New Protein Engineering Approaches for Potentiating and Studying Antibody-Based EGFR Antagonism

By

Alison W. Tisdale

B.S. Chemical Engineering
University of Delaware, 2005

M.S. Chemical Engineering
University of Minnesota, 2009

Submitted to the Department of Biological Engineering in partial fulfillment of the requirements for the degree of

Doctor of Philosophy in Biological Engineering

at the

Massachusetts Institute of Technology
~~January 2019~~
June 2019

Signature redacted

Signature of Author: ....
Alison W. Tisdale
Department of Biological Engineering
Jan. 23, 2019

Signature redacted

Certified by: ....
K. Dane Wittrup
C.P. Dubbs Professor of Chemical Engineering and Biological Engineering
Thesis Advisor

Signature redacted

Certified by: ....
Forest White
Professor of Biological Engineering
Graduate Academic Program Chair

Thesis Committee Members

Doug Lauffenburger, Ph.D. (Chair)
Professor of Biological Engineering, Chemical Engineering, and Biology
Massachusetts Institute of Technology

K. Christopher Garcia, Ph.D.
Professor of Molecular and Cellular Physiology, and Structural Biology
Stanford University School of Medicine

New Protein Engineering Approaches for Potentiating and Studying Antibody-Based EGFR Antagonism

By

Alison W. Tisdale

Submitted to the Department of Biological Engineering on January 23, 2019 in partial fulfillment of the requirements for the degree of Doctor of Philosophy in Biological Engineering

Abstract

A variety of cancers are marked by the over-expression and over-activity of the EGF receptor (EGFR), rendering this protein an attractive therapeutic target. Anti-EGFR therapeutics are a mainstay of clinical practice for the treatment of colorectal, lung and head and neck cancers but efficacy is limited and response rates low. Opportunities for improving EGFR antagonism include higher potency inhibition of ligand binding, inducing receptor downregulation, or creating synergistic therapeutic combinations.

The Wittrup lab has previously made significant advances in EGFR antagonism by demonstrating the therapeutic potential of inducing receptor downregulation through multi-epitopic targeting. The lab has also pioneered the use of a novel protein scaffold, called Sso7d, for yeast surface display-based libraries and selections. In the first part of this work I show that a combination of traditional yeast display techniques with simple but novel in silico approaches can be applied to derive a panel of Sso7d binders against EGFR with diverse paratopes. I demonstrate the superior EGFR inhibition of antibody-Sso7d fusions in vitro, and discuss the lessons learned from applying these proteins in vivo.

In the second part of this work I use a structure-guided yeast display approach to create a novel research tool, a minimally modified verstion of cetuximab called "mCetux", which essentially enables in vivo experiments of cetuximab. I apply this antibody tool in vitro and in vivo in a new and highly relevant model system for colorectal cancer and subsequently discuss future opportunities for its use.

Thesis Advisor: K. Dane Wittrup
C.P. Dubbs Professor of Chemical Engineering and Biological Engineering

Acknowledgements

I would like to thank my advisor, Dane Wittrup for his guidance throughout these years. My academic and professional development has been influenced by Dane at many points in time, dating back to well before I was a student at MIT, and at every encounter I have found his actions to be in support of my goals and interests. I would also like to thank my thesis committee members, Doug Lauffenburger and Chris Garcia, for their insight as my project has evolved. Their wisdom, expertise and feedback have helped shape my many experiments into a cohesive body of work of which I am proud.

I would also like to thank all of the Wittrup lab members for making my time at MIT and in the Wittrup lab a great experience. You are a friendly, hard-working and incredibly talented group of individuals and sharing an office and a lab with you has been a privilege. I would particularly like to thank those both within and outside of the Wittrup lab who have contributed substantially to this work. Jonathan Braverman, with whom I collaborated for all of the mCetux studies in organoids; he designed and made the organoids and ran all of the in vitro growth studies. He was also a wonderful collaborator and a wealth of information; I appreciate all of his effort on our joint project. Adrienne Rothschilds and Noor Momin made invaluable contributions that enabled essentially all of my own in vivo work. They were selfless with their time and resources and patient with my many questions. Thanks also to all current and former lab members I shared the Wittrup lab with over the years, you have all contributed to my project and scientific maturation and I'm very grateful! Special thanks to Monique Kauke for making me laugh, commiserating with me, supporting my research with ideas and experimental help, and for generally being great.

I am thankful for the help of several talented visiting students and MIT UROPs who contributed to my research: Doreen Konning, who was around for the early days of Sso7d engineering; Emily Miaou, an undergraduate summer student who was the first to demonstrate downregulation by antibody-Sso7d fusions; Vincent Wedekind, a high school summer student who helped build the first mCetux yeast libraries; and Jordyn Mann and Alexis Cuellar, MIT undergraduates who were jacks-of-all-trades around the lab – performing in vitro experiments, producing protein and generally doing anything asked. I would also like to acknowledge the support of funding sources, including the NIH/NIGMS Biotechnology Training Grant.

Lastly, I would like to thank my family. From my parents and siblings, I am grateful for their love and support throughout my extended PhD journey. And most of all, I am beyond thankful for Will, Maddie and Bryn. They make every day of my life great and I love them unconditionally and forever.

Table of Contents

Chapter 1 – Introduction and Thesis Overview

EGFR structure and biology

The epidermal growth factor receptor (EGFR/HER1/ErbB1) is a multi-domain, cell-surface receptor that is a member of the ERBB family of receptor tyrosine kinases. Other members of this family include HER2 (ErbB2/Neu), HER3 (ErbB3) and HER4 (ErbB4). This family of receptors regulates signaling pathways that control critical cellular functions such as proliferation, migration and differentiation. Their general mechanism of action is ligand binding mediated by the extracellular domain, dimerization (either homo- or hetero-dimerization[1,2]) and subsequent initiation of a signaling cascade through phosphorylation of their intracellular kinase domains. Though the ErbB family members are structurally and functionally similar, distinct features help differentiate their unique contributions to the complicated signaling network they regulate – for example, there are no known ligands for the HER2 receptor; HER3 is kinase dead and can only be activated by trans-phosphorylation.

EGFR is large (~170kDa) and heavily glycosylated, with eleven canonical N-linked glycosylation sites. Glycosyation of the receptor has been shown to affect ligand binding kinetics, stabilize receptor conformation and orientation within the membrane, and regulate dimerization[3–7]. The receptor has an extracellular domain, a single transmembrane domain, and an intracellular portion that consists of a juxtamembrane region, a tyrosine kinase and a tyrosine-rich C-terminal tail. The extracellular domain (ECD) is divided into four subdomains: I (residues 1-165), II (residues 166-310), III (residues 311-480) and IV (residues 481-620). Crystal structures of the EGFR ECD exist in both monomeric and dimeric forms, and show two distinct receptor conformations (Figure 1.1)[8–10]. In the unliganded state the receptor is autoinhibited or "closed", tethered by intramolecular contacts between domains II and IV. In the presence of bound ligand

the receptor conformation is untethered, or "open" and dimerization contacts in domain II are exposed. The ligand is bound between domains I and III, with domain III contributing most of the binding energy of the interaction.

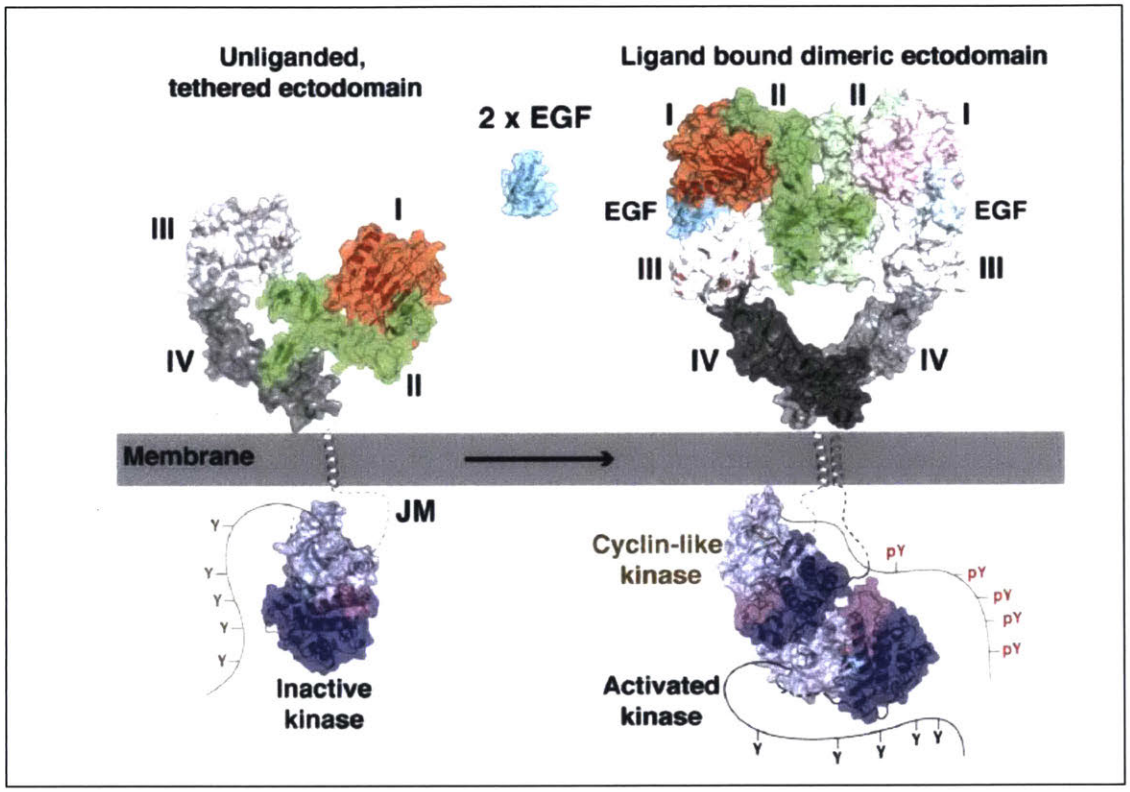


Figure 1.1 – Two conformations of EGFR and mechanism of activation. EGFR exists in a tethered conformation in the absence of ligand. When ligand is bound, the conformation is open and stabilized, allowing dimerization and intracellular phosphorylation. Figure is reproduced from Ferguson, et al[8].

A number of ligands can bind EGFR, including EGF, amphiregulin (AREG), epigen (EPG), and TGFα, all of which bind only EGFR among ErbB family members; EGFR can also bind ligands which cross-react with other ErbB receptors, including heparin-binding EGF-like growth factor (HB-EGF), betacellulin (BTC) and epiregulin (EPR)[2]. Though all of these ligands bind the same site on EGFR they are capable of stimulating divergent biological responses. This functional selectivity is thought to be a consequence of subtle differences in bound conformations that are ligand-specific. Small changes in extracellular conformation mediated by specific ligands

are propagated through the membrane to the kinase domains, where the effect is distinct patterns of phosphorylation and subsequent biological activity[11]. The complex process of ligand binding, dimerization and allosteric regulation of EGFR has been extensively reviewed[12–14].

Ligand binding and dimerization promotes interaction between two receptor intracellular regions and allows autophosphorylation of the kinase domain[12,15,16]. Major phosphorylation sites following EGF-mediated activation include tyrosines 1086, 1148 and 1173[17]. Recruitment and phosphorylation of downstream effector proteins follows, initiating signaling through one of many well-characterized pathways, as shown in Figure 1.2. These include the mitogen-activated protein kinase (MAPK), phosphatidylinositol 3-kinase (PI3K), janus kinase (JAK) / signal transducer and activator of transcription (STAT), and protein kinase C (PKC) pathways[13,18,19]. Flux through this network results in transcription of target genes and regulation of cellular processes such as proliferation, differentiation, survival/apoptosis, and migration/motility.

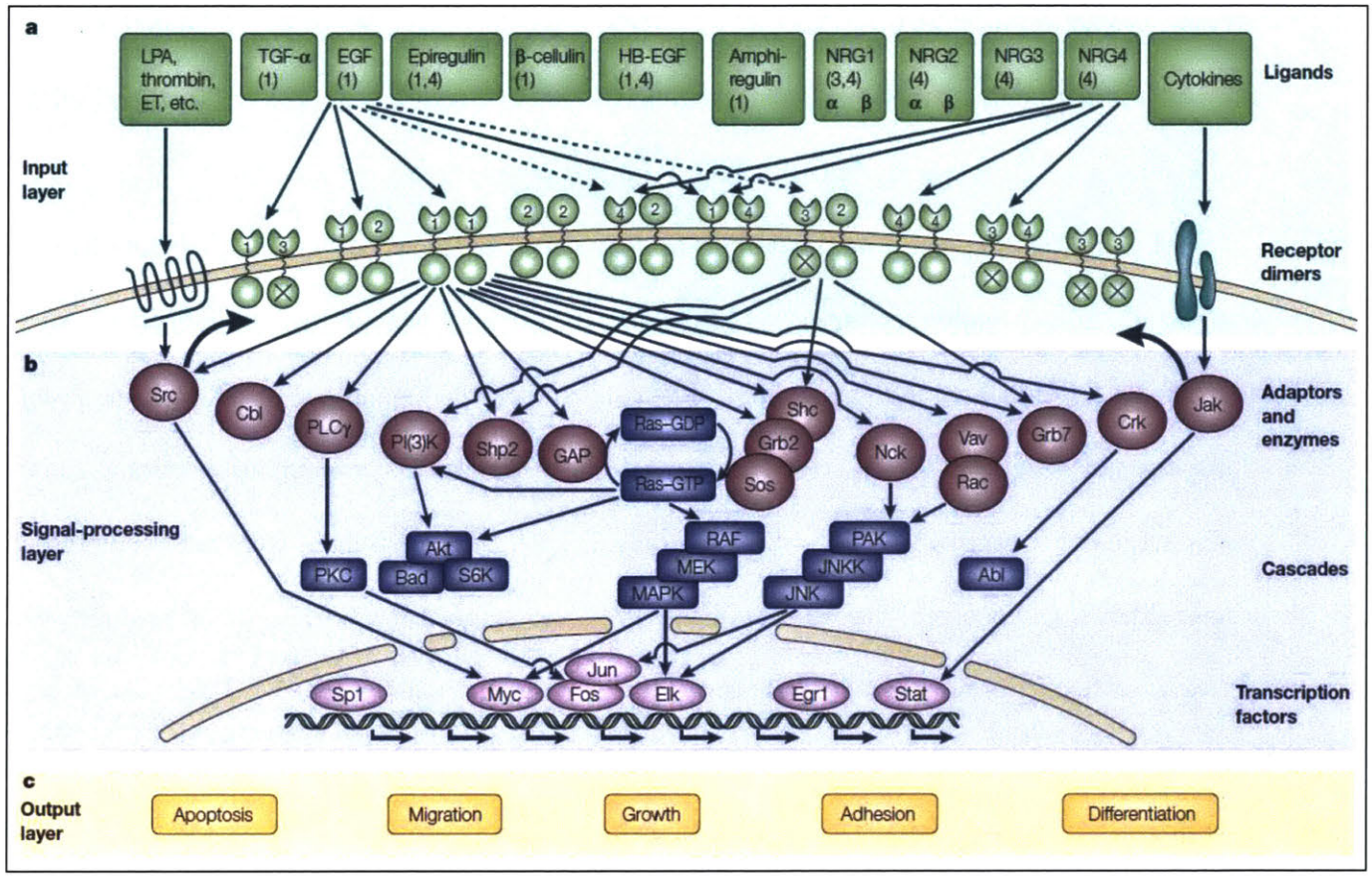


Figure 1.2 – Signaling pathways downstream of ErbB receptors. Reproduced from Yarden and Sliwkowski[18].

Implications of dysregulated EGFR in cancer

EGFR is a validated therapeutic target for the treatment of cancer and dysregulated signaling is known to be oncogenic. Cancerous cells can gain advantage by circumventing or hijacking normal EGFR function in a variety of ways. In some cancers EGFR is overexpressed; normal cells roughly $0.5\text{-}1x10^4$ copies of EGFR, but in tumor cells this can be over 10-fold higher[20]. Overexpression can occur as the result of genetic amplification, as a consequence of a hypoxic tumor microenvironment or as a result of impaired receptor trafficking; in any case, increased density on the cell surface raises the likelihood of dimerization in the absence of ligand binding and ultimately results in ligand-independent signaling[14,16,21,22]. Similarly, overexpression of ligands can result in dysregulated autocrine stimulation of the receptor[23].

Other cancers are driven by activating mutations in EGFR. Mutations in the kinase domain are particularly prevalent in non-small-cell lung cancer (NSCLC)[24,25], and approximately 80% of NSCLC mutations can be accounted for by the exon 19 deletion of the LREA amino acid sequence and the L858R point mutation in exon 21[26,27]. These malignant mutations are transforming, but also confer sensitivity to certain EGFR therapeutics such as small-molecule kinase inhibitors, and result in a state of oncogene addition. Another common oncogenic mutation is the result of a genetic rearrangement that produces a truncation of the EGFR ECD. Known as EGFRvIII (Δ2-7) this receptor lacks amino acids 6-273, resulting in constitutive, ligand-independent, low levels of activation and conferring increased invasiveness and proliferation to affected cells[28–30]. This mutation is commonly found in glioblastoma but has been observed in other tumor types as well[31,32].

Existing anti-EGFR therapeutic strategies

EGFR was the first receptor linked to human cancer, and since then decades of research effort have been invested in developing EGFR inhibitors. Those that have gained FDA approval can be divided into two classes: small-molecule inhibitors which target the intracellular tyrosine kinase domain (TKIs), and monoclonal antibodies (mAbs) which bind to the extracellular region of the receptor [33]. TKIs reversibly or irreversibly compete with ATP for binding to the catalytic domain of the kinase and in doing so block downstream signaling. The first TKIs approved were erlotinib and gefitinib in the early 2000s[34] for the treatment of NSCLC and pancreatic cancers. Development of next generation TKIs has continued since then, with the most exciting advances being the development of drugs that remain efficacious in the context of the acquired mutations that typically make tumors refractory to first-generation TKI treatment. The most recent FDA approval of such a drug was for osimertinib in 2017[35].

An alternative approach to EGFR antagonism is using monoclonal antibodies. Antibodies achieve the same functional effect of blocking downstream signaling, but do so by impairing ligand-receptor binding at the cell surface. Additionally, antibody binding causes receptor internalization and downregulation[36,37], a distinguishing mechanism from TKIs. While both approaches are effective against wild type EGFR, the mutational landscape of a particular cancer may affect the utility of each drug differently. For example, EGFR harboring the kinase-inhibitor resistant T790M mutation remains responsive to cetuximab[38] while the S492R mutation in the EGFR ECD confers resistance to cetuximab, but not to kinase inhibitors or even to panitumumab[39].

The two clinically approved antibodies against EGFR are cetuximab (Erbitux), a human-mouse chimeric IgG1 and panitumumab (Vectibix), a fully human IgG2. In patients, these antibodies are employed either as single-agents or in combination with chemotherapies, for the

treatment of lung, colorectal, and head and neck cancers. Treatment with cetuximab results in statistically significant, if modest, increases in overall survival[40–44]; in the clinical trial that led to FDA approval of cetuximab in colorectal cancer, the overall response rate when used in combination with chemotherapy was only 23%, with median duration of response of 5.7 months. In head and neck cancer the addition of cetuximab to chemotherapy increased overall survival by about 3 months.

The major toxicity associated with anti-EGFR antibodies is dermatological; 80-95% of metastatic colorectal cancer (mCRC) patients treated with cetuximab or panitumumab develop a papulopustular skin rash, typically affecting sun-exposed surfaces of the body[45]. Interestingly, data from clinical trials showed a positive correlation between rash severity and clinical response[41,46,47]. Other noted toxicities include fatigue, hypersensitivity and gastrointestinal sensitivities[41,45].

Mechanism of action of EGFR-targeting antibodies

It has long been appreciated that the EGFR-targeted antibodies cetuximab and panitumumab derive at least part of their anti-tumor activity from direct steric blocking of receptor-ligand binding; this is the purpose for which they were developed. The mechanism underlying this function can be seen in the crystal structure of the cetuximab Fab region in complex with soluble EGFR extracellular domain (ECD), shown in Figure 1.3. This structure shows the antibody binding at domain III of the receptor ECD, at least partially occluding the known ligand binding site[8,48]. Consequently, these antibodies decrease ligand-dependent receptor phosphorylation and suppress signaling through the numerous pathways downstream of EGFR[19]. In vitro assays and in vivo studies have shown decreased cellular proliferation, the induction of apoptosis, cell-cycle arrest, delayed tumor growth, decreased angiogenesis and inhibition of metastasis to be some of the consequences of cetuximab exposure.

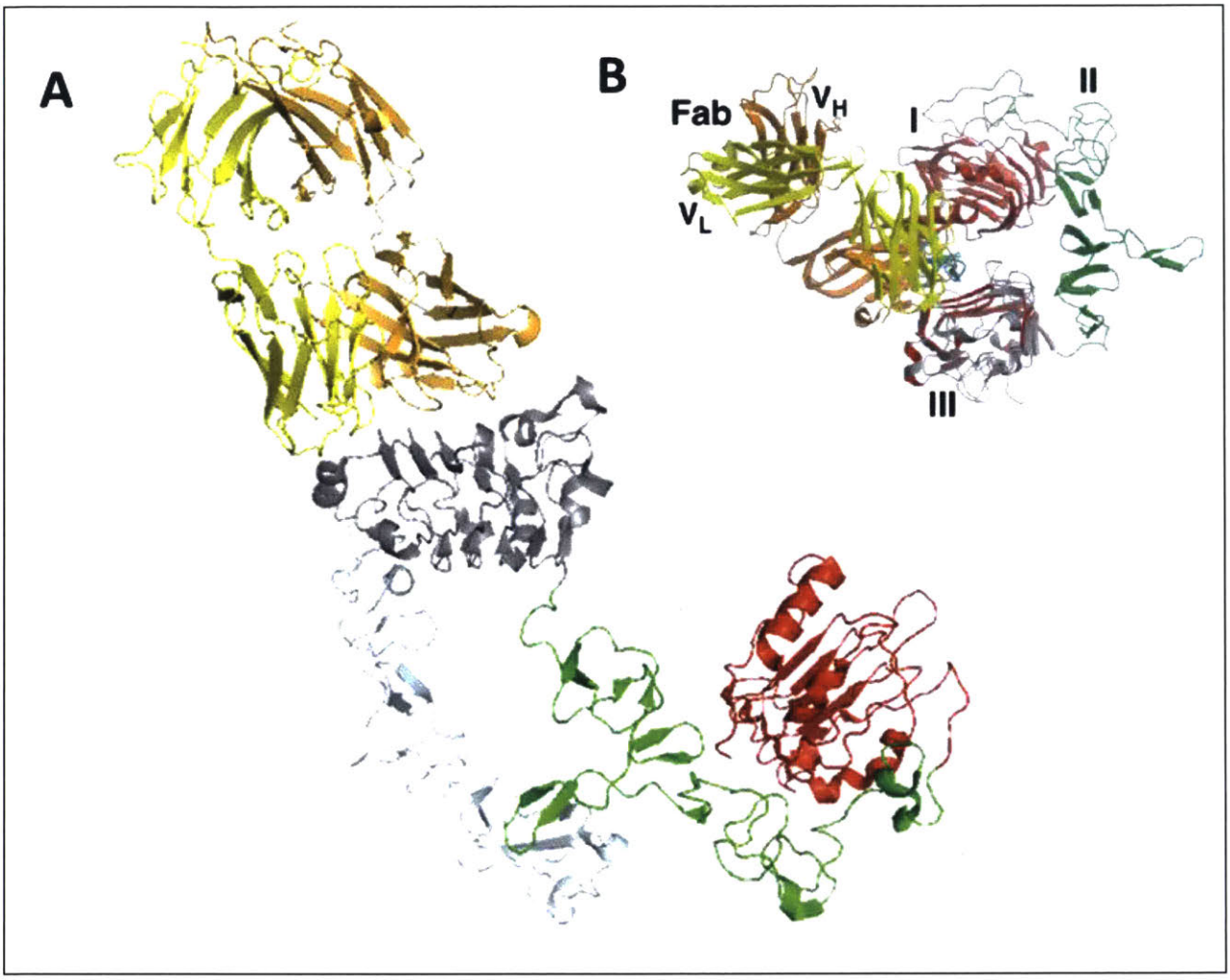


Figure 1.3 – Crystal structure of EGFR in complex with cetuximab Fab showing mechanism of EGFR inhibition (adapted from Li, et al[48], PDB: 1YY9). **A** – Cetuximab Fab binds to domain III of the EGFR ECD (domain I: red; domain II: green; domain III: gray; domain IV: pale cyan). **B** – Domain III and the Fab from A are superimposed on the structure of ligand (EGF)-bound receptor domains I-III. Figure from Li, et al.

Additionally, EGFR antibodies can engage the innate immune system and initiate ADCC through interactions between their Fc regions and FcγRs on immune cells. A similar mechanism of action has been shown for other antibodies in cancer including rituximab (Rituxan), an anti-CD20 monoclonal antibody used for B-cell lymphoma[49], and trastuzumab (Herceptin), an anti-HER2 anibody for breast cancer[50], yet early understanding of cetuximab was that immune-mediated effects were minimal, since growth inhibitory activity of the antibody in xenograft

models was preserved when only a Fab fragment was used[51]. But subsequent in vitro studies with cetuximab demonstrated that the antibody could in fact mediate ADCC against a variety of tumor cell lines[52], and suggested that this mechanism constituted an important component of the overall anti-tumor effect observed in patients. Consistent with this hypothesis is the finding that induction of apoptosis required Fc receptor-positive cells[53], and that FcγR polymorphisms affecting Fc binding correlated with clinical response to cetuximab[54,55]. It has even been suggested that the differences in antibody isotype between cetuximab and panitumumab, which result in distinct profiles of innate immune cell stimulation, can explain the reduced clinical effect of panitumumab[56].

