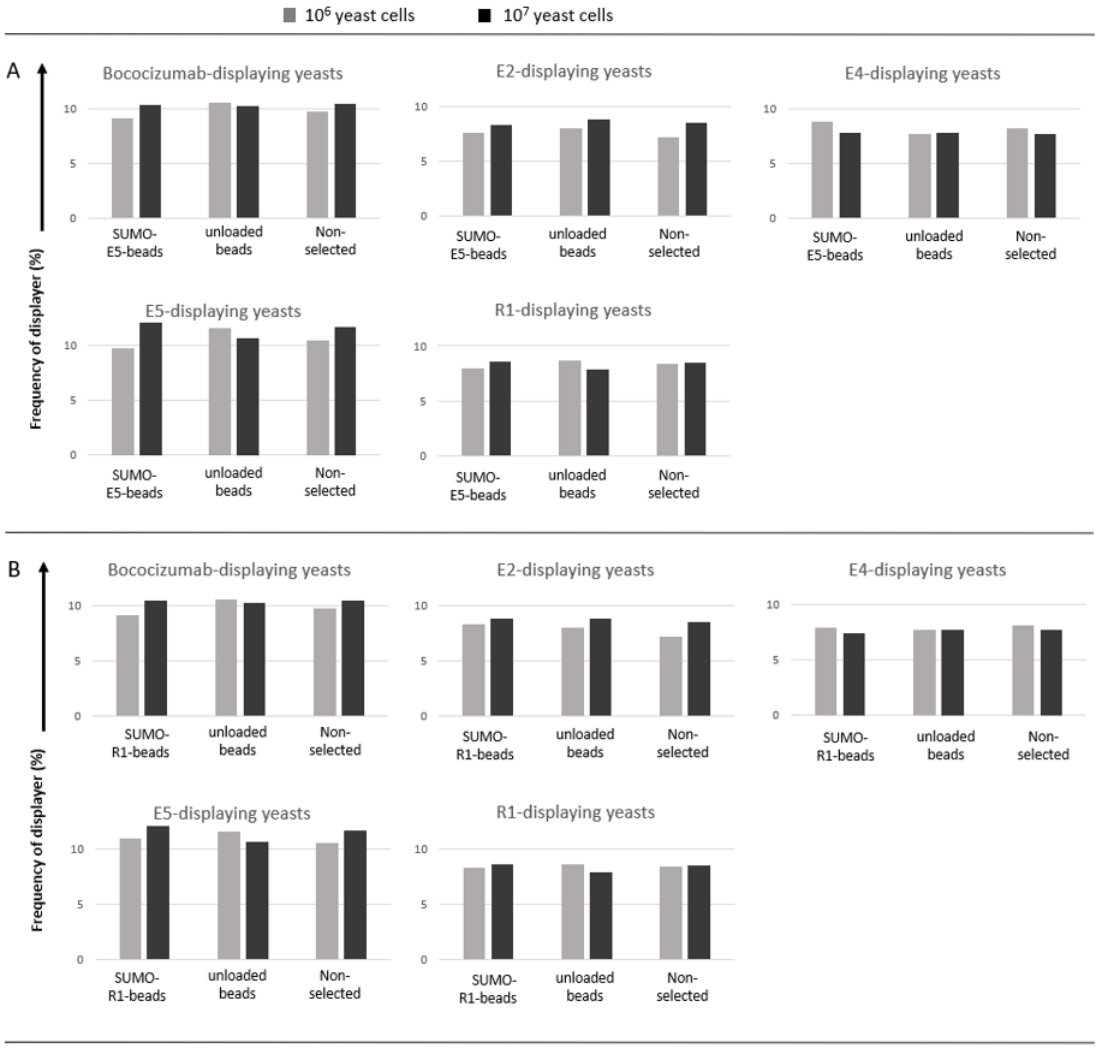
- In the first condition, the YSD sample was subjected to our assay, where the respective fusion-protein was immobilized on the surface of the magnetic beads (Figure 4.9: SUMO-E5-beads, SUMO-R1-beads, MBP-S-beads, SUMO-L-beads).
- Second, a negative selection with bare beads was performed. This condition was important, because it allowed us to evaluate if possible effects were coming from the beads themselves (Figure 4.9: unloaded beads).
- In the third and final condition, for every combination the frequency of displayers present without any kind of selection was quantified. Thereby, the fraction of displayers in the sample before and after a selection could be compared (Figure 4.9: non-selected).