Even more recently, research has continued to elucidate other cetuximab-mediated immune effects, in particular the ways in which cetuximab engages and depends upon components of the adaptive immune system. Multiple studies have demonstrated the importance of CD8+ T cells for cetuximab's anti-tumor effect[57,58], and the ability of cetuximab to trigger immunogenic cell death[59]. Additionally, the inhibition of EGFR signaling provided by cetuximab can influence adaptive immune responses by potentiating the induction of MHC expression on cancer cells[60], reducing immunosuppressive PD-L1 expression[61,62], and reshaping the tumor microenvironment by altering cytokine and chemokine expression[63].

It is worth noting that knowledge about cetuximab and the adaptive immune system was scarce until relatively recently because in vivo studies of cetuximab were limited to human xenograft models in immune-deficient mice. Investigations which probe the function of a fully competent immune system require new tools; while some advances have been made, such as the immune-reconstituted xenograft model used by Yang et al to interrogate both innate and adaptive

immune responses to cetuximab[58], even more physiologically faithful in vivo systems would provide additional benefit. This is the motivation for the work in Chapter 3.

Current limitations and opportunities for improved EGFR targeting

The further development of improved EGFR-based therapeutics is ongoing. Strategies for improving efficacy are primarily focused on achieving more potent blockade of the EGFR signaling network and overcoming acquired resistance. Typical approaches include increasing receptor downregulation (as is discussed more in Chapter 2), multi-specific targeting of ErbB family members, and deploying combinations of existing EGFR inhibitors. More recently, combinations of anti-EGFR antibodies and immunotherapies such as check-point inhibitors have begun to be explored[64]. EGFR targeting is also important for therapy outside the context of signal antagonism. For example, antibody drug conjugates (ADCs)[65] and CAR-T cells[66] and bispecific T-cell engagers (BiTEs)[67] can all take advantage of EGFR overexpression in tumors.

Thesis overview

This thesis details two distinct but related efforts in which the Wittrup lab's expertise in protein engineering was leveraged to tackle the inadequacies of current EGFR-based approaches for cancer treatment. In Chapter 2 we explore the hypothesis that superior signaling blockade and tumor control could be achieved by a single agent that targets multiple EGFR epitopes simultaneously. To this end, novel fusion proteins were built by using the cetuximab antibody as a "backbone" and fusing to it EGFR-binding moieties derived from a novel scaffold library based on a small, hyperstable protein called Sso7d.

The discovery and engineering of the EGFR-targeting Sso7ds was achieved by yeast surface display, and employed some standard library screening strategies as well as a novel in silico approach for identifying low-frequency clone families from enriched populations. Extensive affinity maturation was carried out to optimize the EGFR binders, and this exercise revealed recurring patterns that were translated into specific recommendations for best practices in future Sso7d engineering projects.

Following affinity maturation the engineered Sso7d clones were fused to cetuximab to create multi-epitopic targeting constructs; the activity of these fusions was explored extensively in vitro, and characterized in a mouse model of glioblastoma.

Chapter 3 presents a project in which the aim was not to create a novel therapeutic moiety, but rather to develop a tool which could be used pre-clinically to support the further elucidation of anti-EGFR antibody biology. We also envisioned a potential application of this reagent to be the development of rationally designed combination therapies; anti-EGFR antibodies combined with immunotherapies are a particularly attractive field of study.

To this end, the yeast surface display platform was utilized to engineer cetuximab for murine EGFR cross-reactivity. By employing a structure-guided, saturation mutagenesis strategy it was possible to develop a minimally mutated antibody that bound both mouse and human EGFR with comparable affinity. This activity of this new antibody was tested against cetuximab in a panel of in vitro assays to validate its comparability. Finally the antibody was applied to a particularly relevant mouse model system of colorectal cancer.

Chapter 2 – Multi-Epitopic Targeting of EGFR by Sso7d-Antibody Fusions

Introduction

EGFR is a validated therapeutic target for the treatment of cancer, owing to its over-expression and over-activity in a variety of tumor types[21,22,31,32,68]. Clinically approved drugs designed to inhibit EGFR signaling have been in use since the early 2000's and can be divided into two types: small-molecule inhibitors which target the intracellular kinase domain, and monoclonal antibodies which target the extracellular region and block ligand binding[33]. Cetuximab is one such antibody which is employed, either as a single-agent or in combination with other therapies, for the treatment of lung, colorectal and head and neck cancers. Treatment with this drug results in statistically significant increases in overall survival, but these benefits are relatively short-lived; for example, in the clinical trial which earned cetuximab FDA approval for colorectal cancer, the overall response rate when used in combination with chemotherapy was only 23%, with median duration of response of 5.7 months[40–44]. Clearly there remains room for improvement.

The Wittrup lab has previously engineered antibody-based EGFR-targeted therapies which have achieved potent signaling blockade in vitro and anti-tumor efficacy in mice. The most promising construct was a fusion protein in which two unique fibronectin-based EGFR-binding domains were attached to different termini of cetuximab. The full protein thus incorporated 6 total binding paratopes capable of simultaneously engaging three unique epitopes on the EGFR extracellular domain. This topology, unlike cetuximab alone, caused the formation of large clusters of receptors which were shown to be downregulated through a mechanism of decreased endosomal recycling rate[69]. The construct did not agonize EGFR, and instead potently blocked signaling as measured by phosphorylation of downstream effector proteins Erk and Akt. Studies in nude mice showed that in vitro downregulation correlated with in vivo tumor control and that the fusion

protein was superior to cetuximab in multiple mouse tumor xenograft models, notably including K-Ras and B-Raf mutant cell lines[70].

This prior work was an important demonstration not only of a molecule with immediate therapeutic potential, but also an example of strategies for augmenting traditional molecularly targeted therapies like antibodies. While cetuximab alone derives its efficacy from signal inhibition via ligand competition and induced cell death through Fc-dependent mechanisms such as ADCC and CDC, the cetuximab fusion added to these functions the third mechanism of downregulation. To qualitatively evaluate the relative contributions of each of these functions mutations were made to ablate either Fc or variable domain binding. Some key results showing amount of receptor downregulation and tumor control for each of the constructs are provided in Figure 2.1. The data indicate that while downregulation and effector functions both contribute to efficacy in vivo, the most critical property seemed to be ability to compete ligand. As is shown in the figure, a mutation in the cetuximab variable region that decreases ligand competition by 7-fold (HND-LCAx) results in a complete loss of tumor control in all models tested despite only minimal losses in downregulation and presumably no change in effector function.

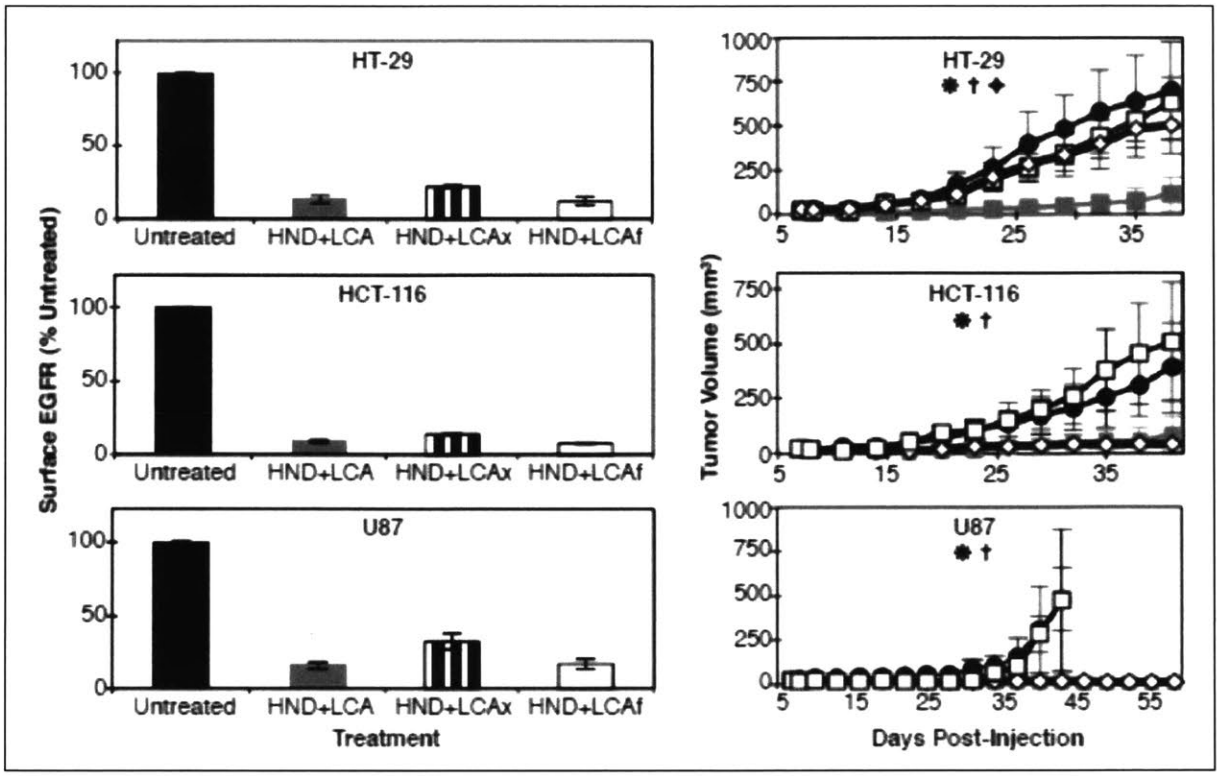


Figure 2.1 – Data from Spangler, et al. (2012), showing in vitro downregulation (left) and in vivo tumor xenograft growth (right) following treatment with cetuximab fusions and related mutants in three different cell lines. PBS (filled circles); HND+LCA (filled triangles) = fusion construct, unmutated; HND+LCAx (open squares) = Ab variable domain mutant; HND+LCAf (open diamonds) = Fc mutant.

This result is perhaps unsurprising in the context of the well-understood phenomena of ligand redundancy and signal amplification that occur in these receptor networks. It has been shown clearly both by experiment and mathematical modeling that signaling input to the EGFR network, i.e. ligand binding, is amplified through a series of downstream intracellular phosphorylation events before culminating in the output response[71–73]. This has important implications for EGFR-targeted therapies because it suggests that even when the vast majority of ligand binding is inhibited at the cell surface the amount of downstream signaling, and therefore cellular response, may be only minimally changed. It further implies that even small advances toward complete ligand inhibition over existing targeted therapies could result in greatly enhanced

anti-tumor efficacy. Indeed, a study using an oligoclonal anti-EGFR antibody mixture achieved only slightly better inhibition of ligand-mediated EGFR phosphorylation than cetuximab (~90% maximal inhibition for the mixture compared to 84% with cetuximab alone), but a corresponding 65-fold decrease in ERK phosphorylation[74]. Further, when a mathematical model was used to evaluate the factors that contribute most to the efficacy of EGFR-targeted therapeutics it was determined that cetuximab efficacy was particularly sensitive to ligand binding[75]. Overall, these findings suggest that a next generation multi-epitopic-targeting antibody fusion with enhanced ligand blocking could demonstrate superior anti-tumor efficacy.

An additional motivation for developing an improved EGFR antagonist would be the opportunity to explore combination treatments with various immunotherapies. A number of recent reports have demonstrated that EGFR signaling not only impacts tumor cell growth directly but also has immunomodulatory effects[56,58–62]. To the extent that cetuximab modifies the tumor immune environment and primes the immune system for response to therapies like checkpoint inhibitors, a more potent antagonist may provide an even more efficacious combination.

One opportunitiy for improvement upon the previous work was more practical in nature, concerning the potential developability of the cetuximab-Fn3 fusion molecule into a drug. Specifically, the cetuximab-Fn3 fusion was expressed relatively poorly from mammalian cells and affinity purified product was dominated by large aggregate species. Biophysical properties such as thermal and solution stability are critical characteristics of any protein-based therapeutic being considered for commercial development, as these parameters can affect downstream manufacturing processes such as purification and formulation and can impact the long-term storage potential of drug products. High yields from cellular expression systems are also desirable

in order to reduce production costs. Because of the unfortunate characteristics of the cetuximab-Fn3 fusion this molecule likely has limited commercial potential.

Antibodies, such as cetuximab, typically perform well against these developability criteria and their success in these categories is one reason for their growing representation among both approved pharmaceuticals and those in development. The tenth type-III domain of human fibronectin (Fn3), which serves as the basis for the cetuximab fusion partners, is less predictable. The WT Fn3 domain has high thermal stability (T_m ~80°C) and solubility, and isolated binders based on this scaffold typically, though not always, have favorable biophysical properties[76,77]. However, experience in the Wittrup lab suggests that while monomers of Fn3 are often stable and well-behaved, preparation of soluble Fn3 dimers is significantly more difficult, as these proteins are highly prone to aggregation. The tendency of Fn3 domain oligomers to aggregate could derive from the native function of cellular fibronectin, which forms insoluble fibrils that make up the extra-cellular matrix; the process of fibril formation is known to be mediated by fibronectin self-assembly[78,79]. In any case, the effect that Fn3 domains have on protein aggregation when used as fusion proteins is not definitively known.

It is possible, then, that future embodiments of an antibody-based EGFR antagonist could benefit from replacing the Fn3 domains with binders conferring more favorable biophysical properties. One possibility is the Sso7d protein, a DNA-binding protein native to the hyperthermophilic archaebacteria *Solfolobus solfataricus*. It is small in size (~7 kDa, 63 amino acids) and has extremely high thermal stability, with a T_m of ~98°C[80]. It is a globular protein with topology similar to a eukaryotic SH3 domain, 5 β-strands and a C-terminal α-helix. It has also previously been shown to be a suitable scaffold for directed evolution and selection of specific binders against multiple targets using yeast display[81]. Nine residues on a face containing three β-sheets, (the nine positions are shown in Figure 2.2 in pink) are the sites at which diversity is incorporated.

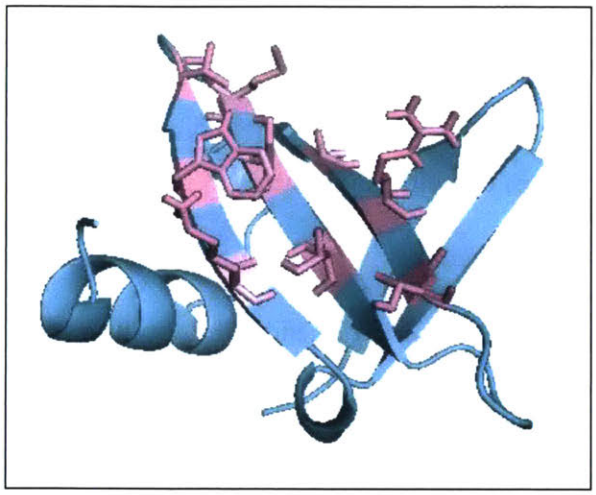


Figure 2.2 – Structure of the Sso7d scaffold protein, with the nine randomizable amino acids shown in pink. (PDB ID: 1BNZ)

Recent work in the Wittrup lab has focused on further improving the Sso7d scaffold by engineering a reduced-charge, truncated version which exhibits decreased non-specific sticking to cell surfaces without significantly changing the T_m. This optimized scaffold was used to build two naïve libraries each with different amino acid diversity at the nine randomized positions. One library, Sso18, allows all amino acids except cysteine and proline; the other, Sso11, incorporates only eleven amino acids which were chosen to mimic the distribution of amino acids in protein-protein interactions[82].

The improved Sso7 libraries were validated in selections against multiple targets including EGFR, K-Ras and mouse serum albumin (MSA)[82,83]. Specific binders were able to be identified

for all targets and the clones were amenable to affinity maturation using standard error-prone PCR libraries and yeast display methodologies. Additionally, the individual clones were found to be well-expressed in *E. coli*.

This work has opened the door for the possibility of using these libraries to identify novel binders that may serve as useful tools for building EGFR antagonists. This foundation, combined with the precedent that antibody fusions can be potent inhibitors of EGFR signaling both in vitro and in vivo, serves as the basis for the work that follows. This chapter begins by describing standard yeast display methods applied to the selection of Sso7d clones which fit the non-cetuximab competitive critera, and then introduces a novel, complementary, in silico approach which was used to augment the discovery and reveal low-frequency clone families. A number of trends and patterns observed during the subsequent process of Sso7d directed evolution are then discussed, as these lessons may impact future engineering efforts. Engineered Sso7d binders were ultimately used to build cetuximab-Sso7d fusions and the activity of these constructs was characterized in vitro. Following promising in vitro results, the chapter shows the disappointing lack of efficacy of these constructs in patient-derived xenograft mouse models, and discusses potential causes and implications for future embodiments of EGFR antagonists.

Results

Selection of non-cetuximab competitive Sso7d binders by yeast display

Two previously constructed and validated Sso7d yeast libraries, each consisting of approximately 10^9 unique members and called S11 and S18, were used for selecting Sso7d clones which bind the extracellular domain of EGFR. Keeping in mind the ultimate goal of multi-epitopic targeting, it was also required at the outset that these proteins not overlap with the cetuximab binding site on EGFR. To this end, a sorting scheme was developed as is shown in Figure. 2.3. First, the libraries were enriched for all Sso7d binders to EGFR using positive and negative bead selections and FACS sorts – this is referred to as the "all binders" stage. When sorting by flow cytometry, high antigen concentrations and generous gating strategies were used to ensure all Sso7d clones which bound the antigen were captured.

Then the collected populations were mutated by error-prone PCR and sorted simultaneously for both higher affinity and non-cetuximab competitive binding – the "competition" stage. During these steps the EGFR-Fc antigen was used at a lower concentration to limit the selections to higher affinity binders, and the antigen was pre-incubated with an excess of cetuximab in order to block binding of Sso7ds to this epitope. This competitive sorting scheme was first validated by flow cytometry before it was utilized in selections to ensure that the effect of excess antibody was as expected and gating strategies were appropriate. Added cetuximab had the anticipated effect: cells were either unaffected by its presence or had their binding completely ablated, as is shown in Figure 2.4.

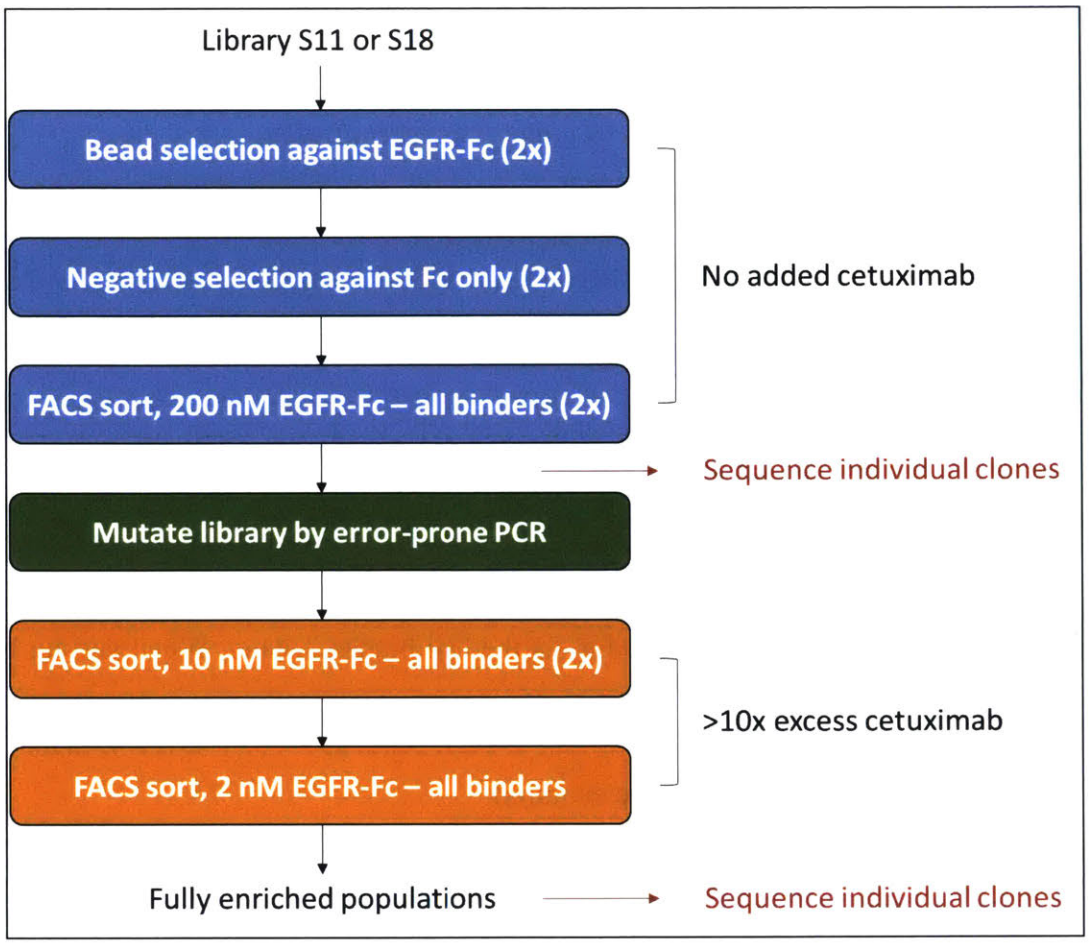


Figure 2.3 – Initial sorting scheme for identifying high-affinity, non-cetuximab competitive EGFR binders from an Sso7d library by yeast display. First, an "all binders stage" (blue) was carried out to identify all EGFR-binding Sso7ds, using high concentrations of EGFR-Fc antigen and generous gating. The resulting population was then mutagenized and a second, "competition stage" (orange) was performed using lower antigen concentration and an excess of pre-incubated cetuximab. Sequencing steps are indicated.

Selections for high affinity, non-cetuximab competitive binding were carried out in this manner until the mutated libraries were fully enriched for a population that bound to 2 nMEGFR-Fc in the presence of cetuximab; this took three sorts (Table 2.1). The resulting clones were then analyzed by sequencing individual yeast colonies after the third sort. When ~30 sequences were obtained in this manner the result was a largely homogeneous population and a dominant consensus clone for each library (Figure 2.5). Certain paratope positions contained more diversity than others and the

consensus for the two libraries was not identical but qualitatively these populations looked very similar. In particular, paratope positions 1, 3, 8 and 9 were common between the two libraries and seemed to be the defining features of a single, overall family. Position 5 was largely conserved within each enriched library but distinct between the two; in fact, the entire center beta strand, containing paratope positions 4-6, differed between the two libraries. Among all the clones sequenced for both libraries, only one was found that was completely distinct from the observed consensus (Figure 2.5, bottom right).

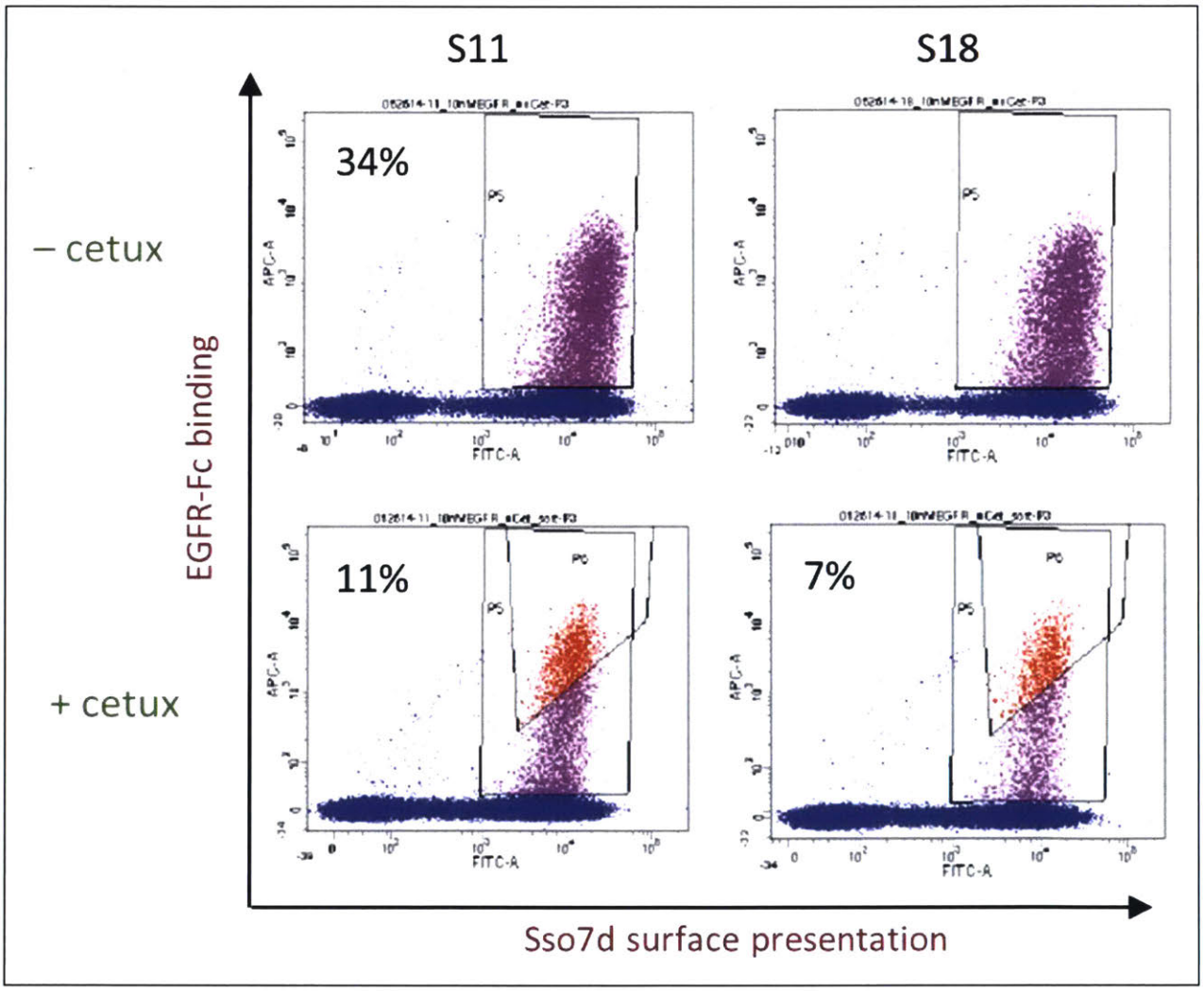


Figure 2.4 –Effect of cetuximab pre-incubation on Sso7d library binding to EGFR-Fc. In one representative sort the percentage of yeast cells binding to 40 nM EGFR-Fc drops by approximately 75% for library S11 and library S18, when 200 nM cetuximab is pre-incubated with

EGFR-Fc for 15 minutes. Average fluorescence intensity of remaining binding population is unchanged. Percentages indicate number of expressing cells that are positive for binding.

Library	1.0		1.1		1.2		1.3	
	S11	S18	S11	S18	S11	S18	S11	S18
– Cet. (% pos)	34	35	69	64	74	73	86	88
+ Cet. (% pos)	11	7	60	58	59	59	79	83
% non-comp	31	20	88	89	80	81	92	95

Table 2.1 – Evolution of EGFR-Fc binding population during sorts for high-affinity non-cetuximab competitive binding. Percent of Sso7d-expressing yeast cells which were positive for binding in the absence or presence of cetuximab competition are shown, for each subsequent sort. The percent of binding cells that are non-cetuximab competitive is calculated as (% pos with cetux) / (% pos without cetux). Fully non-competitive binding was acheieved after only two sorts, but a remaining fraction of non-EGFR binding cells took longer to remove.

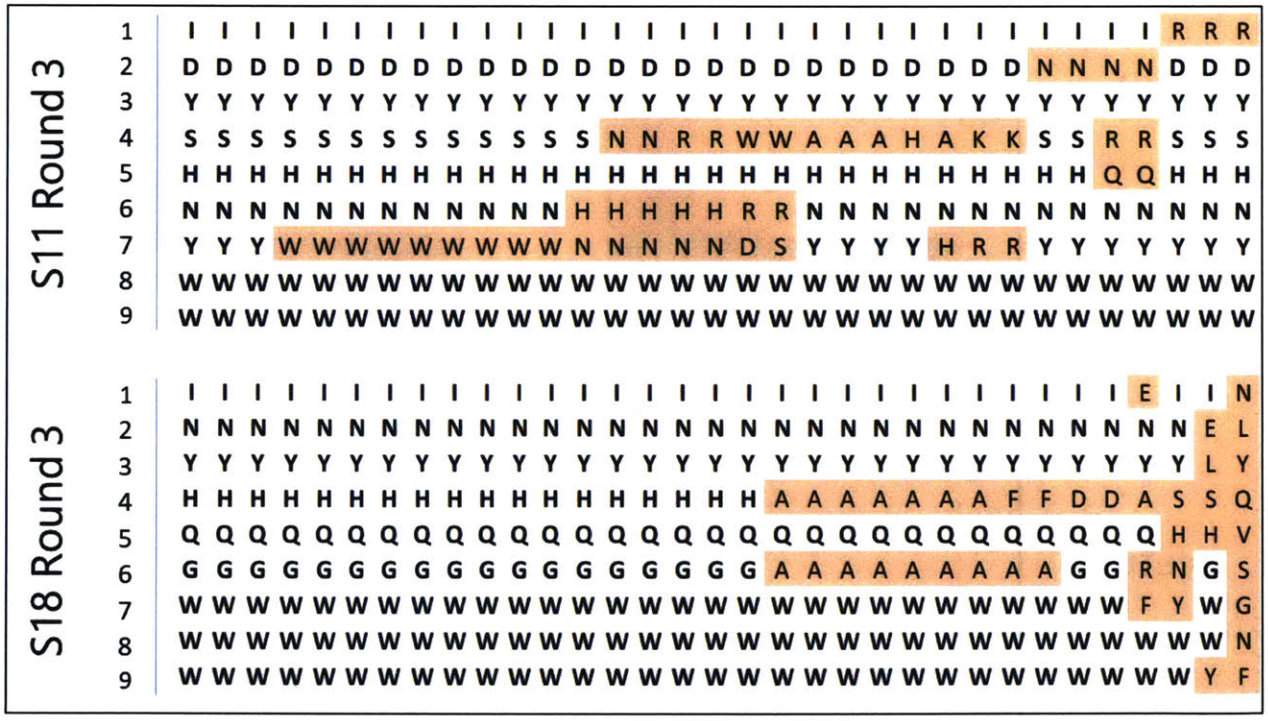


Figure 2.5 – Individual Sso7d clones sequenced after four sorts for high-affinity, non-cetuximab competitive binders as described in Figure 2.1. Each column represents an individual Sso7d, identified by its nine paratope residues, labeled 1-9. The most frequent amino acid at each position within each library is defined as the "consensus", and deviations from this are highlighted in orange.

Intuition suggested that none of these highly similar consensus clones shown in Figure 2.5 (i.e. everything except NLYQVSGNF) were likely to be epitopically distinct. To test this explicitly, a subset of Sso7d clones were expressed solubly as SUMO-fusions and were binned by performing cross-blocking experients on EGFR-expressing A431 cells. As shown in Figure 2.6 every consensus clone tested in this assay was capable of blocking binding of biotin-labeled IDYSHNYWW when present in excess. Even the difference in consensus between the two libraries (i.e. the distinct center beta strand) was not enough of a distinguishing feature to constitute a distinct epitope. The only exception was the NLYQVSGNF clone, which had no effect on binding of any other Sso7d clone tested. Thus, the final yield of this sorting scheme was two clone "families", and importantly, only two tools that could be used to help build a multi-epitopic binding protein.

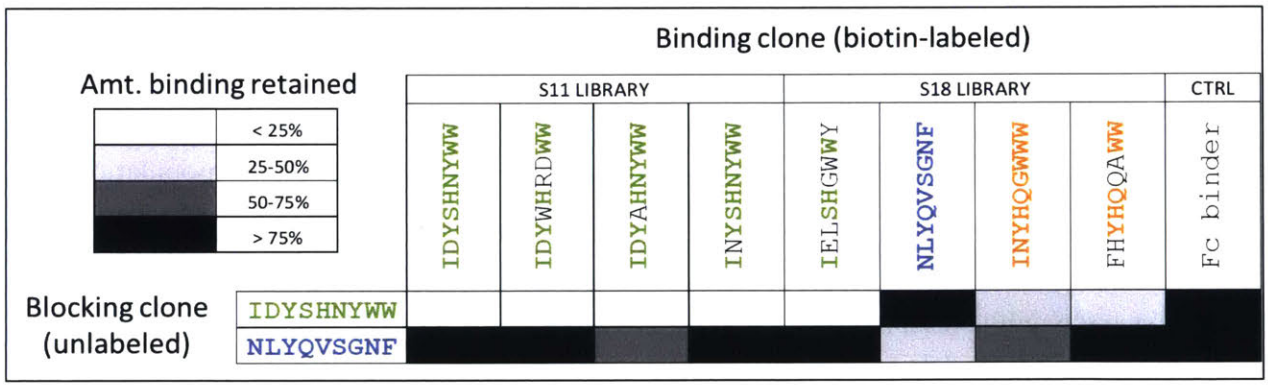


Figure 2.6 – Cross blocking of non-cetuximab competitive Sso7d clones derived from first round of sequencing. EGFR-expressing A431 were incubated with indicated blocking clone for 15 minutes before addition of biotin-labeled Sso7d binding clone. Amount of binding clone fluorescence was analyzed by FACS. Data are presented as amount of binding clone retained in the presence of blocking clone, normalized to binding with no blocking.

In silico analysis to uncover unique EGFR binding clone families

Previous work from the Wittrup lab demonstrated that targeting three epitopes is sufficient to drive EGFR downregulation and antagonize receptor signaling, and thus two Sso7d clones against non-cetuximab epitopes are the minimum required. But this work also showed that not all

combinations of epitopes are functionally equivalent. Thus, despite successfully identifying two clones from the initial sorting scheme, we sought a way to probe the library more deeply.

As discussed above, analysis of the sequences after the "competitive" stage of sorts revealed near complete homogeneity; but sequencing of the populations immediately after the "all binders" stage (i.e. at the first sequencing step in Figure 2.3) revealed a different distribution, as shown in Figure 2.7. This picture of the libraries, selected without regard for cetuximab competition or for affinity, is more diverse. And while there is clearly some bias within each library toward what ultimately emerged as the consensus clone, particularly for S11, the number of non-consensus sequences is much higher at this early stage. This finding is important because it suggests that the limited breadth of the ultimate non-cetuximab competitive population was not due to some intrinsic property of the Sso7d scaffold or the naïve libraries themselves – there are many unique Sso7d paratopes which are capable of specifically engaging EGFR. Rather, the limitation seems to be that the designed sorting scheme was too aggressive and had excessively narrowed the clonal diversity. Consequently, returning to this "all binders" population and designing a different subsequent selection scheme might yield additional non-cetuximab competitive clone families.

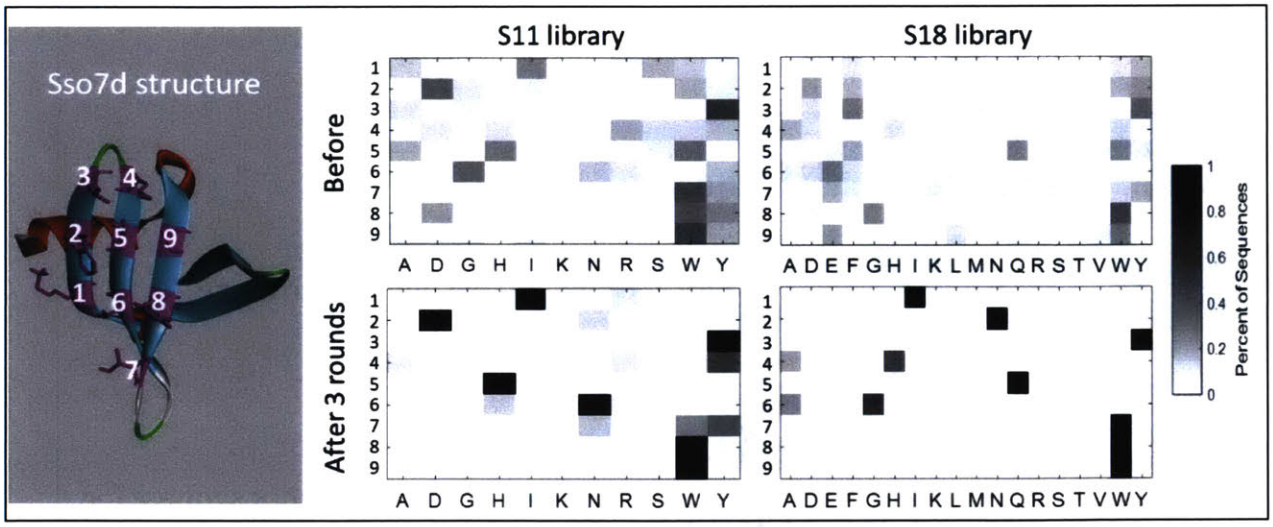


Figure 2.7 – Distribution of paratope amino acids before and after sorting for high affinity, non-cetuximab competitive binding to EGFR-Fc, sampled by sequencing single yeast clones. Sso7d structure with paratope (magenta residues) numbering scheme is shown at left. Heat map of amino acid frequency from 33 sequences for the S11 and S18 libraries is shown at right. "Before" population is after isolation of all EGFR binders but before mutagenesis and sorting with cetuximab competition. "After" is after three competitive sorts. (See scheme in Figure 2.3 for more detail.)

The alternate selection strategy that was designed, shown in Figure 2.8, involved a number of protocol changes. First, because a reasonably diverse population was still observable after the "all binders" stage, this population was used as the starting point for subsequent sorts. Because selection for high affinity may have unnecessarily limited the diversity during the "competitive" stage this criteria was abandoned. The starting population was used as-is rather than being mutagenized by error-prone PCR and sorts were performed with higher antigen concentration, also to reduce the bias toward high affinity binders. Additionally, in order to prevent the emergence of a dominant clone over multiple selections, only a single sort with cetuximab competition was performed, and the entire positive population was collected.

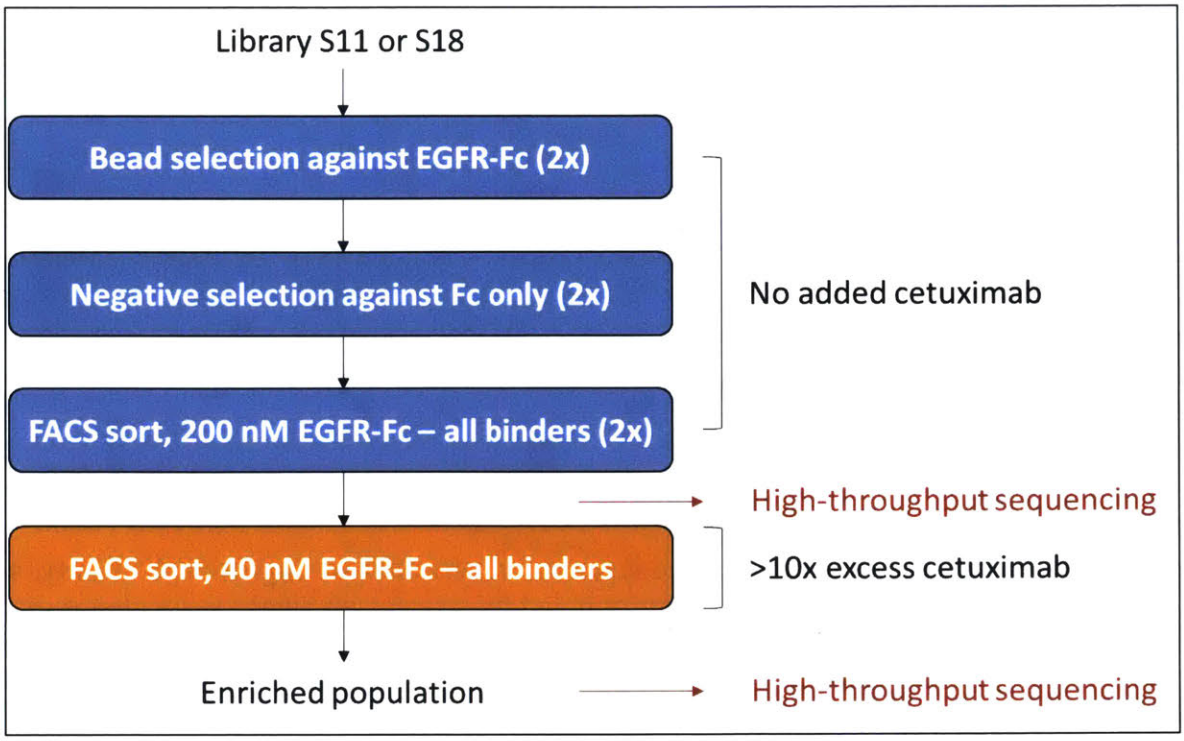


Figure 2.8 – Alternate scheme for selecting non-cetuximab competitive clone families. After the "all binders" stage from Figure 2.3, libraries were sorted one more time for all binders in the presence of cetuximab competition. Pre- and post-sort populations were sent for high-throughput sequencing.

The final change was a shift from sequencing single colonies to sequencing the entire library using next-gen sequencing. This approach was deemed most appropriate for two reasons: first, if the library does contain non-competitive clones that are very rare, this should allow their discovery without requiring analysis of an excessive number of individual clones. Second, the libraries can be sequenced both before and after the sort, and thus a change in frequency of each clone can be calculated. This information is useful because while there is presumed benefit to characterizing populations after only one sort, the accompanying risk is a likely higher rate of false positives; enrichment values can be used to mitigate this risk and reduce the false discovery rate.

A single competitive sort was performed according to these changes, and the resulting plots are shown in Figure 2.9. As was previously observed, the effect of cetuximab competition on an

individual yeast cell was either nothing or complete ablation of binding, again indicating that the competitive scheme was effective as designed. DNA was then isolated from the yeast populations both before and after the sort and analyzed on the Illumina MiSeq. A total of 10 million reads, roughly 2.5 million for pre- and post-sort for both the S11 and S18 library were acquired and analyzed.

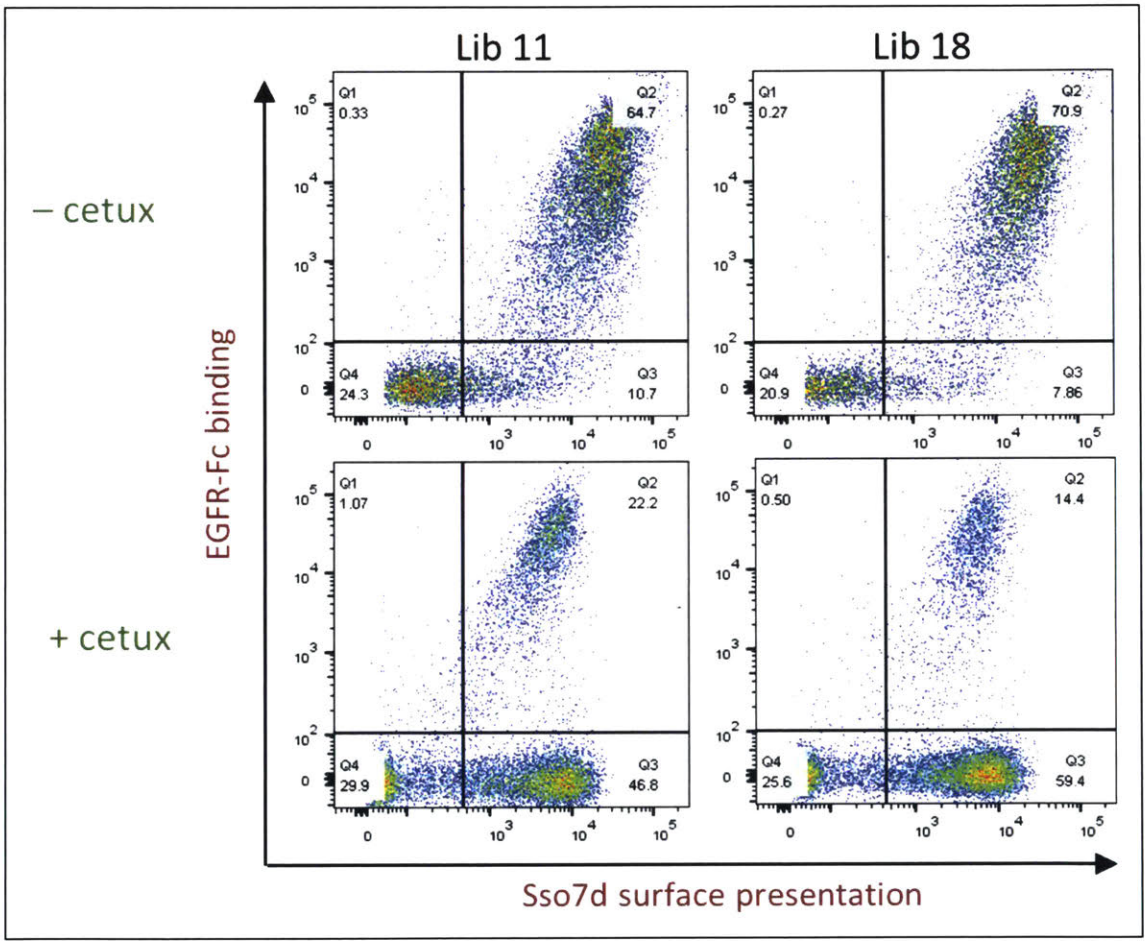


Figure 2.9 – Sort for non-cetuximab competitive population to be analyzed by high-throughput sequencing (HTS). Populations after the "all binders" stage in Figure 2.3 were subjected to a single sort for binding to 40 nM EGFR-Fc in the presence of 400 nM cetuximab (bottom row). All cells in the upper right quadrant were collected by FACS. Shown for comparison is the binding of the population without competition (top row). The top right quadrant of each plot represents a population that was analyzed by HTS.

Analysis of aligned full-length sequences began with simple characterization of the amino acid distributions in the pre- and post-sort libraries. Heat maps of amino acid content at each paratope position are shown in Figure 2.10. Overall, these data are surprisingly similar to Figure 2.7. The "before" distributions in the two figures are nearly identical, despite the fact that one plot used sequences from 33 colonies picked by hand and the other used 2.5 million reads obtained by HTS. In both cases the libraries appear to contain more diversity than just a single family, but the distribution is clearly not uniform. Figure 2.10 shows obvious bias in the "before" population, as does Figure 2.7, toward the previously identified consensus family even though the selections lacked cetuximab competition and were designed to retain all EGFR binders. Also as in Figure 2.7 this bias is more apparent for library S11 than S18, which is likely a simple reflection of the fact that fewer paratope combinations are possible in the library made with fewer amino acids.

Surprisingly, however, the "after" populations in Figure 2.7 and 2.10 are also highly similar, both showing clear convergence on a dominant clone family. In the former case this outcome had been attributed to repeated selections and perhaps unnecessary stringency for high affinity binders. But the altered sorting approach used to generate the data in Figure 2.10 was designed specifically to avoid this outcome and retain as much diversity as possible. The fact that the distribution still turned out to be largely homogeneous means that this finding is not simply the result of overly aggressive sorting criteria, but rather is because the vast majority of the diverse Sso7ds engage the cetuximab/ligand binding site of EGFR, and these are effectively eliminated during even a single competitive sort. Such a finding could either be consistent with the idea of an epitopic "hot spot", or could suggest intrinsic limitatons of the Sso7d scaffold, as is discussed further below.

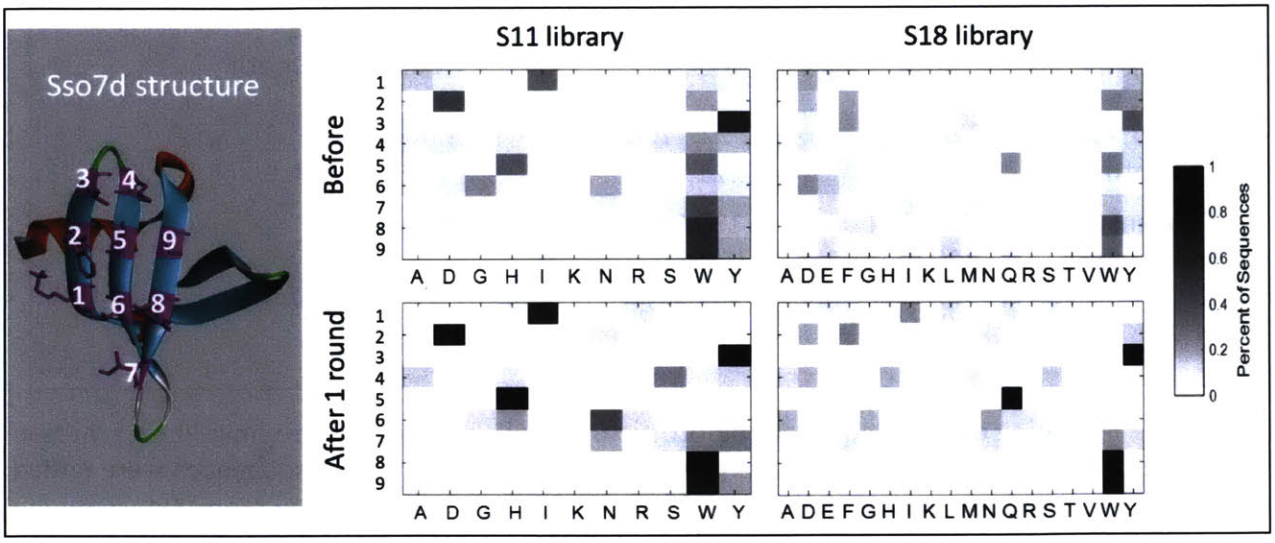


Figure 2.10 – Distribution of paratope amino acids before and after a single sort for non-cetuximab competitive binding to EGFR-Fc, sampled by HTS. Sso7d structure with paratope (magenta residues) numbering scheme is shown at left. Heat map of amino acid frequency is calculated based on the ~2.5 million aligned, full-length sequences for each population from the HTS data. "Before" is after isolation of all EGFR binders but before sorting with cetuximab competition. "After" is population following a single competitive sort in which all positive events were collected.

The "after" data in Figure 2.10 indicate that the vast majority of sequences in the sorted population were members of the consensus family already characterized and represented by the IDYSHNYWW clone. Further, the data in Figure 2.6 suggested that every member of this family was likely to bind the same epitope of EGFR. Since the goal of the HTS approach was to find additional epitopically-distinct families, the next obvious step in the HTS analysis was to prune the sequence set by removing any clones that fit the observed consensus. To this end, a four amino acid motif was defined for each library based on the heat maps in Figure 2.10 and all clones matching this motif were discarded. Figure 2.11 shows that, as expected, this motif accounted for the vast majority of clones in the sorted populations. In both libraries it represented over 90% of the total sequences, while in terms of unique sequences it constituted 55% of the S11 population and 80% of the S18 population.

# seqs (# unique)	S11 Lib	S18 Lib
Aligned, full-length	2,620,825 (91,239)	2,522,128 (221,735)
Consensus removed	126,354 (40,055)	185,925 (44,043)
Enriched only	96,371 (35,146)	159,836 (39,496)

Consensus motif

	1	2	3	4	5	6	7	8	9
S11:	-	-	Y	-	H	-	-	W	W
S18:	-	-	Y	-	Q	-	-	W	W

De-enriched ≡ $\log_2(f_{sorted} / f_{unsorted}) < -1$

Figure 2.11 – Pruning of HTS sequences before hierarchical clustering. Approximately 2.5 million reads were obtained for each of the S11 and S18 post-sort libraries. These sequences were culled by removal of a consensus motif, defined as shown at right, where numbering represents paratope amino acids, and by removing de-enriched sequences using the defined cutoff. Table shows number of total sequences and corresponding number of unique sequences, in parenthesis, at each stage of analysis.