For every combination of fusion-protein on the beads and binder-displaying yeasts, those three conditions were analysed with two different total yeast cell numbers in the sample, 10^6 and 10^7 (in Figure 4.9 displayed in grey and black, respectively). That allowed us to investigate potential influences of the cell numbers on the assay. Flow cytometric detection was performed via the displayers' HA tag as described above (subsubsection 3.2.5.2).

Figure 4.9 includes the average values from two bead selection experiments for every combination possible in our dataset. The three conditions measured for every combination with two different yeast cell numbers are respectively included in one bar plot. In Figure 4.9A, the frequency of displayers of all the different binder-displaying yeast cells after exposure to SUMO-E5 coated beads, unloaded beads as well as the non-selected condition are shown. Equally, Figure 4.9B displays the data from the bead selection with SUMO-R1 on the beads, Figure 4.9C represents the data from the bead selections with SUMO-L and Figure 4.9D summarizes the generated data of the bead selection experiments with MBP-S. As mentioned above, every bead selection experiment was performed twice and respectively showed the same trend. In the Figure 4.9 below average values of both measurements are demonstrated.

Given that E2 and E4 are assumed to be non-sticky binders, we did not expect to see a reduction in the frequency of displayers after exposure of the YSD sample to unloaded beads nor the sticky protein coated beads. Accordingly, the frequency of displayers (approximately 10 %) displayed on the y-axis was the same for all the conditions. This was true for all the different proteins immobilized on the beads.

For the sticky binders bococizumab, E5 and R1 displayed by yeast cells an exposure to the beads followed by a negative selection should lead to a reduction in the fraction of displayers according to our hypothesis. Unfortunately, such a trend could not be observed for those binders. Apart from little fluctuations the frequency of displayers detected in our samples was not altered in the selected conditions compared to the non-selected ones. Thus, no effect caused by the bead selection assay was noticed. Again, this was true for all the sticky proteins that we coupled with the beads.


10^6 yeast cells
10^7 yeast cells
A
Frequency of displayer (%)
Bococizumab-displaying yeasts
E2-displaying yeasts
E4-displaying yeasts
E5-displaying yeasts
R1-displaying yeasts
SUMO-E5-beads
unloaded beads
Non-selected
B
SUMO-R1-beads