Next, the frequency of each individual clone both before and after the sort was used to calculate an enrichment ratio: $enr \equiv \log_2(f_{sorted}/f_{unsorted})$. It was hypothesized that clones which were de-enriched, or less frequent in the sorted library than the unsorted, were likely to be clones acquired during the sort as a result of noise or imperfect separation, and that these should be removed from consideration. This assumption was validated using a previous study in which Sso7d binders against EGFR had been derived from this same naïve library without concern for epitopes[82]. Traxylmayr, et al reported four Sso7d clones which were competitive with cetuximab; when the enrichment ratios for these sequences were calculated using the HTS data it was found that for all four $enr < -3.8$, indicating substantial decreases in frequency. Thus, a conservative cutoff for de-enrichment was defined as $enr < -1$, and sequences with ratios below this threshold were removed. This eliminated a further ~20% of total sequences and ~10% of unique sequences for each library.

The remaining unique sequences were then rank-ordered by enrichment ratio and the top 2000 were considered for further analysis. This arbitrary cutoff was applied mostly for

convenience and to reduce computing power required for subsequent steps, but was still sufficient to capture most of the sequences of interest since it included everything with an enrichment ratio of $enr > 0.46$ for S11 or $enr > 0$ for S18.

Finally, these 2000 most enriched unique sequences from each library were hierarchically clustered based on sequence similarity using a Gonnet scoring matrix, with cluster sizes determined based on intra-cluster inertia. Clusters were represented as sequence logos and were further manually refined to combine any which had the more than three of the nine amino acids the same. The final results of the clustering analysis are shown in Figure 2.12.

Eight clusters, four from each library, were derived from this phylogenetic grouping. In the S18 library two clusters matched with previously observed sequences shown in Figure 2.5: cluster 18.1 had been shown to overlap with the IDYSHNYWW consensus (Figure 2.6), and cluster 18.3 was the second non-cetuximab competitive epitope already discovered. Interestingly, the overall frequency of these families within the HTS data was slightly above 1% – this matches roughly with the discovery rate of these families using just a small number of single clones, where their observed frequency was 1/66. The final S18 cluster, 18.4, was distinct and novel but was found to bind specifically to human Fc, not EGFR; the clones represented by this family were apparently not removed during the negative bead selections against Fc during the "all binders" stage. They were favored for selection and enriched during the competitive sort because the antigen used was EGFR-Fc. In library S11 the 2000 sequences clustered were much more similar overall, as indicated by the shorter linkage heights in the phylogenetic trees. In fact, most of these sequences fit within a single large family, cluster 11.1, which was very similar to the IDYSHNYWW consensus and assumed to target the same EGFR epitope. However, the remaining four clone families, 11.2, 11.3, 11.4 and 18.2, appeared to contain sequences not

previously observed. They were all present at low to very low frequencies within the sorted library, making it unsurprising that they had not been discovered when sequencing colonies individually.

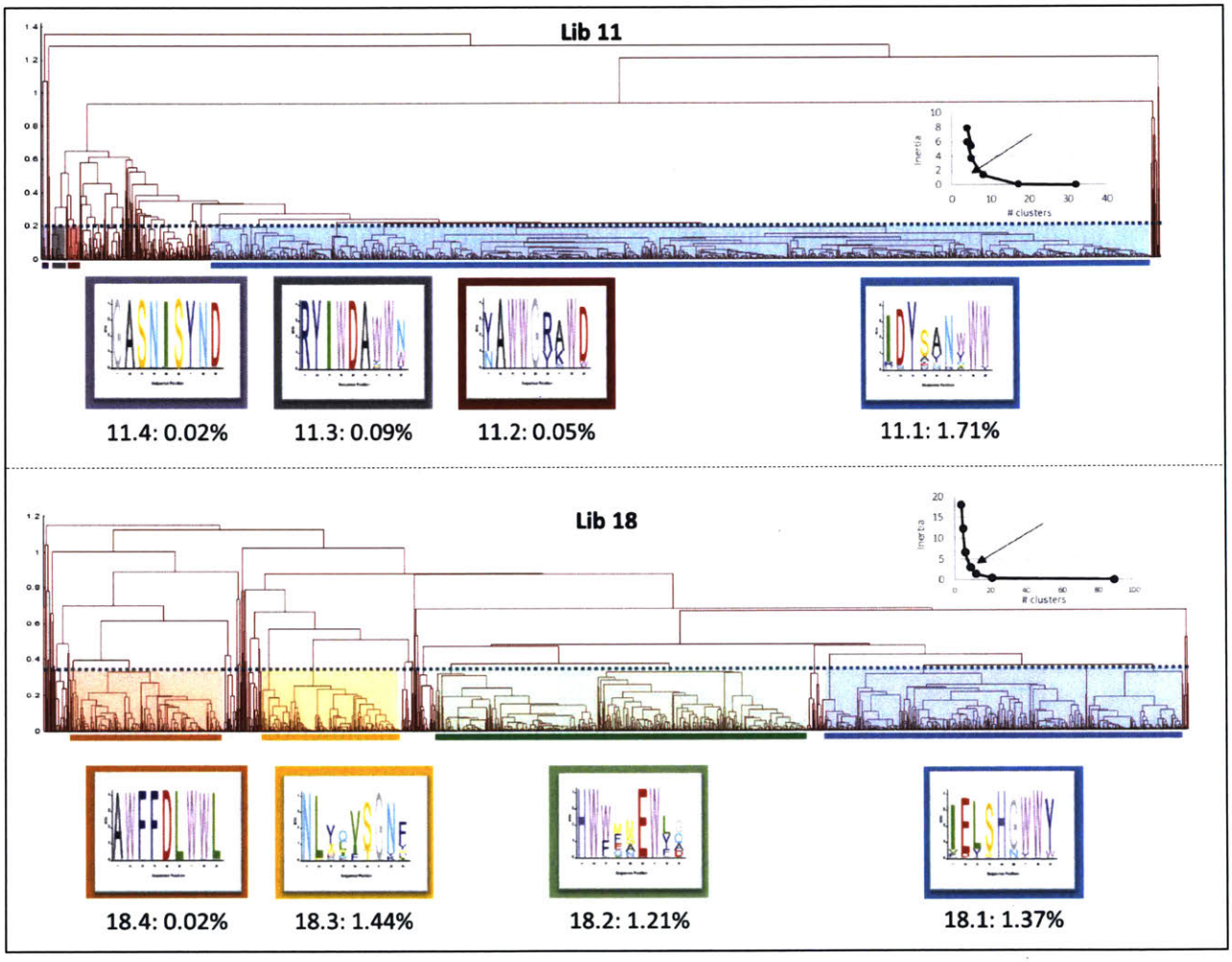


Figure 2.12 – Clustering results and clone families derived from clustering of top 2000 unique enriched paratope sequences. Top, for each library: Phylogenetic tree based on sequence similarity by Gonnet scoring matrix. Sequences are individual points along x-axis; sequence dissimilarity is indicated by linkage height. Blue dashed line indicates cutoff for determining clusters based on calculation of intra-cluster inertia (at the "elbow" of the inset plots). Highlighted areas are final cluster families (colors are arbitrary). Bottom, for each library: Families are represented as sequence logos. Colors correspond to location in the phylogenetic tree. Numbers below the logos are the cluster number and the overall frequency of the family within the HTS data.

A representative sequence from each new clone family was selected and these were combined with the two sequences previously identified by single clone sequencing to constitute

the candidate panel of non-cetuximab competitive EGFR-binding Sso7d clones. All six clones had a neutral or nearly neutral net charge, though all but clone B included at least one charged amino acid within the paratope. In accordance with previous observations by Traxlmayr, et al., the panel of binders was strongly enriched with aromatic residues; in four of the six clones almost half of the paratope positions (4/9) contained an aromatic side chain.

When these clones were tested for EGFR binding on the surface of yeast all were positive for binding (Figure 2.13). The MFIs varied greatly among the clones, likely reflecting differences in affinity, but in each case the extent of binding was unchanged in the presence of cetuximab, confirming that all clones were non-cetuximab competitive. Thus, by deeply sequencing the libraries and analyzing the results in the context of previously known patterns (i.e. removing the dominant consensus clone family from consideration) this in silico approach was effective in identifying rare, distinct clone families from what otherwise appeared to be a largely homogeneous population. The approach expanded the tool set available for making multi-epitopic targeting therapeutics from two clones to six.

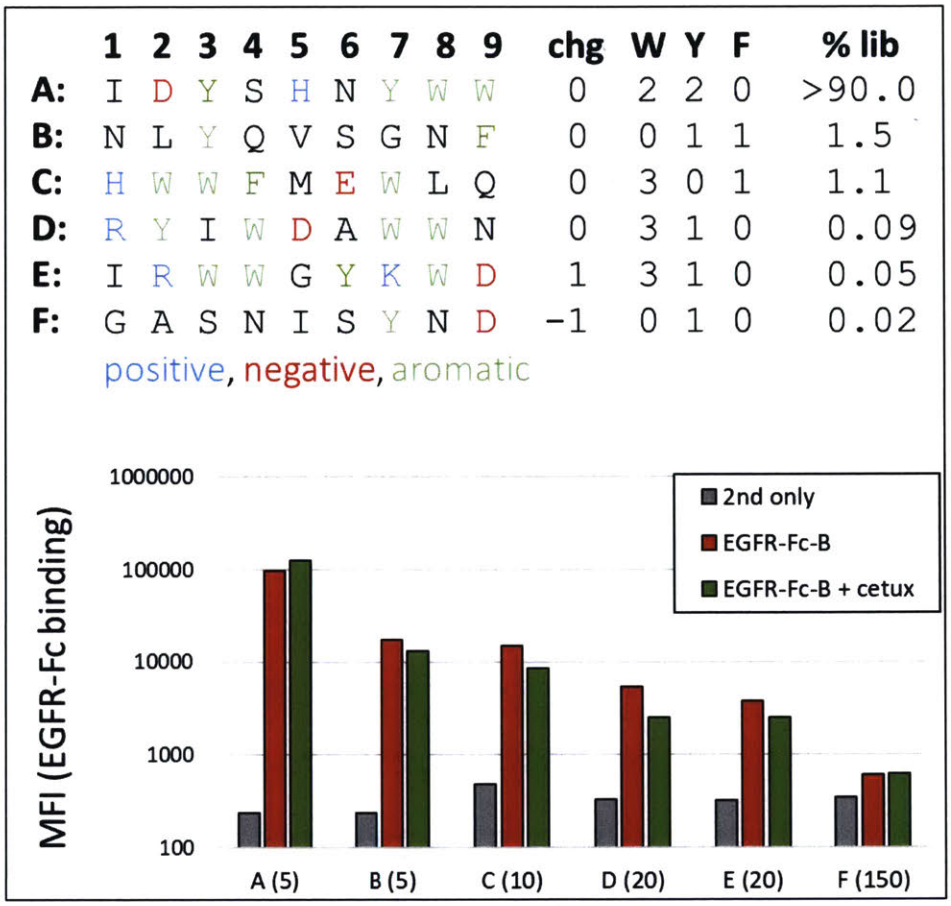


Figure 2.13 – Non-cetuximab competitive panel of Sso7d binders. **Top** – A representative sequence from each clone family was chosen for further characterization. Net paratope charge, hydrophobic content and frequency within the HTS data set are shown. **Bottom** – EGFR binding of yeast-displayed Sso7d to soluble EGFR-Fc, displayed as mean fluorescence intensity (MFI). Sso7d clone is indicated by letter, concentration (nM) of EGFR-Fc shown in parentheses. Competition binding used 10-fold excess of cetuximab.

Affinity maturation of EGFR-targeting Sso7ds

Affinity of the six Sso7ds for EGFR was variable, and roughly corresponded with frequency within the library. Binding to a bivalent EGFR construct could be detected for all Sso7d clones (Figure 2.13), but when a monovalent EGFR construct was used (EGFR-monoFc, see Methods) the affinity was too weak to measure for several clones (Figure 2.14). No specific affinity threshold is required for the purpose of multi-epitopic targeting, but intuition suggests that the higher the Sso7d affinity for EGFR, the better. This is particularly true if one desired consequence

is increased ligand binding inhibition. To this end, the Sso7d clones were affinity matured using yeast display.

Affinity maturation was achieved for five of the six clones over several successive rounds of sorts, each time using error-prone PCR to mutate the entire Sso7d gene. FACS selections first used the bivalent EGFR-Fc construct as antigen and then the monovalent version in later rounds to increase the stringency of selections. Sequence evolution and affinity at each round are shown in Figure 2.14. The affinity of all clones was improved over the course of three or four rounds of maturation, but the round-by-round improvements varied dramatically both within and among clone families. For example, clone A, which was derived from the single clone sequencing after error-prone PCR and already included two backbone mutations (Q56L and Q61R), started at the highest affinity of the panel but barely improved over four rounds. Clone B made small but steady progress of approximately 3-fold decrease in K_d per round; and still others, such as clone D, changed in large increments during some rounds, but none in others. In general, progress seemed to slow when the affinities reached the double-digit nanomolar range, and while it was possible to obtain a picomolar affinity binder (clone D), the slightly weaker affinities of the remaining clones are more in line with what was obtained by Traxylmayr et al and Kauke et al for EGFR, MSA and K-ras binders. The difficulty obtaining binders with sub-nanomolar affinities may be reflective of intrinsic limiations of the Sso7d scaffold itself.

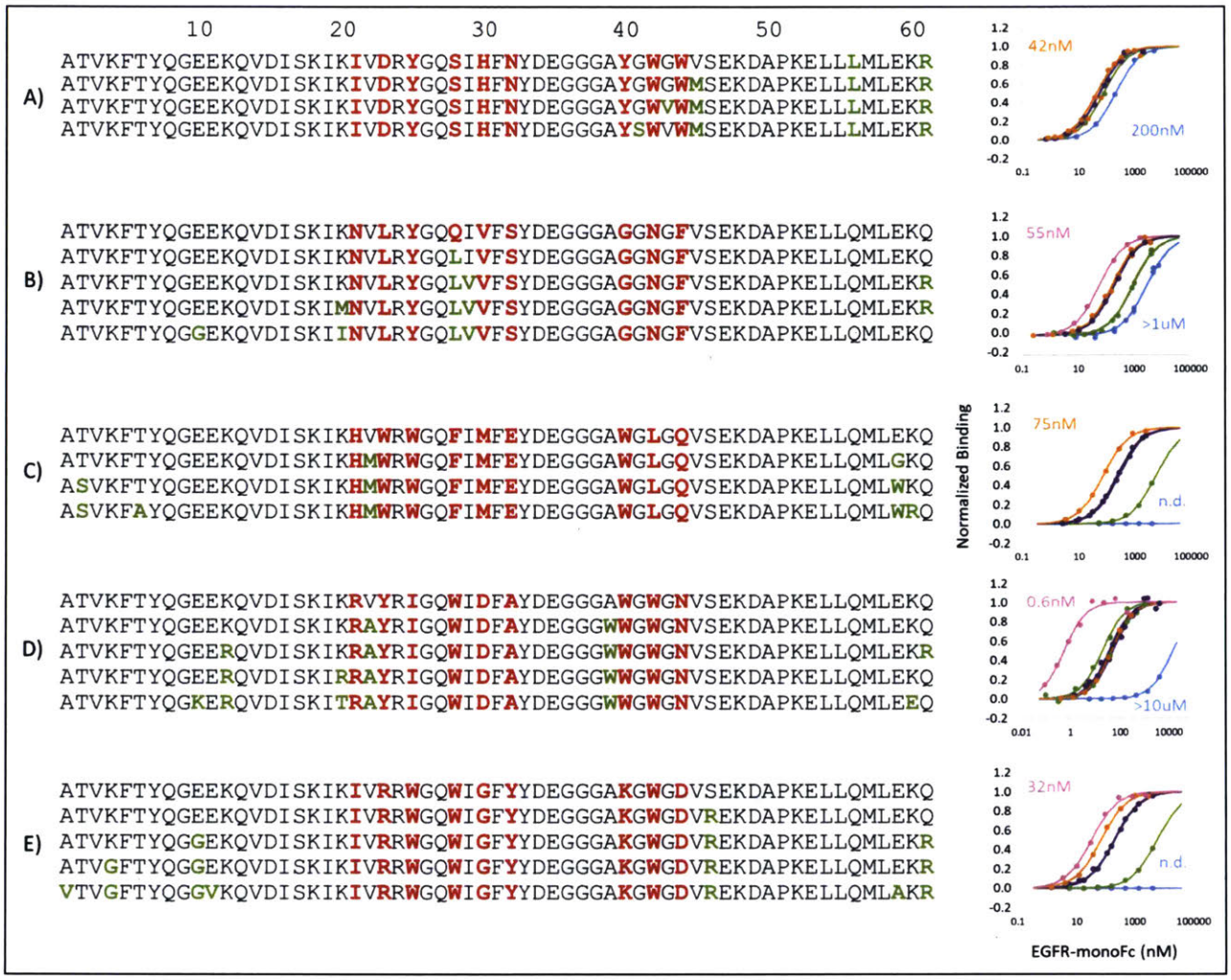


Figure 2.14 – Affinity maturation of Sso7ds. For each clone (indicated by letter at left), the sequence evolution and corresponding monovalent EGFR (EGFR-monoFc) titration binding curves are shown. Parental sequence – chosen either from single clone sequencing or from in silico analysis – is listed first with paratope residues in red. Green indicates a mutation from the original library scaffold as a result of error-prone PCR. Titration curve coloring indicates number of affinity maturation rounds and is as follows: blue = parental, green = first round, purple = second round, orange = third round, pink = fourth round. Points are data and curves are fit to monovalent binding isotherm. Starting affinity for each parental clone indicated in blue in bottom right corner, final affinity of matured clone indicated in appropriate color in top left corner.

Double-digit nanomolar K_ds like those of Sso7d clones A, B, C and E after four rounds of evolution would likely be insufficient for potent antagonism in the context of a monovalent protein, but as cetuximab fusions these Sso7ds would have the benefit of a highly avid format. Weighing this factor against the increasing difficulty of generating tighter binders, it was decided that this

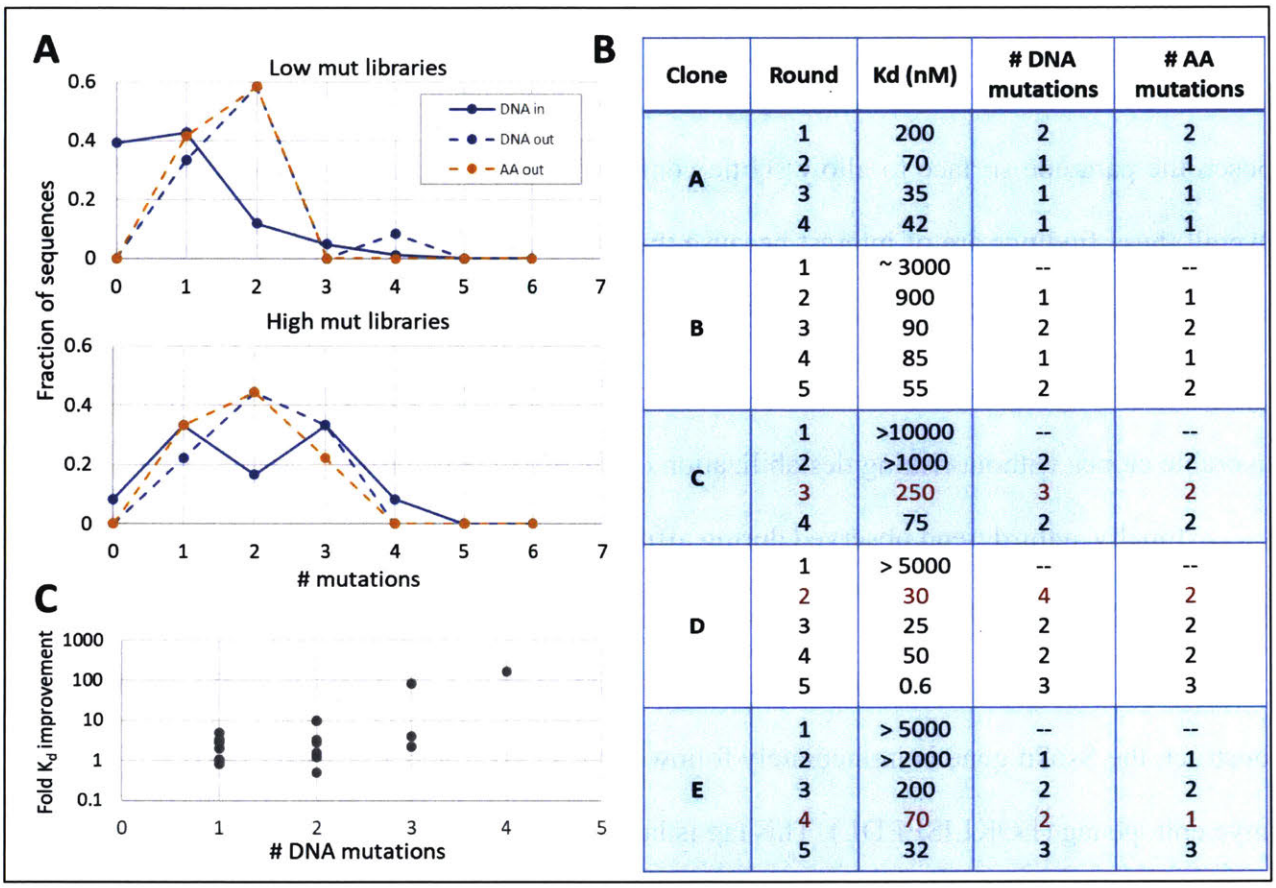
moderate affinity range was sufficient for the purpose of building a multi-epitopic targeting construct. Ultimately, affinity maturation was concluded with the affinities of the five clones ranging from 0.6 – 75 nM, all representing substantial improvements from the parental sequences.

Insights from Sso7d affinity maturation – implications for future engineering

The directed evolution detailed here for the five Sso7d clones required a total of 20 rounds of library building, selection and sequencing. In addition to generating the desired affinity for enabling EGFR antagonism, this extensive protein engineering exercise also revealed a number of trends that are likely unrelated to EGFR. These findings are unrelated to the subsequent development of cetuximab-Sso7d fusions but are presented here because they might constitute actionable lessons for future Sso7d engineering efforts.

The first observation is the impact that mutation rate has on the affinity maturation process. Studies of antibody scFv affinity maturation have shown that optimal mutation rate for error-prone PCR is 1-9 amino acids per scFv gene. As an estimate for optimal mutation rate for Sso7d affinity maturation a simple scaling argument was used – since the Sso7d gene is roughly a quarter of the size of an scFv, mutagenesis was performed aiming for a mutation rate of approximately 1-2 errors per Sso7d. However, when mutation distributions were analyzed after the first two rounds of selections on each lineage (i.e. after the first 10 libraries had been built and sorted) it was clear from the enriched clones that higher mutational burden was preferred. Most of the input library had 0-1 mutations per clone, but the output clones most frequently had 2 mutations (Figure 2.15 A). Further, there appeared to be a correlation between the number of mutations and larger fold improvement in K_d (Figure 2.15 C). This inefficiency in construction was corrected for subsequent rounds, which were built with additional amplification cycles during the error-prone PCR step, and as the input distribution widened so did the output (Figure 2.15 A).

Also interesting was the finding that in three of the libraries, the most improved clone had more DNA mutations than amino acid, meaning that multiple gene changes occurred within a single codon. Because of unequal codon degeneracy certain amino acid point mutations can only be realized when two or even all three nucleotide bases are mutated together; for example tryptophan can only be encoded by a TGG codon, and only five other amino acids can reach a tryptophan with a single nucleotide change. The importance of mutations like this seems well illustrated by the clone D lineage, where during maturation from round 1 to round 2 the largest single fold increase was realized (over 100-fold) and the daughter clone had three consecutive mutations in a single codon to make the A39W mutation (see also Figure 2.14 for clone D amino acid sequences). It is, of course, impossible to design a library that can account for all of the extremely low-probability events like this particular illustrative case, but overall the message seems clear: if the intrinsic properties of the Sso7d scaffold (small size, more rigid structure) mean there are fewer higher affinity "solutions" to the energy optimization problem posed during affinity maturation, then it follows that higher mutation rates would improve the chances of generating an improved clone.



B

Clone	Round	Kd (nM)	# DNA mutations	# AA mutations
A	1	200	2	2
	2	70	1	1
	3	35	1	1
	4	42	1	1
B	1	~ 3000	--	--
	2	900	1	1
	3	90	2	2
	4	85	1	1
	5	55	2	2
C	1	>10000	--	--
	2	>1000	2	2
	3	250	3	2
	4	75	2	2
D	1	> 5000	--	--
	2	30	4	2
	3	25	2	2
	4	50	2	2
	5	0.6	3	3
E	1	> 5000	--	--
	2	> 2000	1	1
	3	200	2	2
	4	70	2	1
	5	32	3	3

Figure 2.15 – Mutational outputs of error-prone PCR library inputs and outputs. **A** – Single clone sequences from the input and output of the first 10 libraries (top) and final 10 libraries (bottom) constructed and sorted by error-prone PCR. **B** – Consensus clone K_d (nM) based on yeast surface titration with EGFR-monoFc for each round of maturation. Corresponding number of DNA and amino acid mutations are shown. Red highlighted rounds ended with multiple DNA mutations in a single codon. **C** – Fold K_d improvement plotted versus number of DNA mutations.

The second observation is that the affinity maturation of the Sso7ds rarely involved changes in the paratope. Affinity maturation of antibodies by error-prone PCR can also produce mutations outside the paratope (i.e the CDRs), but common practice is to discard these mutants for fear that constant region mutations could disrupt antibody structure. However, among all five lineages shown in Figure 2.14, only one mutation ever occurred within the nine residues of the Sso7d paratope. Instead, mutations were frequently found adjacent to or between paratope amino acids, where the side chains point inward toward the Sso7d protein core. It is only possible to

speculate on the exact mechanism by which these mutations confer higher affinity, but it is clear that most of them do not contact the antigen, and thus a reasonable hypothesis is that they alter or loosen the paratope surface to allow subtle conformational changes that support better binding. Overall these findings are of interest because they suggest that perhaps a more limited region of the Sso7d gene should be mutated when building libraries for affinity maturation. A higher density of mutations confined to a more limited region of the protein might yield a higher number of favorable clones without risking destabilization of the Sso7d structure.