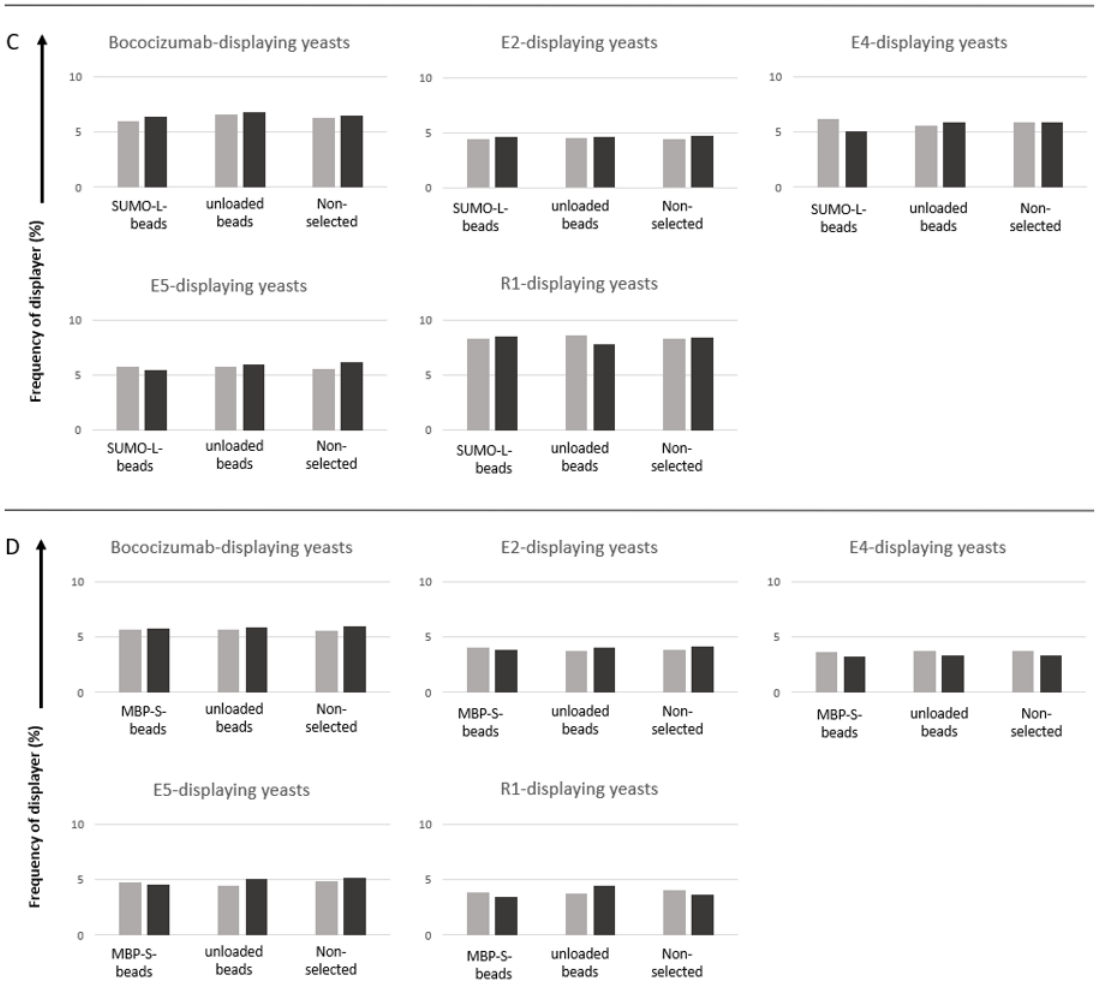


Figure 4.9: Evaluation of the bead selection assay performance with (A) SUMO-E5 coated beads, (B) SUMO-R1 coated beads, (C) SUMO-L coated beads and (D) MBP-S coated beads. Binder-displaying yeast cells were measured in their frequency of displayers (%) after the performance of a bead selection with protein coated beads as well as unloaded beads. Additionally, the frequency of displayers (%) in non-selected samples was measured for comparison. All the conditions were measured twice. Above, the average data of those two analyses is shown.

Chapter 5

Discussion

In this project, we sought to establish a novel bead selection based assay that could be useful in the reduction of non-specific interactions of engineered proteins. We hypothesised that proteins with sticky properties coupled to the surface of magnetic dynabeads should be able to capture non-specifically interacting clones when incubated with a protein library. Subsequently, the non-specifically binding protein could be separated from the rest of the library together with the beads using a magnet. Once established, this so called negative selection assay could be implemented in the developmental process of therapeutic binders and thereby potentially enhance their probability of clinical success.

The sticky proteins coupled to the beads and subsequently tested in the bead selection assay consisted of two Sso7d based binders [18, 19], a fibronectin based binder [16] and a scFv based on the antibody sirukumab [23]. As all of those binders were expected to show non-specific binding and to be prone to aggregation, they were expressed in fusion with a stabilizing protein, namely MBP for the sirukumab based scFv and SUMO for all the other proteins.

To evaluate if the bead selection assay was able to deplete non-specific binders we tested it with a known set of proteins. The set included some proteins that showed sticky behaviour while for others more favorable characteristics were observed before, i.e. no stickyness and little or no propensity for aggregation. Yeast cells displaying those proteins were incubated with the protein coated dynabeads. In the resulting flow cytometric data we could not observe any consistent decreases in the fraction of displayer. This was true for the sticky yeast display samples as well as the non-sticky samples and suggests that the bead selection based assay is not suitable for the reduction of non-specifically binding content of a library.

The success of this method highly depends on the characteristics of the proteins and their quality after soluble expression. Therefore, to ensure that we can replicate previously determined properties like aggregation behaviour and stickiness, we performed some protein quality evaluations. SEC-MALS measurements of the sticky fusion proteins demonstrated that the binders show non-specific binding and aggregation.

The FN3 based lysozyme binder (SUMO-L) was previously shown to be partially

in a significant oligomeric state by Hackel in a study published in 2008. The binder was measured with SEC-HPLC and found to be 80% monomeric [16]. Similarly, our SEC-MALS measurements of SUMO-L suggest a part of the protein to be in a monomeric state, while an additional earlier eluted peak indicates the tendency of the protein to self-interact. Nevertheless, aggregational behaviour can also be caused by intermolecular disulfide bonding of free cystein residues. Consequently, further investigations should evaluate the contribution of the two cystein residues present in the amino acid sequence of the binder.

The Sso7d based EGFR binder SUMO-E5 eluted as single peak in our SEC-MALS measurements. This indicates that the protein was successfully purified and monomeric. Equally to previous projects the monomeric peak exhibited an asymmetric form. [18]. Such peak tailing is associated with non-specific binding behaviour. Presumably, sticky proteins interact to a certain extent with the inert chromatography column leading to the right-tilted shape of the measured peak. In the case of SUMO-E5, we expect the sticky characteristics to come primarily from its aromatic amino acid content.