Finally, a third trend observed during affinity maturation was the propensity for alterations at the C-terminal end of the Sso7d sequence. In particular, it was striking that in four of the five lineages evolved, the final two amino acids were mutated from a KQ to a KR. In the yeast display construct, the Sso7d gene is immediately followed by a two amino acid linker (GS) and then the cmyc epitope tag (EQKLISEEDL). This tag is labeled during sorts with an anti-cmyc antibody to measure display and ensure selection of full-length constructs. Inspection of the Sso7d crystal structure reveals that the C-terminus lies right at the plane of the binding paratope pointing toward where a bound target protein would be (Figure 2.16 A). If a cmyc tag was added at that location, the presence of an anti-cmyc antibody could easily obscure binding of the Sso7d to its target. In yeast a dibasic doublet like Lys-Arg is a recognition site for the KEX2 protease; thus, the functional consequence of the KQ to KR change would be introduction of a cleavage site and elimination of the cmyc tag from the displayed construct, leaving an Sso7d clone with an unobstructed C-terminus. This finding suggests that clones with a C-terminal epitope tag and associated anti-tag antibody are less fit in a contest for high affinity binding. In further support of this hypothesis is the finding that stop codons were also frequently enriched at the final amino acid position (data not shown), the result of which would be the same as for the cleavage site. Again,

these sequences appeared in multiple lineages, though they were always discarded and were not used as templates for subsequent affinity maturation libraries. It is worth noting that these observations are not specific to EGFR as both stop codons and KR motifs were observed for MSA binders[82], K-Ras binders[83] and other projects within the Wittrup lab (data not shown).

An immediate remedy for the obstruction caused by labeling the cmyc tag was to use instead an antibody against the N-terminal HA epitope tag to monitor Sso7d display during selections. This sort scheme is risky because the presence of the HA tag does not guarantee full-length display, so to mitigate this risk a single initial sort can be done after error-prone PCR library construction to select for binders that have both HA and cmyc display. Then, during binding selections the HA tag can be used without fear of propagating truncated clones. This approach was used for the final two rounds of selections for each Sso7d clone, though it is hard to know how the presence of the C-terminal obstruction during the early rounds of selections affected the overall trajectory of evolution. Alternatively, the Sso7d could be expressed from a yeast plasmid in which the orientation of the Aga and Sso7d were switched. In this orientation the C-terminus of the Sso7d would not have an antibody to compete with, though it would still followed by a linker sequence and would be closer to the Aga2 protein.

The optimal situation, however, would be one in which the C-terminus of the Sso7d were located away from the paratope interface altogether. This would allow inclusion and labeling of the cmyc tag without fear of obstructing the binding interaction. To explore the feasibility of modifying the Sso7d to enable this presentation, preliminary experiments were conducted in which the C-terminal sequence of Sso7d was truncated. Before the terminal Q61 of Sso7d are residues K60 and E59, both of which are outward-facing and charged, and accordingly not anticipated to be contributing much to structural stability. If the Sso7d sequence were terminated before E59 this

would move the new truncated C-terminus onto the back side of the scaffold; a cmyc tag in this location would likely be less disruptive. To assess structural stability in the context of C-terminal truncations, an early version of the Sso7d clone B was used as a test case. When the ΔQ, ΔKQ, ΔEKQ, or Δall (in which the entire helix, K52-Q61, was removed, expected to be destabilizing) truncations were made with the clone B paratope, FACS binding experiments showed that all constructs could be secreted and expressed as measured by the presence of both HA and cmyc tags. Truncation of up to the first three amino acids also did not disrupt EGFR binding compared to the original clone B; however, the Δall truncation was unable to bind EGFR, suggesting as expected that the expressed construct was not properly folded.

To see whether the truncations affected thermal stability, expressing yeast were heated for 5 minutes and then incubated with EGFR to measure binding. None of the truncations affected thermal stability compared to the original full-length version; only above 85°C was binding lost for any of the truncations. This assay measures only irreversible unfolding, as yeast are cooled before binding to EGFR is measured, and transient, reversible denaturation would not be detectable by flow cytometry; a proper assessment of changes to thermal stability mediated by the truncations would require a biophysical measurement such as DSC. Further experiments would certainly be warranted to verify the observations presented here, but nevertheless these assays suggest that modification or truncation of the Sso7d scaffold may be feasible, and could benefit subsequent Sso7d engineering efforts.

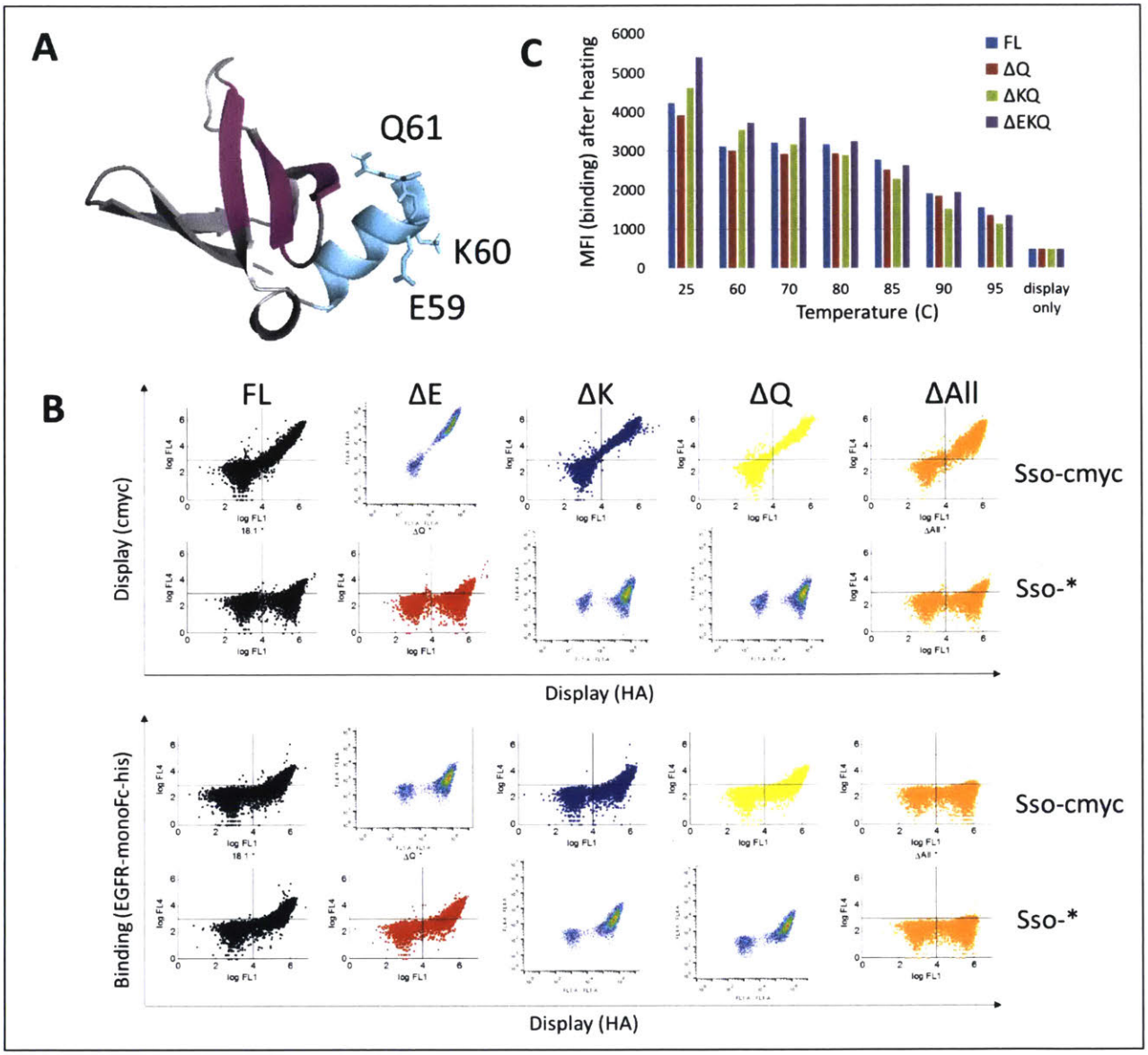


Figure 2.16 – Preliminary experiments exploring the feasibility of using a truncated Sso7d scaffold. **A** – Sso7d structure. Paratope beta sheets are colored magenta and the E, K and Q residues that were truncated are shown in sticks. Cyan indicates the C-terminal helix that was deleted in the Δall truncation. **B** – Flow cytometry to measure the EGFR-Fc binding and display of the truncated variants of an early version of the Sso7d clone B. HA epitope tag is N-terminal, cmyc tag is C-terminal. FL indicates full-length. Each truncation was made with and without the C-terminal cmyc tag. **C** – Pilot experiment of Sso7d thermal stability on yeast. Cells were heated to indicated temperatures for 5min before they were cooled to room temperature and binding to EGFR was assessed.

Epitope binning of affinity matured Sso7ds

The panel of five affinity matured Sso7ds was next assessed by cross-blocking experiments to determine the number of distinct epitopes targeted by the binders. The five Sso7d clones were binned against each other as well as cetuximab, the EGFR ligand and three EGFR-targeting Fn3 domains whose epitopes had been previously mapped; the data are shown in Figure 2.17. These experiments showed that all of the Sso7d clones could bind in the presence of cetuximab, as expected, and that all clones could self-compete, confirming the specificity of binding. However, the five clones appeared to separate into only two unique epitope bins, one containing clones B and D, the other containing A, C and E.

Clones B and D overlapped with Fn3A, known to bind EGFR in domain I; they also competed with EGF, suggesting an interaction with the region of domain I containing the secondary ligand site (this interface can be seen in the ligand-bound EGFR conformation presented in Figure 1.1). In further support of this hypothesis is the finding that point mutations in this region affected binding of clone B and binding of Fn3A (Figure 2.17). Sso7d clones A, C, and E did not compete with ligand, and in fact all seemed to bind more tightly when EGF was present. This could be the result of allosteric changes induced by ligand binding – perhaps the open conformation stabilized by the interaction with EGF exposes the Sso7d A/C/E epitope and enhances binding. These clones also did not overlap with Fn3A or with Fn3B, known to bind domain at the N-terminal end of domain III, or with Fn3D which binds the interface between domains III and IV. Their epitope remains unknown.

The binning experiment confirmed previous observations that clone A and clone B – the two Sso7ds discovered by single clone sequencing – targeted distinct epitopes. However it also showed that none of the clones discovered by the in silico approach, despite their distinct

sequences, represented additional epitope bins. Because binning is merely an indication of whether or not simultaneous binding is feasible, rather than a rigorous definition of epitope, it is still technically possible that the overlapping Sso7ds within a given bin bind distinct sites within EGFR. In support of this possibility is the finding that despite identical binning results, clones B and D behaved differently with respect to the individual point mutations. Mutations that ablated clone B binding in yeast and transfected CHO-K1 cells had no affect on clone D. Further, when clones B and D were tested against other EGFR domains expressed on yeast the binding patterns were different (data not shown). This type of behavior – overlapping epitope bins but unique epitopes – could be explained by allosteric effects, in which binding of one Sso7d clone alters EGFR conformation to disable binding of a second clone despite distal binding locations. But whether or not the cross-blocking data is the result of direct or allosteric competition, the data clearly show that with respect to simultaneous engagement of multiple Sso7ds – the parameter that matters most for multi-epitopic targeting – this panel of binders is only capable of engaging two EGFR epitopes at once.

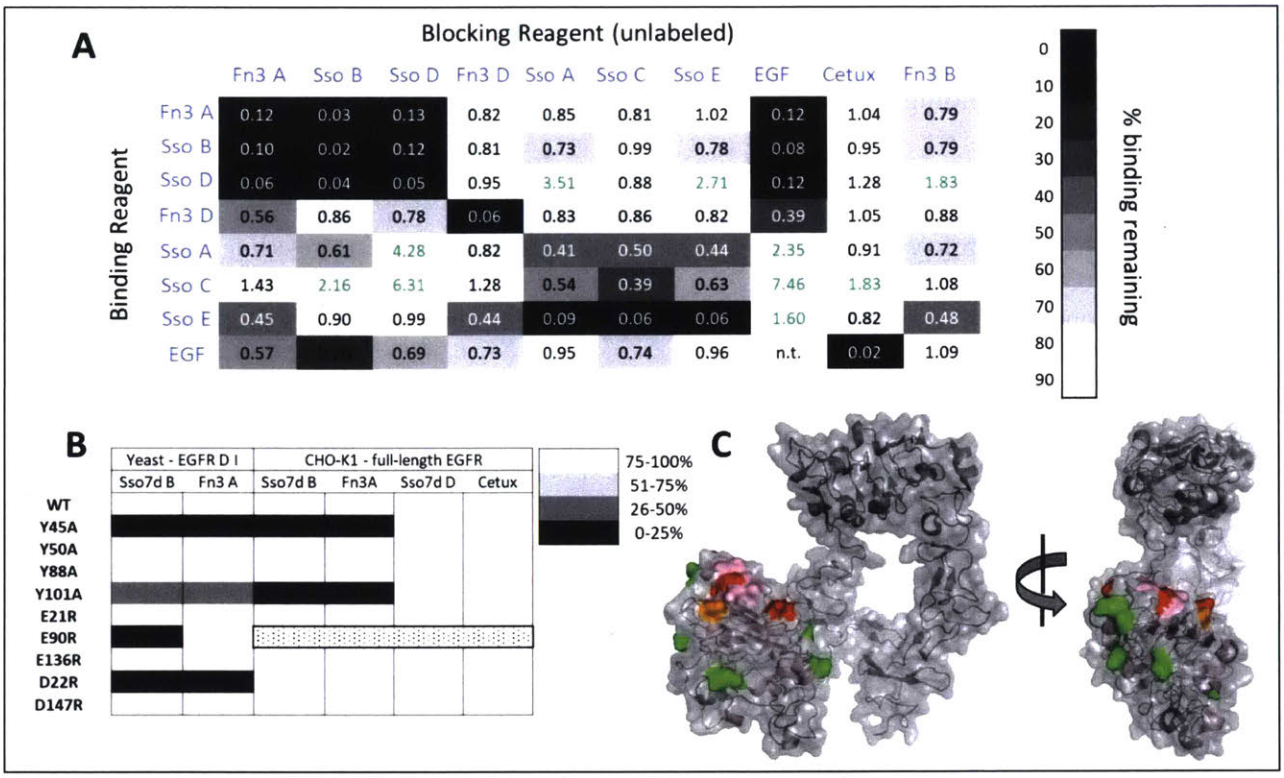


Figure 2.17 – Epitope binning for Sso7d clones, and mapping of clone B. **A** – Soluble Sso7d and Fn3 domains were expressed and purified from E. coli and a portion of each protein was biotinylated. Fluorescently-labeled EGF was purchased. The indicated labeled binding reagent was incubated with A431 cells at a concentration of 50 nM in the presence of 10-fold molar excess of indicated unlabeled blocking reagent. Binding was measured by flow cytometry using streptavidin-647 secondary and normalized to signal in the absence of blocking. Fold changes from unblocked binding are indicated by numbers and colors for each pair. Values in green denote fold changes above 1.5 and signal apparent increases in binding in the presence of competitor. **B** – Indicated point mutants were tested either in the context of EGFR domain I (residues 1-176) expressed on yeast either or as full-length EGFR transiently transfected in CHO-K1 cells. Binding of soluble Sso7d clone B or Fn3A was measured by flow cytometry and data were normalized to wild type binding. Dotted fill for E90R in CHO-K1 indicates that the mutant was not tested. **C** – Point mutations tested in B are indicated on the EGFR crystal structure. Those that did not affect binding are in green; Y45A, E90R, D22R are in red, Y101A is in orange. Other residues mapped to Fn3A are indicated in pink.

In vitro characterization of cetuximab-Sso7d fusions

The number of possible ways of combining cetuximab with any of the five engineered Sso7d binders is larger than is practical to test. Sso7ds can be added by genetic fusion to either antibody terminus on either antibody chain; at any chosen fusion site they can be added as

multimers or as single moieties; and where multimers are desired they can be comprised of clone combinations or a single one used repeatedly. Linker lengths at every junction could also be varied. All of these topological design choices would be expected to affect the geometry of EGFR engagement at the cell surface, though the impact on biochemical efficacy is impossible to predict. Because it is not practical to test the effect of every parameter, previous work with the cetuximab-Fn3 fusions was used to guide initial design choices. In these experiments, effective triepitopic fusions were built using the antibody heavy chain N-terminus and the light chain C-terminus, so this format, or slight variations thereof, was chosen for the Sso7d fusions as well (notation is XH-LX, where H and L indicate heavy and light chains of cetuximab and X corresponds to Sso7d clone lettering). All linker lengths were composed of a single GGGGS repeat.

Five Sso7d clones had been engineered for high affinity, non-cetuximab competitive binding, but the epitope binning data indicated that only two epitope bins were available. Thus, the focus of subsequent explorations in vitro was on only a subset of the Sso7d clones, A, B and D. Clone A, the original consensus clone, and clone B represented the two epitope bins, while clone D was the highest affinity binder and was in silico-derived.

Cetuximab fusions were well-expressed (between 10-15 mg/L) from HEK293 transient cultures without any optimization, a dramatic improvement over the cetuximab-Fn3 fusion. Additionally, SEC traces of Protein A purified products showed that the fusions in this format were entirely monomeric, and remained that way after storage at 4°C for a week (Figure 2.18). This data confirms the original hypothesis that this small, hyperstable scaffold conferred favorable biophysical properties in the context of a fusion protein and that, at least with respect to developability, the antibody-Sso7d fusion was preferable to the former antibody-Fn3 construct.

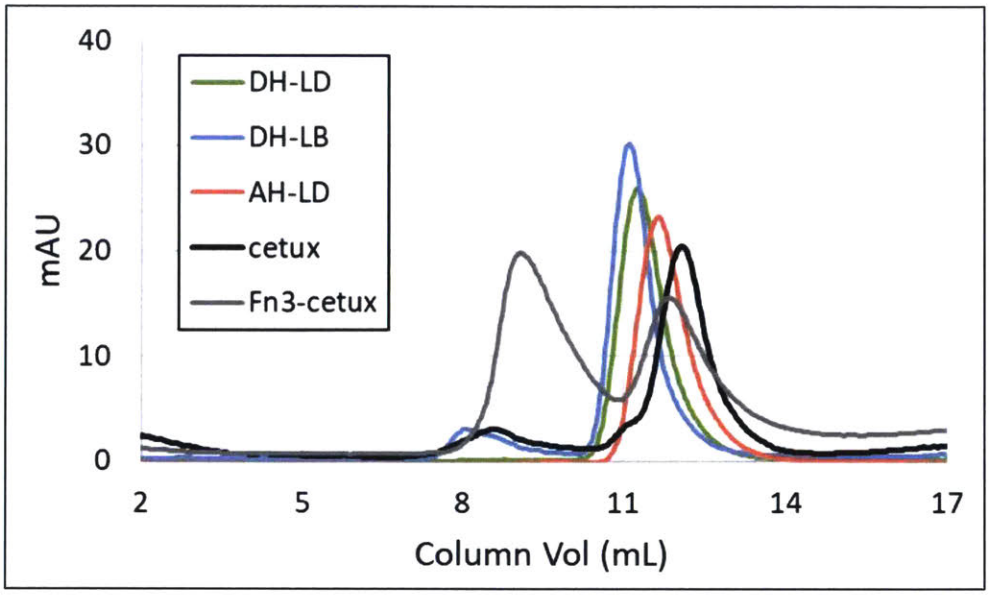
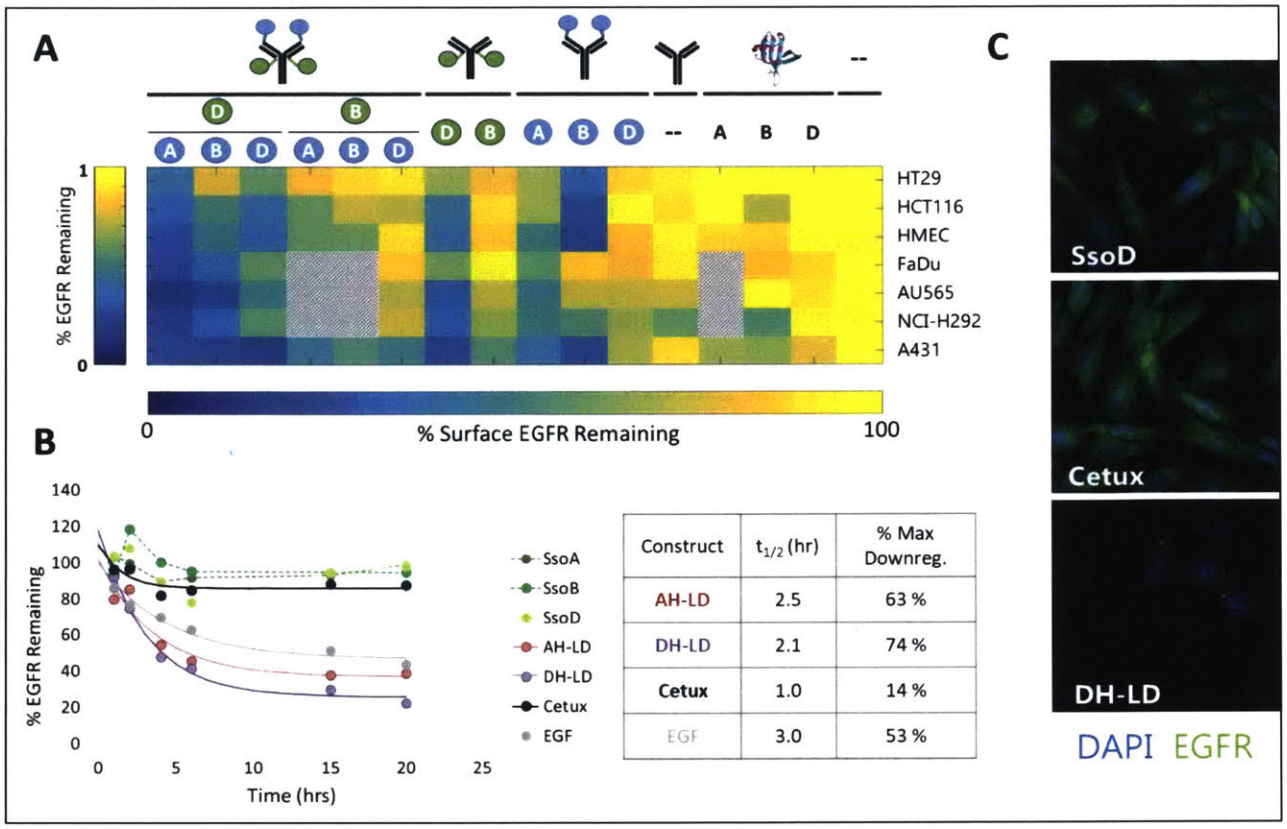


Figure 2.18 – SEC traces of cetuximab-Sso7d fusions in the XH-LX format. Proteins were produced in HEK293 Freestyle cells, then affinity purified by Protein A chromatography. Purified proteins were dialyzed into PBS and stored at 4°C for 1 week before an equal mass of each was analyzed by FPLC.

To determine whether the cetuximab-Sso7d fusions were capable of downregulating EGFR as hypothesized, constructs were incubated with EGFR-expressing cells and the amount of EGFR remaining after incubation was quantified by flow cytometry (Figure 2.19). When tested against a panel of human cancer cell lines cetuximab consistently failed to downregulate EGFR after the overnight incubation, as did monomeric Sso7d binders. The fusions were generally superior to cetuximab alone, with the most consistent and complete downregulation observed for any fusion incorporating the -LD, denoting a light chain with Sso7d D fused to the C-terminus. All three XH-LD antibodies downregulated EGFR, and interestingly, so did the H-LD, which lacks a heavy chain fusion and only has four binding paratopes.



Construct	$t_{1/2}$ (hr)	% Max Downreg.
AH-LD	2.5	63 %
DH-LD	2.1	74 %
Cetux	1.0	14 %
EGF	3.0	53 %

Figure 2.19 – Cetuximab-Sso7d fusions downregulate cell-surface EGFR. **A** – Amount of cell surface EGFR was measured by flow cytometry after a 16hr incubation with indicated construct at 20 nM. Data are normalized to MFI of untreated cells. Graphic along the top indicates protein format – from left to right: XH-LX, H-LX, XH-L, cetuximab alone, Sso7d alone, untreated. The A, B or D notation directly beneath indicates the identity of the Sso7d fusion partner. **B** – Downregulation time course was measured on A431 cells with 20 nM construct at indicated time point and compared to untreated cells. **C** – Confocal microscopy images of CHO-EG cells (CHO-K1 cells stably expressing a human EGFR-GFP fusion), incubated for 5hrs with 20nM construct. EGFR-GRP shown in green, DAPI staining in blue.

The data here also show the effects of fusion placement. H-LD and DH-L, for example, contain the same number of the same paratopes – the cetuximab antibody plus two Sso7d D moieties – and the only difference is the location of their Sso7ds. But this distinction results in strikingly different downregulation patterns, with H-LD having a strong effect and DH-L causing no change in downregulation at all. This is likely due to spatial/geometric constraints imposed by the short linker used to join the antibody and Sso7d, which, in certain formats may have prevented the simultaneous engagement of both moieties with their EGFR targets. The inverse trend of

relatively weak downregulation of -LB fusions but comparatively strong (though somewhat more variable) effect of the BH-L construct may be further evidence of this same phenomenon. Suboptimal formatting perhaps could have been improved with increased linker lengths, which would allow more flexibility between the antibody and Sso7d and could enable Sso7d engagement at more distal epitopes while still permitting cetuximab binding in domain III. Nevertheless, the data clearly confirm the hypothesis that Sso7ds can be used for the purpose of multipepitopic targeting, and that when the fusions are optimally placed a cetuximab-Sso7d fusion can nearly completely downregulate the cell surface EGFR from a variety of tumor cell lines.

Kinetics of downregulation were measured for AH-LD and DH-LD and these were compared to the Sso7ds alone, cetuximab, and the EGF ligand, the latter known to rapidly induce internalization and downregulation as a mechanism of negative signaling feedback (Figure 2.19). Half-lives of internalization were comparable for the two fusions, and both were faster than the ligand; as expected, the antibody and Sso7d monomers had no effect over the time course studied.

Confocal microscopy also showed the dramatic differences in EGFR cellular distribution when constructs were added to CHO cells stably transfected with human EGFR-GFP (Figure 2.19). After a 5hr incubation with constructs, diffuse staining of the EGFR was seen for both the Sso7d clone D and cetuximab alone, while the addition of DH-LD led to a dimmer, more punctate presentation indicative of EGFR internalization.

To determine the effect of fusions on ligand binding to EGFR and subsequent intracellular signaling, three points in the signaling pathway were interrogated. First, EGF binding at the cell surface was measured in the presence of cetuximab, Sso7ds or indicated fusions (Figure 2.20). The experiment was done at 4°C to minimize internalization and limit the analysis to effects of competition at the cell surface. When monomeric Sso7d clones A, B or D were added as ligand

competitors the results confirmed what was previously observed in Figure 2.17. Clones B and D inhibited binding, though the IC_{50} for this was weaker than cetuximab alone, likely because of the lack of avidity and lower monovalent affinity of the Sso7ds for EGFR, while clone A led to tighter EGF binding, an affect previously attributed to allostery. At 30nM no differences between cetuximab and the antibody fusions could be discerned – ligand binding was still undetectable at the highest end of the titration curve – but when the competitor concentration was lowered by an order of magnitude to 3nM the improved inhibition of the fusions was apparent. Both AH-LD and DH-LD inhibited ligand binding more potently than cetuximab, with DH-LD being the strongest competitor. This effect is not due to a reduced number of EGFRs on the surface since the experiment was carried out at 4°C, so the differences are solely attributable to the effects of increased avidity and complementary ligand blocking by the Sso7ds.

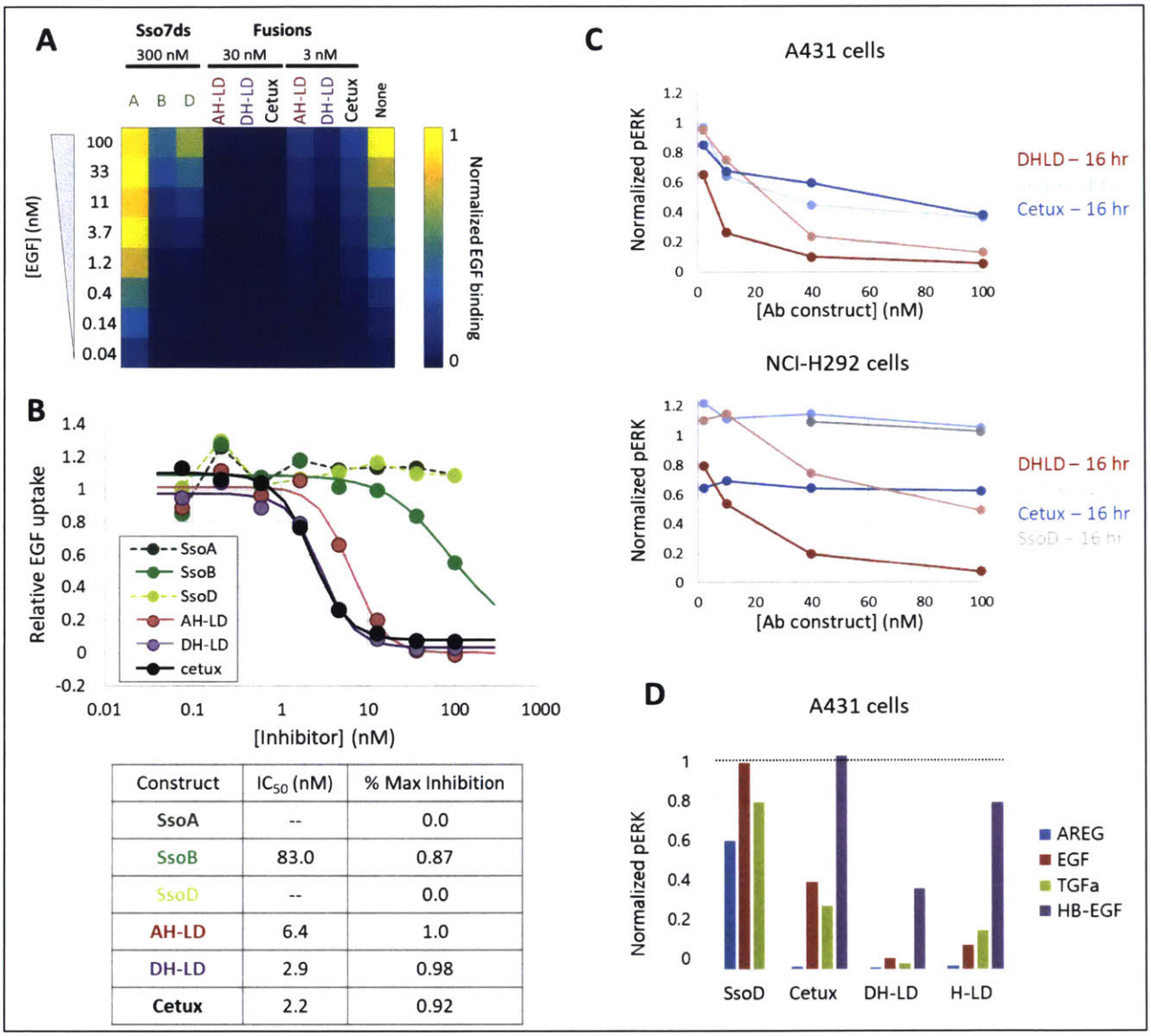
The effect of the fusions on ligand internalization, the subsequent step in the EGFR signaling pathway, was next assessed. Constructs were added to A431 cells for 1 hour before addition of ligand, and after a short incubation the amount of EGF was measured by flow cytometry. To ensure that measured signal was only representative of internalized rather than surface bound ligand, the EGF used in the experiment was labeled with a pH-sensitive fluorophore so that signal was conditional upon cellular uptake. In line with the ability of the fusions to block ligand binding at the cell surface, they also prevented EGF uptake. DH-LD was again superior to AH-LD, but the IC_{50} was not better than cetuximab alone.

However, when phosphorylation of ERK, a downstream effect of EGFR within the MAP kinase signaling pathway, was measured there was a clear difference between cetuximab and DH-LD. Cetuximab had a dose-dependent effect on signaling which reached ~60% inhibition at the highest antibody concentration tested. This curve was unchanged whether antibody was incubated

for 1 hour or 16 hours before stimulation. DH-LD however, was much more potent than cetuximab even after only 1 hour and, at high concentrations, could nearly complete shut off EGFR signaling. Additionally, likely because of downregulation, the inhibitory effect was even more pronounced with increased incubation time. On NCI-H292 cells DH-LD was similarly superior, demonstrating, as expected, that the cetuximab-Sso7d fusions are active against multiple tumor types.

Finally, the fusions were tested to evaluate if the superior signal antagonism resulted in expected phenotypic changes such as altered growth and expression of immune receptors thought to be linked to EGFR signaling (Figure 2.21). In both proliferation assays in A431 cells as well as colonogenic assays in A431 and NCI-H292 cells, DH-LD was superior to cetuximab at limiting cellular growth. Additionally, on three different EGFR-expressing cell lines, both fusions had a more profound impact than cetuximab on the expression of immunosuppressive PD-L1 induced by EGFR signaling. Consistent with other in vitro data, DH-LD was the most efficacious in this experiment, and drove the amount of PD-L1 expression below even the basal, unstimulated levels.

Taken together all of these in vitro results suggest that the DH-LD cetuximab-Sso7d fusion should have superior anti-tumor effects in vivo compared to cetuximab. Similar to the previously engineered antibody-Fn3 fusion, the mechanism of action of DH-LD is a combination of rapid receptor downregulation and potent signaling blockade. This antagonism results in the expected anti-proliferative effect on tumor cells directly and also, excitingly, appears to potentiate the immune-related effects of EGFR inhibition suggesting synergistic combinations with other immunotherapies might be possible.



Construct	IC_{50} (nM)	% Max Inhibition
SsoA	--	0.0
SsoB	83.0	0.87
SsoD	--	0.0
AH-LD	6.4	1.0
DH-LD	2.9	0.98
Cetux	2.2	0.92

Figure 2.20 – Superior EGFR signaling antagonism by cetuximab-Sso7d fusions. **A** – Inhibition of EGF-FITC binding to A431 cells was measured in the presence of either Sso7d or antibody fusion. Cells were incubated with competitor at concentration noted along the top, with EGF-FITC at concentration indicated along left side. Binding was carried out at 4C for 1hr then cells were read washed and read by flow cytometry. Data are normalized to maximum EGF binding (100nM) in the absence of competitor. **B** – EGF uptake was measured by incubating A431 cells with conditionally-fluorescent pHrodo Green EGF in the presence of competitor at indicated concentration. Data (circles) were fit (solid lines) assuming a single binding site and direct (non-allosteric) competition; dashed lines indicate no fit was calculated. IC_{50} values and max inhibition from the fit are shown in the table. **C** – Phosphorylated ERK was measured by ELISA on A431 and NCI-H292 cells by incubating constructs with serum-starved cells for indicated times, then stimulating signaling with 5nM EGF for 5 min. Cells were then washed and lysed, and lysates were used in the ELISA. Data are normalized to cells stimulated in the absence of competition. **D** – Inhibition of signaling was measured by pERK ELISA as in C, but stimulation was done with a variety of EGFR ligands (concentrations listed in Methods section).

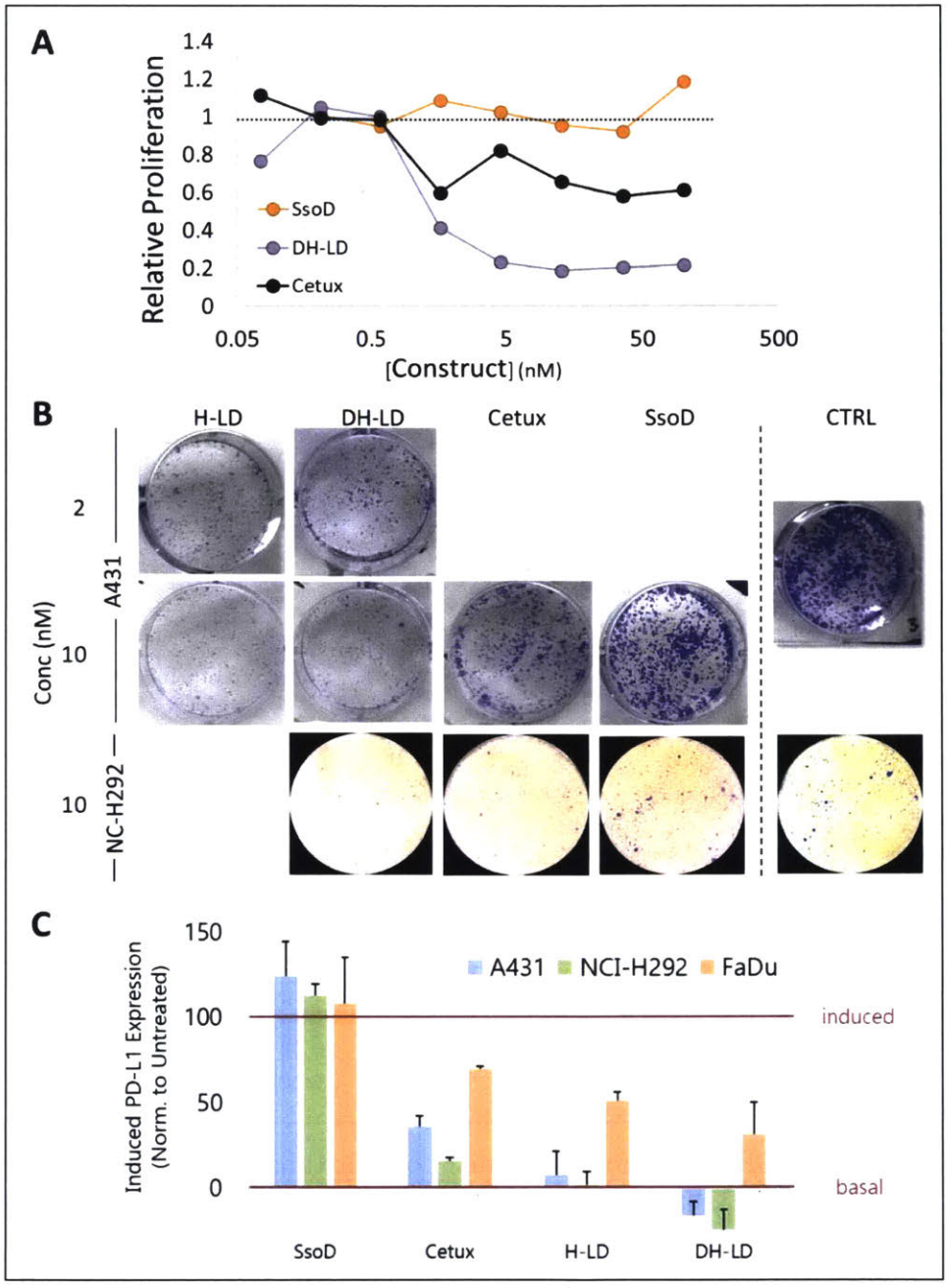


Figure 2.21 – Phenotypic responses of tumor cells antagonized by cetuximab-Sso7d fusions. **A** – A431 cells were plated in 96-well plates at 5,000 cells/well and incubated with indicated constructs at various concentrations for 5 days. Cell growth was quantified by adding XTT reagent and reading absorbance at 592nm. **B** – Cells were plated at 1,000 cells/well in a 6-well plate. Constructs were added and cells were incubated for 7 days before crystal violet staining and imaging. **C** – Effect of the fusions on PD-L1 expression induced by EGF was measured. Constructs were added to serum starved cells, and after an overnight incubation EGF was added for an additional 12-16 hours. Induced PD-L1 expression was measured using an anti-PD-L1-APC antibody and reading by flow cytometry. MFI were normalized to controls that were un-induced and lacked inhibitor ("basal") and induced and lacked inhibitor ("induced").

In vivo efficacy of cetuximab-Sso7d fusions

Based on the promising in vitro results of DH-LD and H-LD these fusions were tested for their ability to control tumor growth in mice. The model tumor system chosen was a patient-derived glioblastoma xenograft (PDX) expressing amplified wt EGFR. A tumor was established in the flank of athymic nude mice and the efficacy of cetuximab and the two fusions were compared. Doses were chosen with the goal of saturating EGFR within the tumor, since both downregulation and signal antagonism and were hypothesized to be the driving force for anti-tumor efficacy. A simple mathematical ODE model was used to estimate the requisite dosing (Figure 2.22 A). This model was based on previous work[84–86] and incorporated both transport between blood and tumor as well as the effects of target mediated clearance; this process was expected to be particularly important for the pharmacokinetics of the antibody-Sso7d fusions which downregulate EGFR much more rapidly than cetuximab alone. Based on the model predictions a 200ug dose given twice weekly was chosen as the treatment schedule. Under these conditions EGFR was predicted to be saturated within the tumor after a single dose.

However, tumor growth curves for the fusions showed that they did not provide superior tumor growth control compared to cetuximab (Figure 2.22 B). On the contrary, the efficacy of the fusions was significantly worse than the antibody alone, and only barely represented an improvement over the PBS control group. In fact, the trend appeared to be that increasing the number of Sso7d moieties from zero on the antibody to two on H-LD to four on DH-LD corresponded with increasingly worse tumor control. None of the groups were toxic as measured by weight changes.

This result was paradoxical given the overwhelming in vitro evidence that the fusions were superior EGFR antagonists to cetuximab. At worst, even if the Sso7d moieties provided no benefit

in vivo, the expectation would be that the fusion would behave comparably to cetuximab, yet the PDX tumor study clearly indicated otherwise. To reconcile these contradictory findings we questioned whether pharmacokinetic factors could be playing a role. Because cetuximab and all of the antibody fusions contained a human IgG1 Fc, their blood concentrations in mice were straightforward to measure using a human IgG ELISA (Figure 2.22 C). This assay revealed that serum concentrations of DH-LD were dramatically lower than those of cetuximab when tested 48 hours after either the first dose or after the sixth. In both cases they were barely detectable, and were roughly an order of magnitude lower than for the antibody alone. For H-LD the concentration after a single dose was similar to cetuximab, but after a few weeks of treatment the H-LD concentration was somewhere between cetuximab and DH-LD.

Differences in PK profiles between cetuximab and the fusions were anticipated because of the increased downregulation mediated by the fusions, but target-mediated disposition to this extent was unexpected. The model, while simplistic, had been designed to account for the consumption of antibody due to downregulation, and the observed concentration was still well below the prediction for the 48hr time point. To test whether the model had simply been poorly parameterized and verify that consumption within the tumor accounted for the loss of the fusions from circulation, drug concentrations were also measured in mice lacking tumors (Figure 2.22 D). Since the tumors are the only source of human EGFR in a xenograft model, there should be no measurable downregulation in this context and the PK of cetuximab and the fusions should be the same. A 200ug dose of either cetuximab or DH-LD was given to nude mice and serum was collected 48 hours later. When IgG concentrations were measured by ELISA for these tumor naïve mice the finding was the same – DH-LD concentration was reduced to approximately 10% of

cetuximab. This clearly indicates that target-mediated drug disposition by the tumor does not explain the inferior efficacy of DH-LD and H-LD in the PDX model.

An alternative explanation for reduced half-life of the fusions could be immunogenicity of the Sso7ds leading to systemic clearance of H-LD and DH-LD. This explanation could account for the observation that after a single dose the H-LD concentration is roughly equivalent to cetuximab but after prolonged treatment the H-LD concentration has actually decreased. However, an immune response seems unlikely to explain the dramatic reduction of DH-LD at just 48 hours after a single dose, since antibody clearance mediated by an adaptive immune response does not happen on that time scale.

As a final hypothesis, the cross-reactivity for mouse EGFR was questioned. The modeling and in vivo experiments were carried out under the assumption that the only relevant EGFR expression was occurring in tumor, despite the fact that mice, like humans, express EGFR in a variety of tissues. Cetuximab is known to not bind measurably to mouse EGFR but the cross-reactivity of the DH-LD fusion had not been explicitly tested. Early experiments on yeast, in which Sso7d clone D was expressed on the cell surface and binding to soluble murine EGFR was interrogated, showed weak or no measurable interaction. But when the DH-LD fusion and a bivalent Sso7d-D-Fc were measured against a mouse EGFR-expressing cell line binding was more pronounced (Figure 2.23). What appeared to be insignificant Sso7d binding to mouse EGFR as a monomer was much more substantial in the highly avid format of the antibody fusion, and there is even a clear difference in cell binding between a construct with two Sso7ds and one with four.

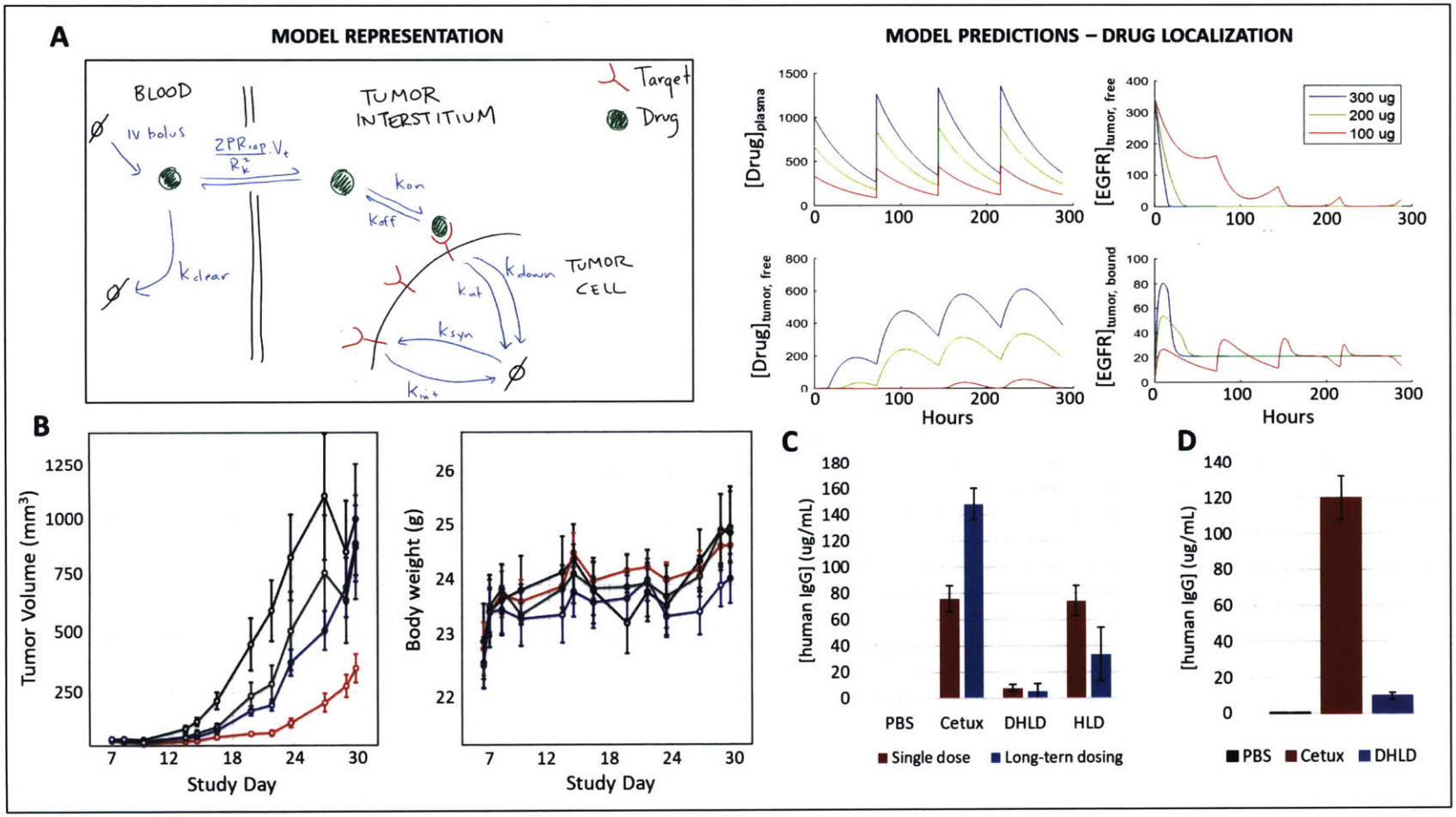


Figure 2.22 – In vivo efficacy study of cetuximab and cetuximab-Sso7d fusion s in human PDX model. **A** – Schematic representation and results from ODE model used to guide dosing decisions. Three doses were simulated, each administered every 72 hours. Model output includes concentrations of drug in plasma, drug free (unbound by EGFR) in the tumor, and both free and bound EGFR concentrations within the tumor. All concentrations are in nM. **B** – Tumor growth curves and mouse weights for efficacy study in PDX model. Doses were 200ug each, administered retro-orbitally twice weekly. Black: PBS, Green: DH-LD, Blue: H-LD, Red: cetuximab. **C** – Concentrations of antibody or fusion present in serum 48 hours after either the first dose or the sixth dose from the PDX efficacy study, as measured by ELISA. **D** – Serum concentrations of cetuximab or DH-LD 48 hours after a single 200 ug dose in tumor naïve nude mice.

Overall this murine cross-reactivity provides an explanation that is consistent with both the tumor-bearing and tumor naïve PK as well as the efficacy results – that endogenous EGFR expression within the mouse acts as an additional and substantial sink that is only a factor for the antibody-Sso7d fusion and not for cetuximab alone. This on-target but off-tumor uptake leads to dramatically less antibody fusion in circulation than cetuximab, which explains why the fusions were inferior (as opposed to just not superior) to cetuximab in the PDX efficacy study.

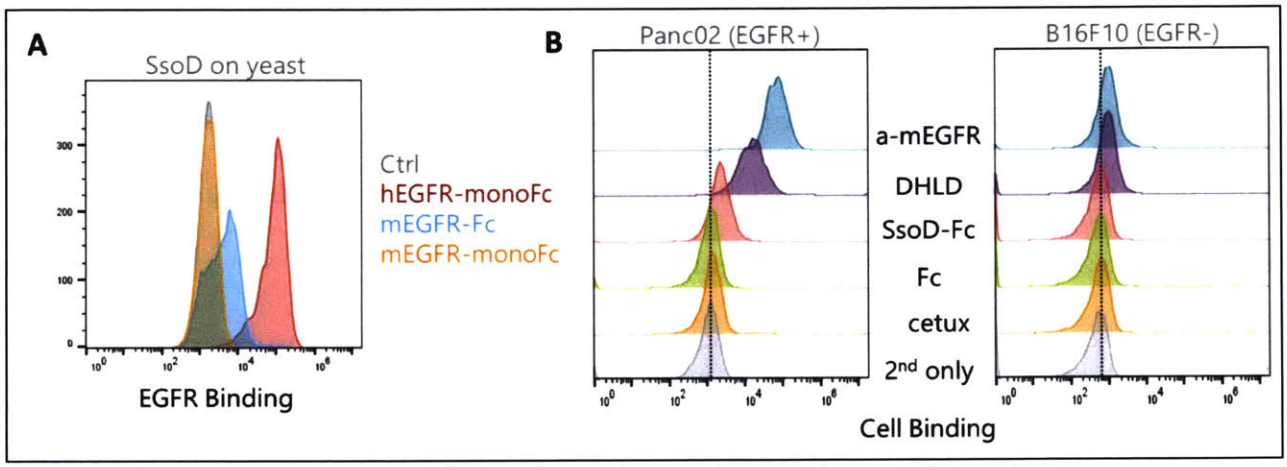


Figure 2.23 – EGFR cross-reactivity of Sso7d clone D. **A** – Human EGFR (hEGFR) or mouse EGFR (mEGFR) in monovalent or bivalent formats were incubated with yeast expressing Sso7d clone D for 1hr at room temperature. Cells were washed then incubated with anti-human-AF488 antibody for 30min at 4°C before washing and reading binding by flow cytometry. **B** – Indicated constructs, at 20nM, were incubated with trypsinized and washed Panc02 cells (known to express murine EGFR) or B16F10 cells (EGFR null) for 1hr at room temperature. Cells were washed, incubated with anti-human-AF488 antibody for 30min at 4°C, then washed again and analyzed by flow cytometry.