The generated data of our scFv based binder MBP-S lead to the suggestion that the protein is highly susceptible to self aggregation. A certain tendency to aggregation was expected as the IgG1 isotype antibody sirukumab had also shown poor performance in the Adimap assays that evaluate specificity and aggregation behaviour. In their measurements this was particularly evident from the Affinity-Capture Self-Interaction Nanoparticle Spectroscopy (AC-SINS) data, where self-association behaviour of antibodies conjugated with gold nanoparticles can be measured at low protein concentrations [23, 43]. Accordingly, we detected a high amount of aggregates by SEC-MALS. Moreover, the HIS-column purification did not yield MBP-S in a pure form but rather a mixture of different proteins, which could indicate co-purification of contaminants due to non-specific interaction with the scFv. Therefore, the expression and purification process of MBP-S should definitely be optimized for future applications.

All of our proteins of interest were successfully biotinylated prior to the coupling with the streptavidin beads. This was evident from the SEC-HPLC measurements performed with each protein in combination with a streptavidin fluorescent conjugate. In the mixed sample the fluorescently labeled streptavidin associated to the biotinylated protein which resulted in earlier elution from the column. Also, the measured absorbance at 647 nm clearly indicates that streptavidin could associate to all of our samples due to the biotinylation.

According to the literature the main contributors to non-specific interactions between proteins are hydrophobicity and positive charge. Therefore, we tried to couple proteins showing such characteristics to the surface of the streptavidin dynabeads. The EGFR binder SUMO-E5 exhibits a high aromatic amino acid content which gives it its hydrophobic properties. More precicely, its nine-residue binding site comprises one tyrosine and phenylalanine residue, respectively, as well as two tryptophans. Especially tryptophan was repeatedly associated with aggregation and is typically found in antibody hot spots as part of hydrophobic patches [36, 44, 45]. Nevertheless, one should consider that a

recent study of Kelly et al. where antibody libraries were screened for non-specific binding characteristics showed tryptophan to be predominantly enriched as part of motifs in the non-specifically binding population rather than as single amino acid [26]. In comparison with SUMO-E5, our second Sso7d based binder SUMO-R1 shows an even higher number of hydrophobic amino acids in its binding interface with four of nine residues being tryptophans. In addition, its binding site also includes two positively charged lysines. As mentioned before, positive charges are associated with non-specific binding and more rapid *in vivo* clearance [27, 24, 32]. Moreover, Kelly et al. found lysine to be enriched in a library that was screened for non-specificity, to a lesser extent than arginine though [26].

A particular feature of the Sso7d scaffold is its rigid binding interface located on the surface of a beta sheet. This is in contrast to natural binding scaffolds like antibodies, which rather interact via their CDR loops. Accordingly, the amino acid composition of Sso7d binders differs considerably from that of typical antibody paratopes and resembles that of energetic hot spots of antibody binding sites [18]. As we could not exclude that this rigidity of the paratope would impede the functionality of the Sso7d based binders in the bead selection assay, we also investigated the performance of SUMO-L and MBP-S on the beads. Both, the human FN3 domain as well as scFvs recognize their targets via exposed loops. As mentioned before, this distinct mode of binding results also in a different distribution of amino acids in the paratope and provides a certain degree of flexibility which might be favorable for capturing non-specific binders [18]. However, we could not observe the desired effects in the bead selection assay for neither of the interaction modes. Also, the slightly increased size of the scFv compared to the other binding scaffolds did not seem to favor the selection of non-specific binders.

It could be anticipated that the fusion to stabilizing proteins such as SUMO or MBP alters the characteristics of the sticky proteins. However, according to a recent publication the affinities for binders either alone or fused to SUMO is very comparable [18]. In the study, different Sso7d binders were measured in their binding affinity to mouse serum albumin. While the Kd of the Sso7d binders only was evaluated using YSD, the affinities of the same binders expressed in fusion with SUMO was determined via biolayer interferometry. Both the assessments show comparable results and little fluctuations may be attributed to the distinct modes of measurement [18]. Equally, the performance of scFvs expressed in fusion with MBP as well as their functionality without fusion partner was investigated in a publication dating back to 2001. According to their data, both proteins showed comparable performance [38]. Undesirable effects caused by the fusion partners SUMO & MBP could also be excluded by associating only the sticky alternative scaffold binders to the beads surface. However, co-expressed fusion partners like SUMO or MBP could also lead to a better distribution of biotin on the protein's surface and thereby function as a linker between the binders and the beads.