Discussion

The objective of this work was to develop an improved EGFR antagonist which had superior anti-tumor activity in vivo. The driving hypothesis was that a molecule capable of simultaneous receptor downregulation and ligand competition would outperform clinically-approved antibodies such as cetuximab. Work within and outside of the Wittrup lab lent support to this approach – downregulation of receptors has been documented using a variety of targeting approaches[69,70,74,87–91]. The new embodiments designed here were intended to advance the multi-epitopic approach by incorporating EGFR binding moieties with favorable biophysical properties. The hope was that this would increase stability, decrease aggregation, and overall improve the developability of the novel therapeutic.

By all in vitro measures, this approach was successful. Sso7ds were isolated and evolved to bind EGFR with moderate to high affinity even in the presence of cetuximab, and fusions of these binders with cetuximab enabled the receptor clustering and downregulation intended. Superior signaling antagonism resulted in reduced cellular proliferation and colony formation in multiple cell lines. Yet in mice in a human PDX model the most promising fusion, DH-LD, was not just ineffective but was inferior to cetuximab alone. This was ultimately attributed to the cross-reactivity of Sso7d clone D. Measurements of its monomeric affinity for murine EGFR showed weak to no binding, but a highly avid antibody fusion format enabled on-target but off-tumor uptake of the fusion, leading to dramatically shorter half-life and faster clearance.

One possibility for overcoming a poor PK profile and restoring therapeutic efficacy is simply to alter the dosing schedule, by increasing the amount of drug given per dose and/or increasing the frequency of administration. But the feasibility of this approach for the particular case of EGFR antagonism is questionable. For one, EGFR is widely expressed on a number of

normal tissues, particularly those of epithelial origin. For therapy administered systemically the dose needs to be high enough to saturate all of the endogenous, normal EGFR in order to ensure adequate tumor uptake. Presumably such a saturating dose exists, since this same phenomenon is encountered and apparently overcome for cetuximab administration in humans in the course of clinical practice, but the difference here is the added turnover as a result of downregulation by the fusions. The combination of a high number of non-tumor target sites with increased degradation rate at each of those sites makes the prospect of saturating dosing daunting. Further, even if a dose regimen existed to ensure adequate tumor uptake of the fusions, dose-limiting toxicity would be a serious concern. Cetuximab is known to cause severe skin rashes and other toxicities in humans, so using a higher dose of a more potent antagonist could exacerbate pre-existing toxicity concerns as much as it would help with anti-tumor efficacy.

A final alternative for overcoming limited tumor uptake would be to consider intratumoral dosing. This approach has gained traction and seen increased clinical exploration over the last few years, particularly within the context of cancer immunotherapy where local administration of a drug can be used to prime a systemic immune response[92]. However, since the mechanism of action of the cetuximab-Sso7d fusion is local blockade of growth signals directly at the tumor site, the therapy could only succeed if every cancer cell were exposed to drug. For metastatic disease in particular this could limit efficacy since it would be impossible to locally administer the fusion to every metastasis. Further, the need for repeat dosing would present a substantial challenge. Thus, this approach is also not suitable for overcoming the poor pharmacokinetic profile of the antibody-Sso7d fusions.

Because of the anticipated dosing/toxicity challenges, ultimately the idea of using the cetuximab-Sso7d fusion as an improved therapeutic was abandoned. Instead this work remains

as another demonstration that downregulation is an effective means to a biochemical end – receptor antagonism is improved through cell-surface clustering and reduced recycling. It is notable, however, that no clinical products have been developed that utilize this mechanism of action. Perhaps the approach would hold more promise for a target not as ubiquitously expressed as EGFR.

Nevertheless, the process of identifying and optimizing the Sso7ds used in this study provided an opportunity to further our understanding of the Sso7d as an engineerable scaffold. Numerous reports have definitively demonstrated that an Sso7d library can produce a binder to a diverse array of targets[81–83,93], just as has been shown for other small protein scaffolds. But this work pushed the boundaries of Sso7d selections by requiring not just any binder to EGFR, but by seeking as many unique clones to as many unique epitopes as possible. This process was surprisingly challenging – in the absence of next-gen sequencing only two sequence-distinct clone families were discovered. The application of deep sequencing and accompanying in silico techniques increased the number of unique clone families, but not the number of epitopes targeted.

One explanation for the limited epitopic coverage achieved here could be the dominance of epitopic "hot spots" on EGFR. It has long been understood that binding free energy is not evenly distributed across a protein's surface and that certain regions are more favored for binding because of enrichment in particular amino acids[94,95]. This phenomenon can affect naïve library selections as has been observed for antibodies as well as phage-selected peptides[96–102]. The finding that approximately 75% of the "all binders" library targeted the ligand/cetuximab epitope of EGFR, and that a staggering 90% of the remaining non-cetuximab competitive population interacted with a single distinct epitope means that greater than 95% of the Sso7d naïve library

covered only two epitopes on EGFR. This certainly supports the idea that hot spots exist on the EGFR surface and that their presence complicated the selection of Sso7d binders for this study.

Alternatively, or perhaps in addition to the above explanation, is the possibility that intrinsic limitations of the Sso7d scaffold contributed to the poor epitopic coverage observed. Among EGFR-targeting antibodies, nanobodies, DARPins and fibronectins, several distinct epitopes have been accounted for without particular effort to broaden coverage[74,87,103–105]. But the Sso7d paratope differs from these scaffolds in terms of its total size – only 9 amino acids form the paratope, as well as its structure – the paratope residues are non-contiguous amino acids distributed among three beta sheets as opposed to a string of amino acids comprising one or more flexible loops. These differences may be functionally important. For example, one analysis of antibody-antigen interfaces revealed that antibodies utilize an average of 19 contact residues to mediate antigen binding[106], a larger number than is even available on the Sso7d paratope. A crystal structure of an Sso7d in complex with K-Ras showed only nine amino acids from the Sso7d formed the contacts with its target[83]. Another study which analyzed antibody-antigen interactions concluded that specificity and affinity are derived predominantly from hydrophobic side chains dispersed among short-chain hydrophilic residues which help the paratope adopt necessary conformations to complement an antigen surface[107]. While the Sso7d binders clearly enrich for hydrophobic side chains, the paratope lacks the context of surrounding short-chain hydrophilic residues. There are amino acids interspersed between the nine designated paratope residues, but these are likely required for core packing and proper folding and do not provide much conformational flexibility. Indeed, the fact that Sso7d affinity maturation by directed evolution frequently resulted in mutations adjacent to or between paratope residues suggests that shifts to the paratope structure were required to enhance antigen interactions.

Further, the results here revealed limitations with respect to Sso7d affinity maturation. While slightly sub-nanomolar affinities have been reported[93] and were achieved for one of the lineages in this study, typical Sso7d affinities, even after directed evolution, are low- to mid-nanomolar or higher[82,83]. Successive efforts toward higher affinity in this study produced only marginal gains once K_d values reaced the mid-nanomolar range, and the best single round gains required a relatively high number of mutations, suggesting that achieving high affinities is difficult because only specific combinations of mutations can provide better binding. This is in contrast to the typical evolution of antibodies and other scaffolds where picomolar affinities are regularly achieved and even femtomolar K_d values have been reported[76,108,109]. Overall these distinct properties of the Sso7d scaffold suggest that it may not be the best suited for applications when broad antigenic coverage, binding to a particular epitope, or unusually high affinity binding are requirements.

However, one benefit of the Sso7d scaffold library was its amenability to in silico discovery methods. Deep sequencing has been utilized in the context of antibody libraries for a variety of applications, including design validation and diversity estimation for naïve libraries[110,111], estimation of clonal affinity following selections[112], and epitope mapping[113–115]. The technology has even been previously applied for the discovery of rare clones[116,117] in a manner similar to the one employed here. But applying NGS to antibodies for the discovery and analysis of clone families is challenging for two reasons. The first is the length of the antibody, which when typically formatted as an scFv with 750-800bp exceeds the read length limit of current NGS technologies; and the second is the complexity of the paratope, which typically can't be fully defined without considering multiple CDR sequences of often varying lengths. To

overcome these difficulties computational approaches are typically limited to a reduced region of the protein, such as the CDR3 sequences, at the risk of lost information.

By comparison, Sso7d analysis is straightforward. The entire gene is approximately 200bp and can be sequenced without even the need to use paired end reads, and the paratope is defined by just a fixed nine amino acid sequence. Even the relatively intuitive and simple hierarchical clustering approach applied here revealed information missing from the traditional single clone sequencing approach. Further refinements to this relatively unsophisticated computational approach would likely yield additional value.

Beyond the theoretical exploration of inherent Sso7d capability, there were also practical lessons which could inform future use or improved design of Sso7d libraries. The most impactful of these was the repeated finding that affinity maturation selections produced clones in which the C-terminal cmyc epitope tag was mutated away, through either introduction of a dibasic KR motif or a stop codon. Initial experiments were performed to explore the feasibility of creating a library in which the Sso7d C-terminus was truncated, in order to move the C-terminus away from the binding paratope interface. Though further experiments would be needed to definitively assess the value of this approach, these preliminary data suggest such a modification might be both possible and valuable.

Methods

Reagents and protein production

Chicken anti-cmyc and chicken anti-HA antibodies for yeast display were purchased from Exalpha Biologics, Inc. Mouse anti-HA antibody (clone 16B12) was purchased from Abcam. Goat anti-chicken IgY-AF488 antibody, streptavidin-AF647, goat anti-human-AF647 and goat anti-mouse-AF647 antibodies were purchased from ThermoFisher Scientific. Human EGF and HB-EGF were purchased from Peprotech, human AREG and TGFα were purchased from R&D Systems. pHrodo Green EGF was purchased from ThermoFisher Scientific.

EGFR-Fc and EGFR-monoFc consisted of the EGFR ECD fused to either wild type human IgG1 Fc or an engineered version in which a glycosylation site is introduced to stabilize monomeric solubility[118] (for sequences see Appendix A). EGFR-Fc, EGFR-monoFc, cetuximab, and cetuximab-Sso7d fusions were all cloned into the gWiz expression vector (Genlantis) using standard cloning techniques. Protein was produced by transient transfection in the HEK293 Freestyle system (ThermoFisher Scientific). Transfections were performed using 1mg plasmid per 1L of culture with cells at a density of 0.5E6/mL; plasmid was mixed with PEI at a ratio of 3:1 (w/w) PEI:DNA in OptiPro SFM media and incubated for 15min at room temperature before addition dropwise to cultures. Transfected cultures were harvested after 7 days, centrifuged and the supernatant was filtered and stored at 4°C until purification. All Fc-containing proteins were purified using rProtein A Sepharose FastFlow resin (GE Healthcare), dialyzed into PBS and stored at -80°C. Sso7d monomers were produced as SUMO fusions and with an N-terminal His tag using the pE-SUMO vector (LifeSensors) as previously described[82]. Proteins concentrations were measured by A280 using a NanoDrop 2000 Spectrophotometer (ThermoFisher Scientific).

Yeast display library construction and selections

Yeast display experiments were generally performed as described previously[119,120]. Construction of Sso11 and Sso18 libraries as well as bead and FACS selections through the "all binders" stage were previously described[82]. For competitive selections cetuximab was pre-incubated with EGFR-Fc or EGFR-monoFc for 30 minutes at room temperature before yeast were added. Concentrations of cetuximab were always at least 5-fold higher than the EGFR antigen, and always at a minimum concentration of 50nM. All error-prone PCR libraries were built according to protocols previously described[120], using primers that ended just outside the Sso7d coding sequence and included ~40bp of overlap with the pCTCon2 backbone plasmid. Final library sizes were always >1E8 transformants. Sso7d display during sorts was monitored by using either chicken anti-cmyc or chicken anti-HA antibody as the primary antibody and goat anti-chicken IgY-AF488 as the secondary. In binding experiments using simultaneous anti-cmyc and anti-HA detection, the anti-HA antibody was mouse anti-HA clone 16B12 with a goat anti-mouse-AF647 secondary antibody.

For sequencing, library DNA was isolated using the ZymoPrep Yeast Plasmid Miniprep II Kit (Zymo Research). For single clone analysis, yeast DNA was transformed into Stellar competent cells (Clontech), plated, and single colonies were picked for sequencing. All sorts were done on a BD FACS Aria (BD Biosciences).

Sso7d heat denaturation experiments on yeast

Yeast expressing the indicated Sso7d construct were heated to various temperatures and incubated for 5min. Cells were immediately cooled to 4°C and then binding was assessed as above for yeast binding experiments, using 20nM EGFR-Fc as the binding antigen and the anti-cmyc or anti-HA antibody for display. Binding was measured by flow cytometry.

High throughput sequencing of yeast libraries and in silico analysis

Yeast miniprepped DNA was prepared for high-throughput sequencing by amplifying genes with primers containing distinct barcodes for each of the library populations (S11 before sort; S11 after sort; S18 before sort; S18 after sort). Sequencing was performed using the Illumina MiSeq system. In silico sequence analysis and clone family clustering was performed using Matlab (MathWorks) and commands from the Bioinformatics Toolbox.

Cell culture

NCI-H292, A431, FaDu, AU565, HT29, HCT116 and HMEC cells were all purchased from ATCC (Manassas, VA) and were cultured in media according to ATCC recommendations. All media and supplements were purchased from ATCC unless noted. CHO-EG cells, stably expressing EGFR-GFP, were established as previously described[121] and maintained in media supplemented with 0.3mM geneticin (ThermoFisher Scientific). All culture media was supplemented with 10% FBS (ThermoFisher Scientific). 0.25% trypsin-EDTA was purchased from ThermoFisher Scientific.

Microscopy

CHO-EG cells were plated on #1 glass cover slips (Chemglass) and incubated with indicated antibody or fusion constructs. After incubation for 5hr cells were fixed with 4% paraformaldehyde for 10 minutes at room temperature and cover slips mounted with DAPI-containing mounting medium (Vectashield, Vector Laboratories) and dried overnight. Images were acquired at room temperature using a GE (Applied Precision) DeltaVision Spectris inverted Olympus X71 microscope with a 60x objective lens, captured with a Photometrics CoolSNAP HQ camera. SoftWoRx software was used for image acquisition and deconvolution.

EGFR downregulation assays

Cells were seeded in 96-well plates in complete medium and incubated overnight. Adherent cells were then serum-starved overnight before constructs were added at indicated concentrations and incubated at 37°C for indicated times. For quantification, cells were trypsinized, washed in PBS+0.1% BSA (PBSA), then incubated with 100nM cetuximab to saturate available EGFR binding sites. Incubation was done for 1hr at 4°C to prevent further internalization or downregulation. Cells were then washed and incubated with goat anti-human-AF647 antibody at 1:1000 dilution in PBSA for 30 min at 4°C. Some published versions of this experiment use a brief acid wash to strip remaining construct from the cell surface before quantifying the amount of receptor, but in this case the use of cetuximab as the detecting reagent enabled quantification without any stripping. Any free EGFR or Sso7d-bound EGFR should be labeled by the added cetuximab (since none of the Sso7ds prevent concurrent cetuximab binding), and both this and any previously cetuximab-bound EGFR would be labeled by the anti-human secondary antibody. Finally, cells were washed in PBSA and resuspended in PBS immediately before reading on an Accuri C6 flow cytometer (BD Biosciences).

EGF binding and uptake assays

Cells were seeded in 96-well plates in complete medium and incubated overnight. For EGF binding competition assays, EGF-FITC was added to wells at desired concentrations and incubated for 2hr at 4°C to prevent internalization caused by ligand binding. Cells were then trypsinized, washed and binding was measured by flow cytometry.

For EGF uptake assays adherent cells were serum-starved overnight before constructs were added and returned to 37°C for indicated times. pHrodo EGF, which only fluoresces at the acidic

pH that occurs during internalization, was then added to the media for 2 hours. For quantification, cells were trypsinized, washed with cold PBSA, and read by flow cytometry.

pERK ELISAs

For measurement of phosphorylated ERK, A431 or NCI-H292 cells were plated in 12-well plates in complete medium and allowed to adhere. After serum starvation, blocking constructs were added as indicated. Signaling was stimulated by the addition of 8nM EGF, 5nM HB-EGF, 10nM TGFα or 15nM AREG to the media for 5 min at 37°C. Cells were then immediately placed on ice, washed with cold PBS and 200uL of RIPA lysis buffer (ThermoFisher Scientific) supplemented with cOmplete™ protease inhibitors (Millipore Sigma, St. Louis, MO) was added. Plates were incubated for 20min at 4°C to allow complete lysis. Lysate was collected, spun at 14,000xg for 10min and the supernatant transferred to a clean tube. Protein concentration was quantified by BCA assay. pERK was measured using the SimpleStep ERK 1/2 (pT202/Y204 + Total) ELISA Kit (Abcam) according to manufacturer's instructions.

Growth inhibition assays

For proliferation assays cells were plated at 5,000 cells / well in a 96-well plate in complete medium and constructs were added at indicated concentrations. Wells were incubated for 5 days at 37°C, then quantified by adding 50uL of prepared XTT reagent to wells, incubating for 1.5hrs and reading absorbance at 592nm.

For colonogenic assays, cells were seeded at 1,000 cells per well in complete medium in a 6-well plate. After an night to allow cells to adhere to plates, constructs were added to wells at desired concentrations and plates were incubated for one week. Wells were quantified by washing with PBS and staining 1.5mL/well of 6% glutaraldehyde, 0.5% crystal violet for 30min at room temperature. Plates were washed with water and imaged.

PD-L1 expression analysis

Cells were seeded at 100,000 cells/well in 12-well plates and allowed to adhere overnight in complete medium. The next day constructs were added in serum free medium and incubated overnight at 37°C for 12-16hrs. EGF was then added at a final concentration of 25ng/mL in each well and again cells were incubated overnight to induce PD-L1 expression. Cells were then trypsinized, washed and incubated with anti-PD-L1-APC antibody (Biolegend) for 1hr at 4°C, washed again and then quantified by flow cytometry.

In vivo efficacy and PK study in patient-derived xenograft models

All animal work was performed at the Mayo Clinic, Rochester, MN, in accordance with IACUC animal protocols. Patient-derived xenografts were inoculated subcutaneously in the flank of immunocompromised NOD/SCID mice on day 0. Tumor size measurements and treatments were begun on day 7; doses were 200ug of drug, injected retro-orbitally twice weekly for three weeks total. Tumor volumes were calculated as $(LxW^2)/2$.

ELISA for cetuximab or cetuximab-Sso7d fusion PK

Blood samples from PDX mice or from tumor naïve nude mice were drawn by retroorbital bleed 48hrs after either the indicated dose; serum was collected after centrifuging and stored at -20°C. Human Therapeutic IgG1 ELISA Kit (Cayman Chemical) was used for analyzing amount of Fc present in mouse serum. Serum was diluted 1:1000 or 1:10000 in the assay and manufacturer's protocol was followed.

Chapter 3 – Engineering a human/mouse cross-reactive cetuximab

Introduction

Cetuximab is a monoclonal antibody which targets and antagonizes the function of human EGFR by sterically blocking binding between the receptor and its ligands. First FDA approved in 2004 it is employed today either as a single-agent or in combination with other chemotherapies for the treatment of lung, colorectal and head and neck cancers[40–44]. Treatment with cetuximab results in statistically significant increases in overall survival, but these benefits are modest; for example, in a clinical trial which supported cetuximab's FDA approval in colorectal cancer, the median duration of response was only 5.7 months when used in combination with chemotherapy.

Beyond the limited efficacy there are other challenges of cetuximab therapy. First, predicting response to treatment remains challenging. Anti-EGFR antibodies such as cetuximab and panitumumab are counter-indicated in patients harboring Ras mutations[122], but even among Ras wild type patients response rates are less than 50% in CRC; this number is even lower (~10-15%) for head and neck cancer, or when cetuximab is used as a second-line treatment[40,123,124]. Molecular and biological mechanisms underlying this selectivity of response are not well understood, and thus the quest for better biomarkers or alternative genetic signatures continues[125–127].

Another limitation of cetuximab therapy is the inevitable development of acquired resistance. Molecular alterations within the EGFR gene, oncogenic muations in downstream effectors of EGFR, or mutations in other EGFR family members can all inactivate cetuximab and lead to disease progression even in the context of positive initial responses. These mechanisms have been extensively reviewed and remain the topic of current investigation[123,128–133].

Suboptimal peri- and intratumoral distribution of cetuximab may also contribute to its limited efficacy. Pharmacokinetic studies of cetuximab in humans have shown dose-dependent clearance rates, a pattern consistent with target-mediated drug disposition (TMDD)[134,135]. These studies informed clinical practice, where cetuximab is administered through intravenous infusion at an initial loading dose of 400mg/m^2, followed by a weekly maintenance dose of 250mg/m^2. This maintenance dose is at the saturating end of the dose-dependent PK curve and is expected based on theoretical and modeling calculations to deliver intratumoral concentations of antibody that are adequate for EGFR inhibition. But a quantitative measurement of cetuximab's tumor uptake is difficult to obtain from human patients, and the impact of its distribution on drug efficacy is incompletely understood. However, studies in mice have shown that heterogeneous antibody distributions are often found in tumors despite high systemic drug doses[136,137] and studies in humans have shown correlations between cetuximab clearance and drug response[134,138]. Overall these findings suggest that a more advanced understanding of cetuximab distribution in tumors might lead to improved treatment paradigms.

As an alternative to molecularly targeted therapies such as cetuximab, immunotherapy has emerged as a powerful new approach for treating cancers. There are multiple classes of immunotherapies with distinct mechanisms of action, but all act generally to harness the power of the immune system against cancer to generate durable cures. In some cancers, treatment with a single immunotherapy such as a checkpoint blockade antibody can yield striking results – for example, when durvalumab, a PD-L1 inhibitor was applied in stage III non-small-cell lung cancer, or nivolumab, a PD-1 inhibitor was used in melanoma, these drugs resulted in dramatic extensions in progression free survival[139,140]. However in some patient subsets and in certain immunologically

"cold" tumors even these revolutionary drugs are ineffective or insufficient alone, and the potential value of combination approaches is becoming clear.

The opportunities and motivations for considering immune-focused combinations with cetuximab are numerous. It has been previously demonstrated that antibodies against tumor-expressed antigens can synergize with other immunotherapies such as systemically administered cytokines, cancer vaccines and checkpoint blockade to create combinations with increased antitumor efficacy[141,142]. In these contexts tumor-targeting antibodies mediate cell killing by recruiting innate effectors through interactions between the Fc and FcγRs, and by helping to recruit and prime tumor-antigen-specific T-cell responses. ADCC has long been appreciated as an essential component of cetuximab's antitumor efficacy[56,143], and more recent studies have elucidated the immunomodulatory effects of EGFR signaling which can be altered through use of EGFR inhibitors like cetuximab[59,61,144–149]. The diverse array of effects mediated by cetuximab suggests multiple promising avenues for potentiating anti-tumor activity through combinations[150,64,151,152].

Overall, then, there is clearly room for improved understanding and utilization of anti-EGFR antibodies. Better measurements of cetuximab biodistribution and tumor uptake could inform new administration methods or dosing regimens; improved genetic or biomarker signatures that correlate with patient response or development of acquired resistance could aid in more effective patient stratification; the immunomodulatory effects of cetuximab treatment could be harnessed to rationally design synergistic combination therapies. All of these are active areas of exploration, but often the tools available for studying these phenomena are not optimal. This is particularly true for pre-clinical studies in mice, since the cetuximab antibody does not bind the murine EGFR protein. As a result mouse models using or studying cetuximab require the presence

of human EGFR; this is achieved either through xenografts implanted in immune compromised mice, or by the introduction of human EGFR into otherwise murine cells. In both cases the model system lacks critical components which could affect study outcomes, such as an intact immune system, relevant levels of endogenous EGFR expression, or native signaling pathways. Clinical studies in humans can obviously provide valuable biological insight but the small scale and retrospective nature limits their utility as well.

Thus, in this study we aimed to utilize yeast display to "murinize" cetuximab and create a relevant mimic for the antibody which could be used in important pre-clinical explorations. To be maximally relevant the murinized antibody would need the same properties as cetuximab – similar affinity, identical binding epitope, and the ability to comparably inhibit EGFR signaling. To date, the only published attempt at a pre-clinical cetuximab mimetic is the 7A7 antibody[153], though the value of that molecule was recently called into question[154]. Our approach was to use the cetuximab sequence as the starting template, rationally design a mutant library and sort for murine EGFR binding.

Results

Validation of cetuximab scFv on yeast

The variable domains of cetuximab were cloned into the pCTCon2 vector for yeast display as an scFv, and expression and binding to human and mouse EGFR were assessed. The construct was well-expressed on the surface of yeast (Figure 3.1) as measured by labeling of the C-terminal cmyc tag. The expected binding profile was also observed – the cetuximab scFv showed clear binding to a bivalent human EGFR-Fc antigen but not to the murine version.