As mentioned above, evaluations of specificity and other 'drug-like' properties typically demand large amounts of pure protein and are therefore limited to a few lead candidates. Thus, considerable efforts have been made over the years to develop high throughput

tools with lower material requirements. In 2012, Hötzel described an ELISA based approach which measures antibody binding to baculovirus particles BVP [46]. BVPs are stable nanoparticles that mimic infected cell surfaces and consist of a complex mixture of natural compounds such as phospholipids and extracellular matrix constituents. A higher score in the *in vitro* assay was shown to correlate with faster clearance rates in cynomolgus monkeys. In terms of *in vivo* clearance predictions, the development of the BVP assay was a step forward compared to traditional biophysical property evaluations like cross-interaction chromatography (CIC) as it could identify more off-target effects in a more high throughput manner [46].

More recently, the company Adimab described a FACS based approach that can be used for the identification of non-specific binding proteins. In addition, this so called polyspecificity reagent (PSR) assay can be applied to counter select for non-specificity. Thereby, PSR which consists of solublilized mammalian cell membrane preparations is used to screen libraries for sticky variants in a high-throughput manner [21]. Equally to our bead selection assay, the investigated proteins are typically presented by yeasts and the assay can be implemented in the binder discovery already during early stages. In performance evaluations of Adimab, the PSR assay correlates well with CIC measurements and the BVP assay developed by Hötzel et al. Those two methods are known as surrogates for solubility prediction of antibodies and *in vivo* clearance, respectively. However, while part of the antibodies that were measured with the three distinct assays by Adimab were CIC positive, BVP positive and could also be identified as sticky by the PSR assay with a sensitivity of 98 %, outliers between the methods exist and performing the multiple methods may still be beneficial if one is not limited by the throughput [21].

As previously mentioned, our bead selection based negative selection assay did not show the expected performance in the depletion of non-specifically interacting proteins. Still, some fairly comparable aspects of the assay and the PSR approach might suggest its functionality in the appropriate set up. Both of the methods use potentially sticky proteins to catch non-specific binders in diversified YSD libraries. In the case of the PSR those are mammalian proteins extracted from the soluble membrane fraction. While this approach brings the advantage of covering various modes of interaction by using a mix of proteins, this leads to considerable lot-to-lot differences. In contrast to those ill-defined compositions, the proteins that we applied for the capturing of sticky binders were previously characterized and deliberately chosen which would make the assay's outcomes better reproducible. Intriguingly, the Wittrup lab explored the usage of well-characterized chaperones as possible replacements for the PSR mix in a later publication and could predict non-specificity with slightly less accuracy [47]. Notably, it would be highly interesting to test such chaperones as well as the PSR itself on the beads.

Moreover, a considerable disadvantage of the FACS based selection of non-specific binders applied in the PSR assay is the necessity of multiple washing steps prior to detection. Given that non-specific associations are described as a very weak type of bonding re-dissociation could eventually happen quite easily while washing. For the bead selection assay, no washing steps are required between incubation and separation of the

beads. Further, the selection is performed immediately after incubation which potentially increases the probability of capturing also interactions with a high dissociation rate (*koff*).

In theory, the most important advantage of the bead selection based approach compared to a FACS based screening are the inherent characteristics of the method. As mentioned in the first chapter, association of an interaction partner on the beads surface in combination with the multivalency that is naturally given by displaying yeast cells (typically between 10,000 and 100,000 POI per cell [9]) causes enormous avidity effects which correspondingly enhances the half-life of bonding. Hypothetically, this should be the ideal foundation for the capturing of low affinity off-target binding and interactions with short half-lives. Additionally, the immobilization intrinsically causes a local concentration enhancement once a yeast cell finds itself in close proximity to a protein coated bead. Highly avid magnetic bead capture offers extremely high throughput and is a very cost effective method [11].

At the bottom, it remains partially occluded why the bead selection based approach introduced here did not lead to the expected reduction of non-specific binders in YSD samples. In our investigations we tested distinct types of alternative scaffolds as well as different modes of interaction. Our proteins exhibited all the typical characteristics associated with non-specific interaction in the literature and next to their capacity to cross interact with other sticky proteins also self interactions were evaluated. Consequently, this leads to the suggestion that a certain variability of molecules might be necessary for capturing sticky protein variants in a diversified protein library as it was successfully shown with the PSR assay. Obviously, our data set was fairly small to assess the functionality of the bead selection based approach and should definitely be enlarged for further investigations. Possibly, proteins with structural and functional similarities to molecular chaperones could be capable to catch non-specifically interacting protein variants. Finally, a potential next step could be to investigate the the suitability of the bead based assay itself by immobilizing the PSR mix or its chaperone proportion identified as functional agent on the surface of the dynabeads.

Overall, we were not able to establish a novel method for avoiding binders with reduced specificity in protein engineering libraries. Nevertheless, the bead selection based approach shows some theoretical advantages compared to FACS based selections, especially for weak interactions like off-target binding. Therefore, future findings might bring it back to the spotlight in the ongoing challenge of predicting therapeutic binders' success before going into the investigations on a clinical level.

Bibliography

[1] Maarten P Schreuder, Stephan Brekelmans, Herman Van Den Ende, and Frans M Klis. Targeting of a heterologous protein to the cell wall of saccharomyces cerevisiae. *Yeast*, 9(4):399–409, 1993.

[2] A Kondo and M Ueda. Yeast cell-surface display—applications of molecular display. *Applied microbiology and biotechnology*, 64(1):28–40, 2004.

[3] Eric T Boder and K Dane Wittrup. Yeast surface display for screening combinatorial polypeptide libraries. *Nature biotechnology*, 15(6):553–557, 1997.

[4] Eric V Shusta, Lauren R Pepper, Yong K Cho, and Eric T Boder. A decade of yeast surface display technology: where are we now? *Combinatorial chemistry & high throughput screening*, 11(2):127–134, 2008.

[5] John McCafferty, Andrew D Griffiths, Greg Winter, and David J Chiswell. Phage antibodies: filamentous phage displaying antibody variable domains. *nature*, 348(6301):552–554, 1990.

[6] Mingyue He and Michael J Taussig. Antibody-ribosome-mrna (arm) complexes as efficient selection particles for in vitro display and evolution of antibody combining sites. *Nucleic acids research*, 25(24):5132–5134, 1997.

[7] Richard W Roberts and Jack W Szostak. Rna-peptide fusions for the in vitro selection of peptides and proteins. *Proceedings of the National Academy of Sciences*, 94(23):12297–12302, 1997.

[8] Eric T Boder and K Dane Wittrup. [25] yeast surface display for directed evolution of protein expression, affinity, and stability. *Methods in enzymology*, 328:430–444, 2000.

[9] Ginger Chao, Wai L Lau, Benjamin J Hackel, Stephen L Sazinsky, Shaun M Lippow, and K Dane Wittrup. Isolating and engineering human antibodies using yeast surface display. *Nature protocols*, 1(2):755, 2006.

[10] Gerald M Cherf and Jennifer R Cochran. Applications of yeast surface display for protein engineering. *yeast surface display*, pages 155–175, 2015.

[11] Margaret Ackerman, David Levary, Gabriel Tobon, Benjamin Hackel, Kelly Davis Orcutt, and K Dane Wittrup. Highly avid magnetic bead capture: an efficient selection method for de novo protein engineering utilizing yeast surface display. *Biotechnology progress*, 25(3):774–783, 2009.

[12] Yik A Yeung and K Dane Wittrup. Quantitative screening of yeast surface-displayed polypeptide libraries by magnetic bead capture. *Biotechnology progress*, 18(2):212–220, 2002.

[13] Jennifer J Van Antwerp and K Dane Wittrup. Fine affinity discrimination by yeast surface display and flow cytometry. *Biotechnology progress*, 16(1):31–37, 2000.

[14] Michele C Kieke, Bryan K Cho, Eric T Boder, David M Kranz, and K Dane Wittrup. Isolation of anti-t cell receptor scfv mutants by yeast surface display. *Protein engineering*, 10(11):1303–1310, 1997.

[15] Daša Lipovšek, Shaun M Lippow, Benjamin J Hackel, Melissa W Gregson, Paul Cheng, Atul Kapila, and K Dane Wittrup. 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(4):1024–1041, 2007.

[16] Benjamin J Hackel, Atul Kapila, and K Dane Wittrup. Picomolar affinity fibronectin domains engineered utilizing loop length diversity, recursive mutagenesis, and loop shuffling. *Journal of molecular biology*, 381(5):1238–1252, 2008.

[17] Nimish Gera, Mahmud Hussain, Robert C Wright, and Balaji M Rao. Highly stable binding proteins derived from the hyperthermophilic sso7d scaffold. *Journal of molecular biology*, 409(4):601–616, 2011.

[18] Michael W. Traxlmayr, Jonathan D. Kiefer, Raja R. Srinivas, Elisabeth Lobner, Alison W. Tisdale, Naveen K. Mehta, Nicole J. Yang, Bruce Tidor, and K. Dane Wittrup. Strong enrichment of aromatic residues in binding sites from a charge-neutralized hyperthermostable sso7d scaffold library*. *Journal of Biological Chemistry*, 291(43):22496–22508, 2016.

[19] Monique J Kauke, Michael W Traxlmayr, Jillian A Parker, Jonathan D Kiefer, Ryan Knihtila, John McGee, Greg Verdine, Carla Mattos, and K Dane Wittrup. An engineered protein antagonist of k-ras/b-raf interaction. *Scientific reports*, 7(1):1–9, 2017.

[20] Ruei-Min Lu, Yu-Chyi Hwang, I-Ju Liu, Chi-Chiu Lee, Han-Zen Tsai, Hsin-Jung Li, and Han-Chung Wu. Development of therapeutic antibodies for the treatment of diseases. *Journal of biomedical science*, 27(1):1–30, 2020.

[21] Yingda Xu, William Roach, Tingwan Sun, Tushar Jain, Bianka Prinz, Ta-Yi Yu, Joshua Torrey, Jerry Thomas, Piotr Bobrowicz, Maximiliano Vásquez, et al. Addressing polyspecificity of antibodies selected from an in vitro yeast presentation system:

a facs-based, high-throughput selection and analytical tool. *Protein Engineering, Design & Selection*, 26(10):663–670, 2013.

[22] Steffen Hartmann and Hans P Kocher. Best practices in assessment of developability of biopharmaceutical candidates. In *Developability of Biotherapeutics: Computational Approaches*, pages 155–174. CRC Press, Boca Raton, FL, 2015.

[23] Tushar Jain, Tingwan Sun, Stéphanie Durand, Amy Hall, Nga Rewa Houston, Juergen H Nett, Beth Sharkey, Beata Bobrowicz, Isabelle Caffry, Yao Yu, et al. Biophysical properties of the clinical-stage antibody landscape. *Proceedings of the National Academy of Sciences*, 114(5):944–949, 2017.

[24] Charles G Starr and Peter M Tessier. Selecting and engineering monoclonal antibodies with drug-like specificity. *Current opinion in biotechnology*, 60:119–127, 2019.

[25] Lindsay B Avery, Jason Wade, Mengmeng Wang, Amy Tam, Amy King, Nicole Piche-Nicholas, Mania S Kavosi, Steve Penn, David Cirelli, Jeffrey C Kurz, et al. Establishing in vitro in vivo correlations to screen monoclonal antibodies for physicochemical properties related to favorable human pharmacokinetics. In *MAbs*, volume 10, pages 244–255. Taylor & Francis, 2018.

[26] Ryan L Kelly, Doris Le, Jessie Zhao, and K Dane Wittrup. Reduction of nonspecificity motifs in synthetic antibody libraries. *Journal of molecular biology*, 430(1):119–130, 2018.

[27] T Igawa, H Tsunoda, T Tachibana, A Maeda, F Mimoto, C Moriyama, M Nanami, Y Sekimori, Y Nabuchi, Y Aso, et al. Reduced elimination of igg antibodies by engineering the variable region. *Protein Engineering, Design & Selection*, 23(5):385–392, 2010.

[28] Deidra Bethea, Sheng-Jiun Wu, Jinquan Luo, Linus Hyun, Eilyn R Lacy, Alexey Teplyakov, Steven A Jacobs, Karyn T O'Neil, Gary L Gilliland, and Yiqing Feng. Mechanisms of self-association of a human monoclonal antibody cnto607. *Protein Engineering, Design & Selection*, 25(10):531–538, 2012.

[29] Amita Datta-Mannan, Arunkumar Thangaraju, Donmienne Leung, Ying Tang, Derrick R Witcher, Jirong Lu, and Victor J Wroblewski. Balancing charge in the complementarity-determining regions of humanized mabs without affecting pi reduces non-specific binding and improves the pharmacokinetics. In *MAbs*, volume 7, pages 483–493. Taylor & Francis, 2015.

[30] Sheng-Jiun Wu, Jinquan Luo, Karyn T O'Neil, James Kang, Eilyn R Lacy, Gabriela Canziani, Audrey Baker, Maggie Huang, Qing Mike Tang, T Shantha Raju, et al. Structure-based engineering of a monoclonal antibody for improved solubility. *Protein Engineering, Design & Selection*, 23(8):643–651, 2010.

[31] Hedda Wardemann, Sergey Yurasov, Anne Schaefer, James W Young, Eric Meffre, and Michel C Nussenzweig. Predominant autoantibody production by early human b cell precursors. *Science*, 301(5638):1374–1377, 2003.

[32] C Andrew Boswell, Devin B Tesar, Kiran Mukhyala, Frank-Peter Theil, Paul J Fielder, and Leslie A Khawli. Effects of charge on antibody tissue distribution and pharmacokinetics. *Bioconjugate chemistry*, 21(12):2153–2163, 2010.

[33] Sara Birtalan, Robert D Fisher, and Sachdev S Sidhu. The functional capacity of the natural amino acids for molecular recognition. *Molecular bioSystems*, 6(7):1186–1194, 2010.

[34] Pilarin Nichols, Li Li, Sandeep Kumar, Patrick M Buck, Satish K Singh, Sumit Goswami, Bryan Balthazor, Tami R Conley, David Sek, and Martin J Allen. Rational design of viscosity reducing mutants of a monoclonal antibody: hydrophobic versus electrostatic inter-molecular interactions. In *MAbs*, volume 7, pages 212–230. Taylor & Francis, 2015.

[35] Laurence Lins, Annick Thomas, and Robert Brasseur. Analysis of accessible surface of residues in proteins. *Protein science*, 12(7):1406–1417, 2003.

[36] Gautier Robin, Yoshiteru Sato, Dominique Desplancq, Natacha Rochel, Etienne Weiss, and Pierre Martineau. Restricted diversity of antigen binding residues of antibodies revealed by computational alanine scanning of 227 antibody–antigen complexes. *Journal of Molecular Biology*, 426(22):3729–3743, 2014.

[37] Donald M Crothers and Henry Metzger. The influence of polyvalency on the binding properties of antibodies. *Immunochemistry*, 9(3):341–357, 1972.

[38] Horacio Bach, Yariv Mazor, Shelly Shaky, Atar Shoham-Lev, Yevgeny Berdichevsky, David L Gutnick, and Itai Benhar. Escherichia coli maltose-binding protein as a molecular chaperone for recombinant intracellular cytoplasmic single-chain antibodies. *Journal of molecular biology*, 312(1):79–93, 2001.

[39] Nikolai N Sluchanko, Kristina V Tugaeva, Yaroslav V Faletrov, and Dmitrii I Levitsky. High-yield soluble expression, purification and characterization of human steroidogenic acute regulatory protein (star) fused to a cleavable maltose-binding protein (mbp). *Protein expression and purification*, 119:27–35, 2016.

[40] Alison L Main, Timothy S Harvey, Martin Baron, Jonathan Boyd, and Iain D Campbell. The three-dimensional structure of the tenth type iii module of fibronectin: an insight into rgd-mediated interactions. *Cell*, 71(4):671–678, 1992.

[41] Siyi Hu, Zhiqiang Zhu, Liangwei Li, Liang Chang, Weifang Li, Liansheng Cheng, Maikun Teng, and Jing Liu. Epitope mapping and structural analysis of an anti-erbb2 antibody a21: molecular basis for tumor inhibitory mechanism. *Proteins: Structure, Function, and Bioinformatics*, 70(3):938–949, 2008.

[42] Schrödinger, LLC. The PyMOL molecular graphics system, version 1.8. November 2015.

[43] Shantanu V Sule, Muppalla Sukumar, William F Weiss IV, Anna Marie Marcelino-Cruz, Tyler Sample, and Peter M Tessier. High-throughput analysis of concentration-dependent antibody self-association. *Biophysical journal*, 101(7):1749–1757, 2011.

[44] Timothy M Lauer, Neeraj J Agrawal, Naresh Chennamsetty, Kamal Egodage, Bernhard Helk, and Bernhardt L Trout. Developability index: a rapid in silico tool for the screening of antibody aggregation propensity. *Journal of pharmaceutical sciences*, 101(1):102–115, 2012.

[45] Neeraj J Agrawal, Sandeep Kumar, Xiaoling Wang, Bernhard Helk, Satish K Singh, and Bernhardt L Trout. Aggregation in protein-based biotherapeutics: computational studies and tools to identify aggregation-prone regions. *Journal of pharmaceutical sciences*, 100(12):5081–5095, 2011.

[46] Isidro Hötzel, Frank-Peter Theil, Lisa J Bernstein, Saileta Prabhu, Rong Deng, Leah Quintana, Jeff Lutman, Renuka Sibia, Pamela Chan, Daniela Bumbaca, et al. A strategy for risk mitigation of antibodies with fast clearance. In *MAbs*, volume 4, pages 753–760. Taylor & Francis, 2012.

[47] Ryan L Kelly, James C Geoghegan, Jared Feldman, Tushar Jain, Monique Kauke, Doris Le, Jessie Zhao, and K Dane Wittrup. Chaperone proteins as single component reagents to assess antibody nonspecificity. In *MAbs*, volume 9, pages 1036–1040. Taylor & Francis, 2017.