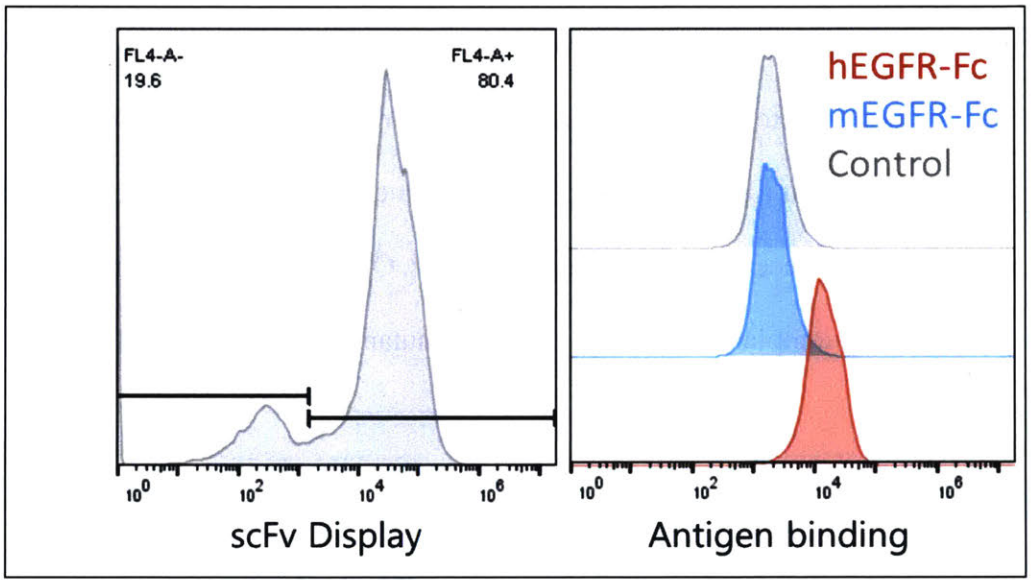


Figure 3.1 – Yeast display of cetuximab scFv. **Left** – Cetuximab expression on yeast surface was induced and full-length expression was assessed by labeling the C-terminal cmyc tag. Cells were incubated with chicken-anti-cmyc primary antibody, followed by an anti-chicken-AF488 secondary antibody before being measured by flow cytometry. **Right** – Cetuximab binding was measured by incubating induced yeast with anti-cmyc antibody and indicated EGFR-Fc construct at 25nM or control. Cells were gated for cmyc+ cells, and histograms of binding MFI are shown.

Yeast display library design and selection of a cross-reactive cetuximab

For generating a cetuximab scFv mutant library a saturation mutagenesis approach was chosen, in which designated codons are mutagenized to ensure the complete sampling of every

amino acid at each mutated position. Because a typical yeast display library has on the order of $1x10^8$ members the number of mutated sites in a given library was limited to six, yielding a maximum theoretical amino acid diversity of $20^6 = 6.4x10^7$ per library. Selection of the mutated sites was guided by examination of the crystal structure of human EGFR-ECD in complex with cetuximab as well as the alignments of mouse and human EGFR.

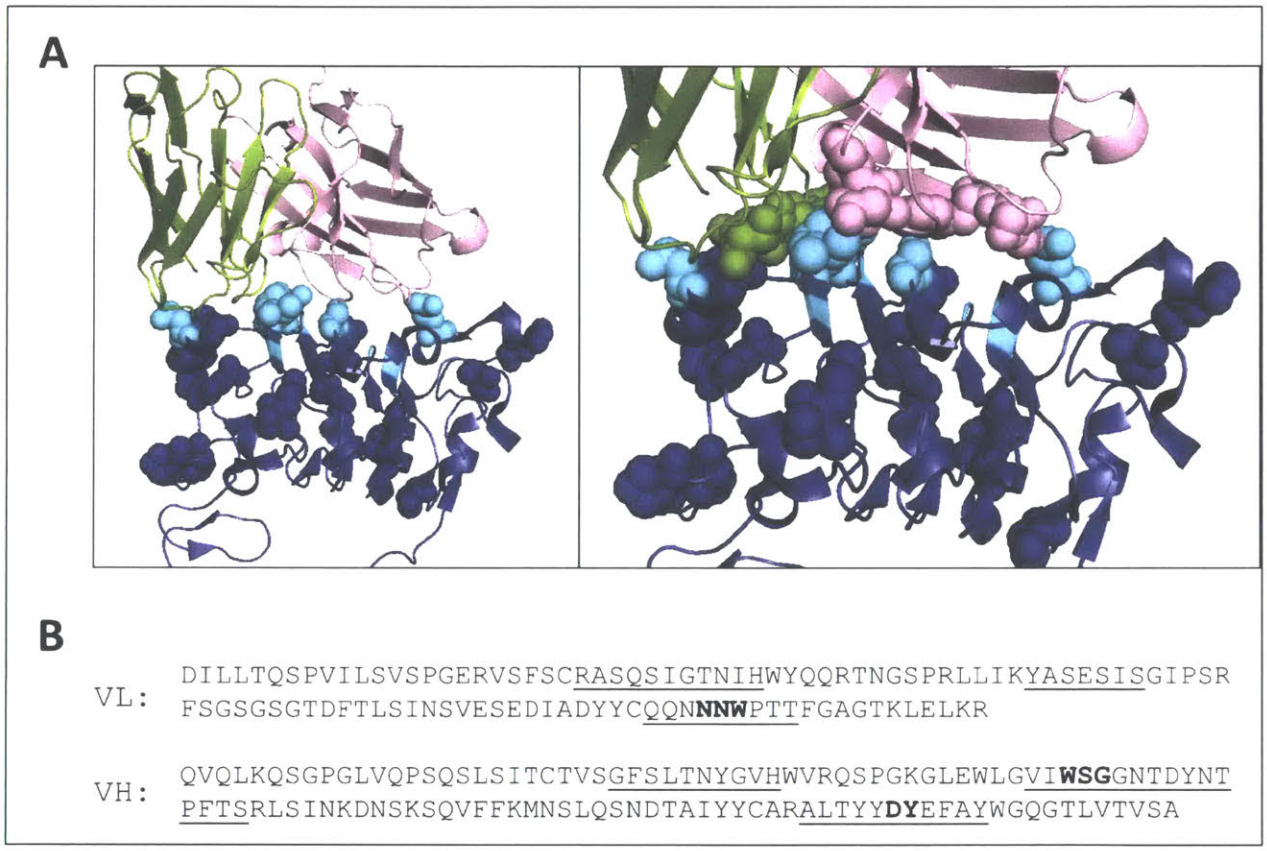


Figure 3.2 – Structural considerations which guided cetuximab mutant library design. **A** – Structure of human EGFR (blue) in complex with cetuximab (green VL, pink VH) from Li et al[48], PDB ID: 1YY9. Left – Spheres indicate the non-conserved residues between mouse and human EGFR domain III. Cyan represents residues that have specific interactions with cetuximab. Right – Pink and green spheres indicate the 8 residues that were mutated in the yeast display libraries. Bottom: Sequences of cetuximab VH and VL, with CDRs underlined. Mutated sites are in bold.

The structure shows that several amino acids important for mediating contact with cetuximab are non-conserved between the two species (Figure 3.2). It was hypothesized that compensatory mutations on the cetuximab side of the interface near these non-conserved sites

might confer cross-reactivity of the antibody. Specifically, eight cetuximab positions of interest were identified, spanning three linearly contiguous regions within CDRs L3, H2 and H3 (Figure 3.2 B).

Three yeast display libraries were built, each with no more than six of these sites mutated at a time, according to the scheme in Figure 3.3. FACS selections were performed on the three libraries using the bivalent murine EGFR-Fc construct as antigen. After four sorts on each library only Library C showed improved binding (Figure 3.4 A); interestingly, while the enriched population showed weak binding at 150nM to bivalent murine EGFR, the affinity for human EGFR was unaltered. A few clones from this population were sequenced and the enriched mutations are shown in Figure 3.5. The only mutations found in this enriched population were located in CDR L3 within the NNW sequence. A single site, N92, was mutated in 14 of 15 sequenced clones, where 11 of those clones contained the N92L mutation and another 3 contained N92M. Maintenance of human binding was likely a result of the fact that none of the cetuximab residues which directly contact EGFR were mutated during this round.

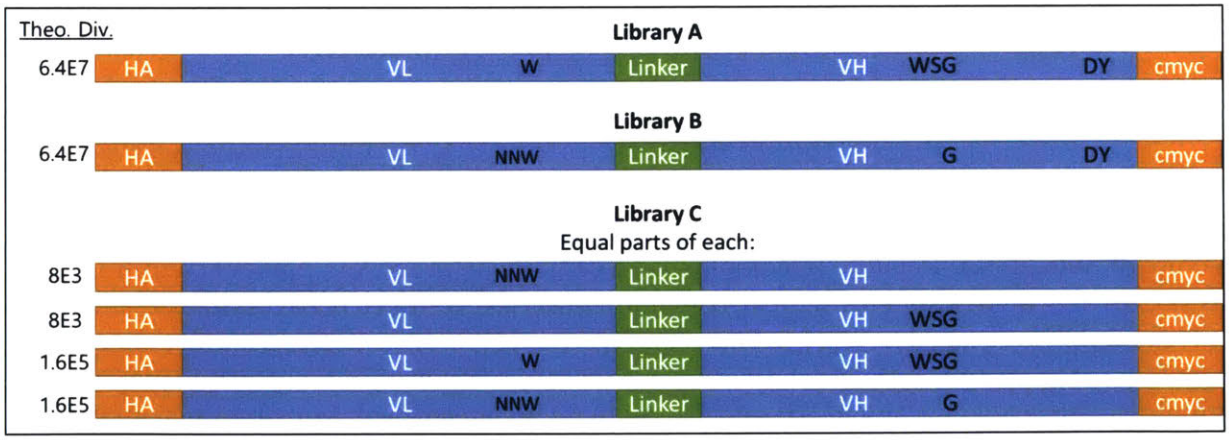


Figure 3.3 – Saturation mutagenesis library design for initial round of sorts for murine EGFR binding. Of the eight identified mutation sites of interest, combinations of no more than six were combined into a single library to limit the theoretical diversity to less than the number of yeast transformants in the library. Mutations were introduced into each contiguous segment using a modified NNK approach (see Methods for more details) and then full-length inserts were

assembled by overlap-extension PCR before transformation into yeast. Three total libraries were built and sorted. Theoretical diversity of each full-length insert is indicated.

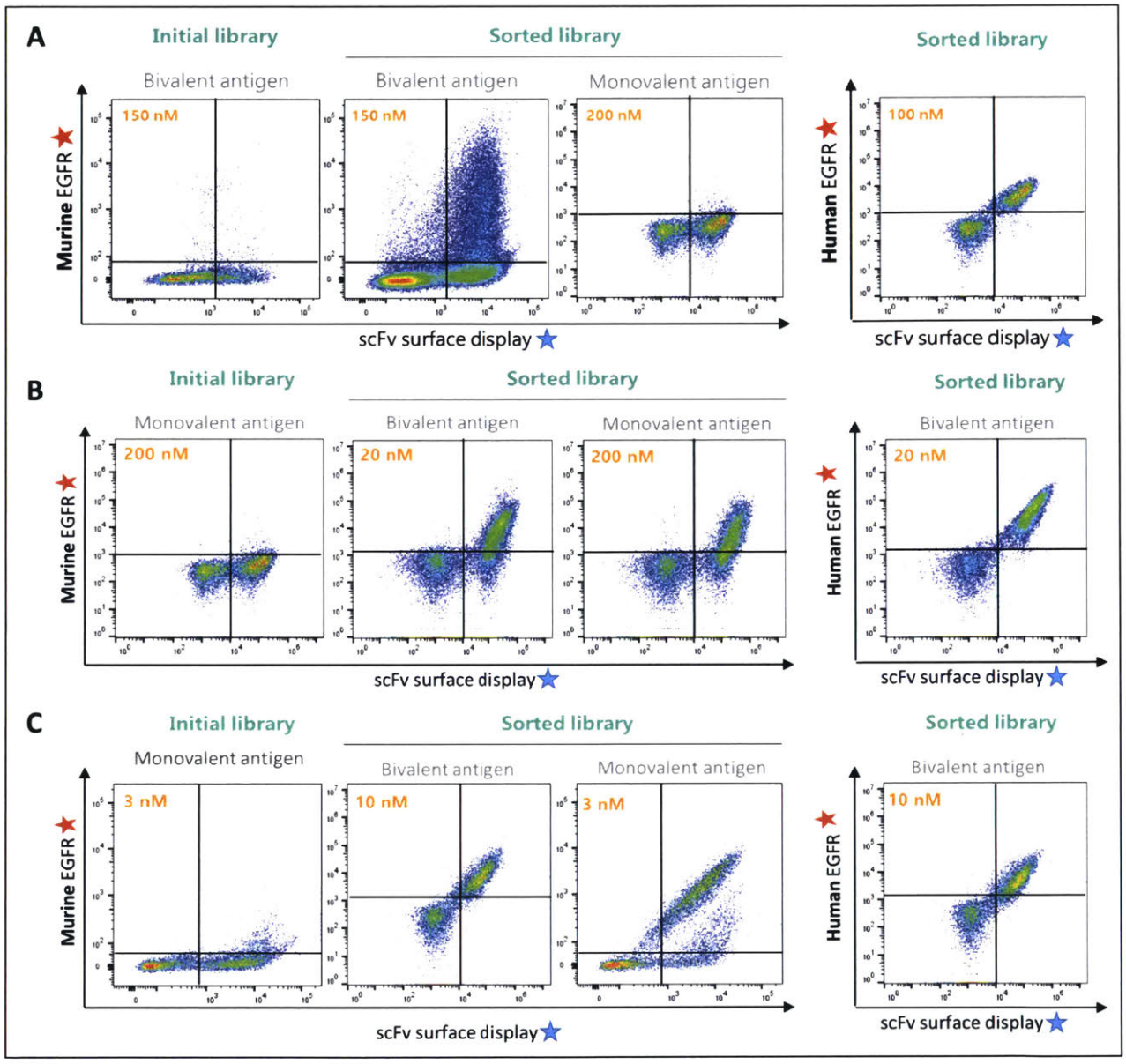


Figure 3.4 – Sort progression during selections for murine EGFR binding. **A** – In Round 1, Libraries A, B and C (from Figure 3.3) were sorted using chicken anti-cmyc as the display antibody and 150nM biotinylated EGFR-Fc as the antigen. After four sorts, enriched populations were tested for binding at the indicated concentrations to murine and human EGFR antigens before sequencing. **B** – The Round 2 library was mutagenized by error-prone PCR before three sorts against lower concentrations of bivalent antigen. Again, enriched populations were tested for binding and cross-reactivity at indicated concentrations before sequencing. **C** – The final, Round 3, library was another round of error-prone PCR and sorting, this time using EGFR-monoFc as the antigen.

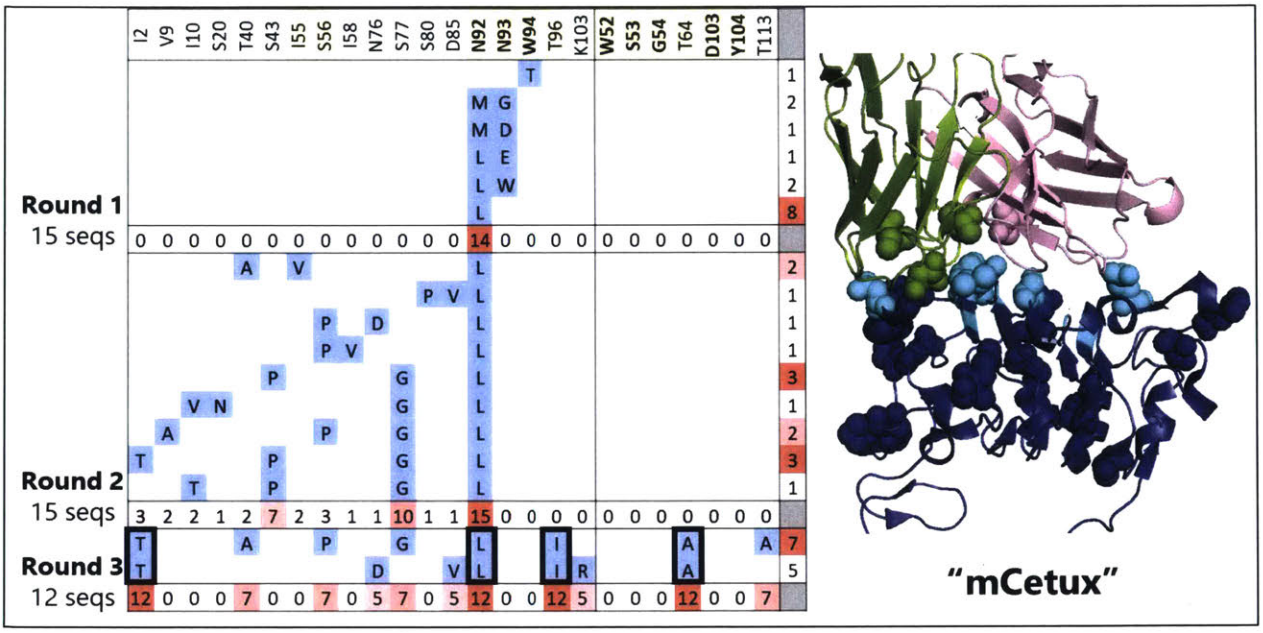


Figure 3.5 – Sequence progression of yeast display sorts during selection of mCetux. Cetuximab amino acids are listed along the top; residues within CDRs are highlighted in yellow; bold positions are those selected for the saturation mutagenesis library in round 1. Rows within each library represent individual clones found; numbering at the bottom of each library represents the number of time a mutation was observed at a particular position out of a total number of sequences acquired (listed to the left). Numbering along the right represents the total number of times each individual clone was observed. Final mutations in mCetux are boxed and shown as green and pink spheres on structure at right. Other coloring in the structure is the same as Figure 3.2.

DNA from this enriched population (a pooled mixture of the sequences shown for Round 1 in Figure 3.5) was used as the template for a subsequent round of affinity maturation using error-prone PCR over the entire scFv gene. To increase stringency of the selections in this second round antigen concentration was lowered in successive sorts down to 20nM. After this round the enriched population had high enough affinity to engage the monovalent murine EGFR construct (EGFR-monoFc) suggesting improvement in K_d, and again, binding to human EGFR was preserved. A variety of new mutations were found by sequencing single clones, including some both within and outside of CDRs (Figure 3.5, Round 2).

A third and final round was constructed, again by error-prone PCR, and sort stringency was increased by switching from bivalent to monovalent EGFR as the antigen. At the end of this

final round, tight binding was observed to low concentrations of EGFR-monoFc, apparently mediated by four consensus mutations that appeared in all 12 sequences analyzed (Figure 3.5, Round 3).

Ultimately a single clone, incorporating only the consensus mutations from the final round of sorting, was identified and constructed; its yeast binding titration against monovalent human and murine EGFR is shown in Figure 3.6. This clone has a monovalent K_d of 3.1nM against murine EGFR, almost identical to the affinity of cetuximab for human EGFR. Interestingly, it has not just preserved but slightly improved binding against human EGFR, even though this property was never explicitly a selection criteria. In addition to the original N92L mutation in CDR L3, which resulted from the saturation mutagenesis library, it has three additional mutations, all of which became dominant during this third round. Two of these three, T96I in the LC and T64A in the HC, were also cetuximab CDR mutations, though not amino acids specifically implicated in the cetuximab/human EGFR interaction. The final new mutation was I2T of the LC, which is not a CDR residue but is located at the interface with EGFR. This evolved antibody clone was called "mCetux" and is the antibody used in studies throughout the remainder of this work.

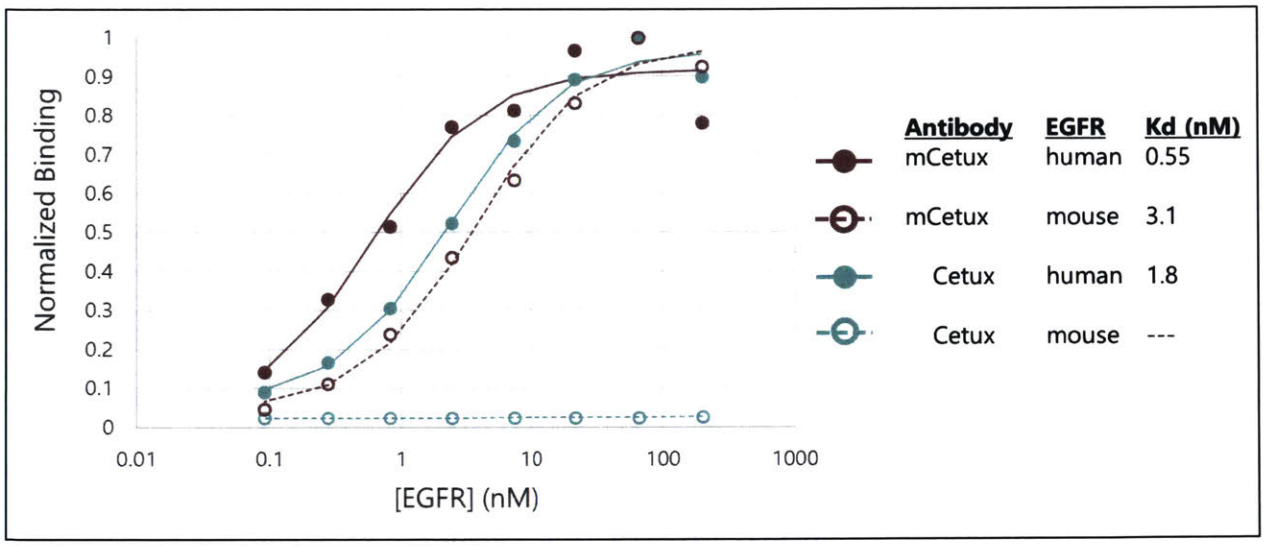


Figure 3.6 – Yeast titrations of mCetux and cetuximab against monovalent human and murine EGFR. Raw data is shown in dots, solid and dashed lines are fitted curves for monovalent binding isotherms. Data is representative of multiple experiments.

Validation of mCetux as cetuximab mimetic

In order to demonstrate that the four mutations in mCetux did not change the antibody binding epitope on EGFR, competition binding assays were performed. mCetux scFv expressed on yeast was bound to murine EGFR or human EGFR, with or without an excess of soluble cetuximab. Figure 3.7 shows that an excess of soluble cetuximab had no effect on the binding of mCetux to murine EGFR, but was able to completely ablate binding to human EGFR. Though an indirect measure, this result suggests that mCetux and cetuximab have at least overlapping, if not identical epitopes, at least for human EGFR. As another measure of epitopic similarity, the ability of the EGFR ligand, EGF, to compete with antibody for EGFR binding was assessed. Excess soluble ligand caused a dramatic reduction (>90%) in binding between mCetux and human EGFR, consistent with the idea that the antibody engages the ligand binding site of the receptor.

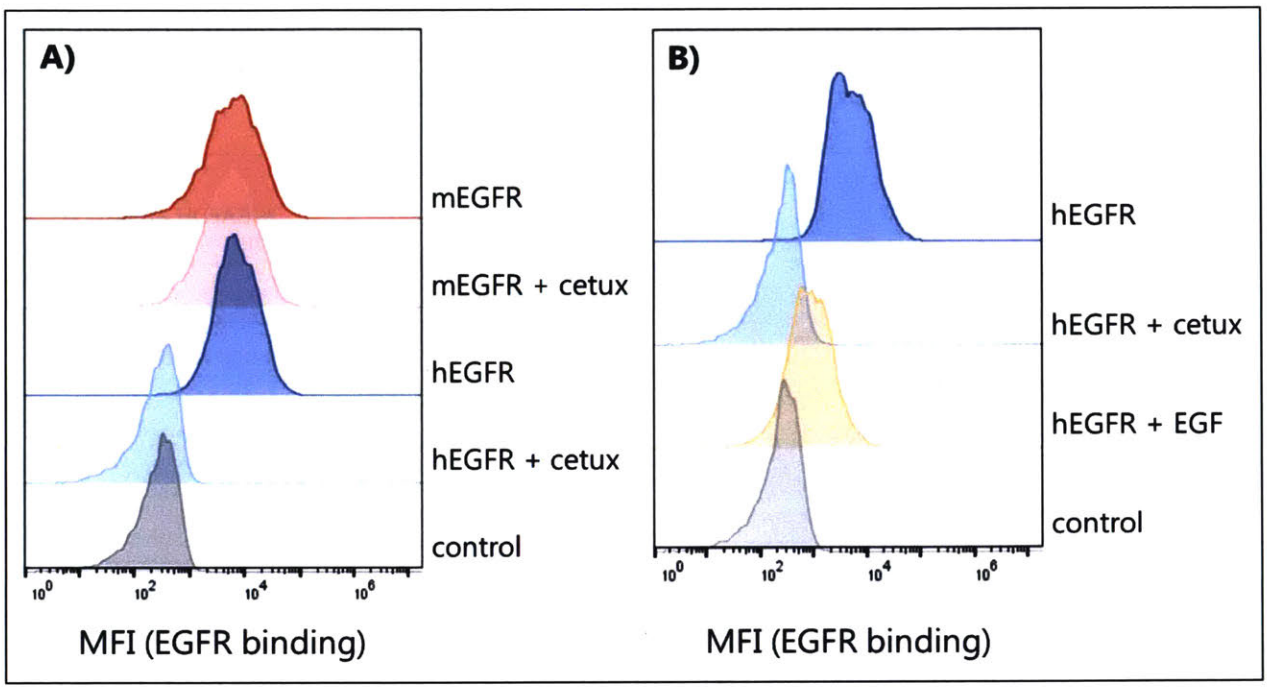


Figure 3.7 – Competition assays to assess mCetux binding epitope on EGFR. **A** – mCetux displayed on yeast was bound to biotinylated monovalent murine or human EGFR (10nM) with or without excess soluble cetuximab (200 nM). **B** – Competition with the EGFR ligand human EGF (200nM) was also assessed on yeast.

The mCetux scFv sequence was reformatted and expressed as a full-length IgG, and the binding of the antibody to EGFR expressed endogenously on cell lines was subsequently evaluated. Cetuximab bound only to the human cell line A431, known to express high levels of EGFR, and the murine MC38 cell line transfected with a "humanized" murine EGFR which has been shown to bind cetuximab (see Methods); no binding was detectable to any murine cell lines, as expected, or to EGFR-null CHO cells which served as the negative control. For mCetux, binding was detectable to both human EGFR expressing cell lines, not to the CHO cells, and among the murine cell lines the results were mixed (Figure 3.8). In order to demonstrate that these varying signals represented specific binding to expressed murine EGFR and not non-specific interactions with the cell membrane, qPCR was done to quantify murine EGFR expression in each cell line.

The correlation of murine EGFR RNA levels with the extent of binding on each cell type confirms that the signals represent specific interactions between mCetux and EGFR.

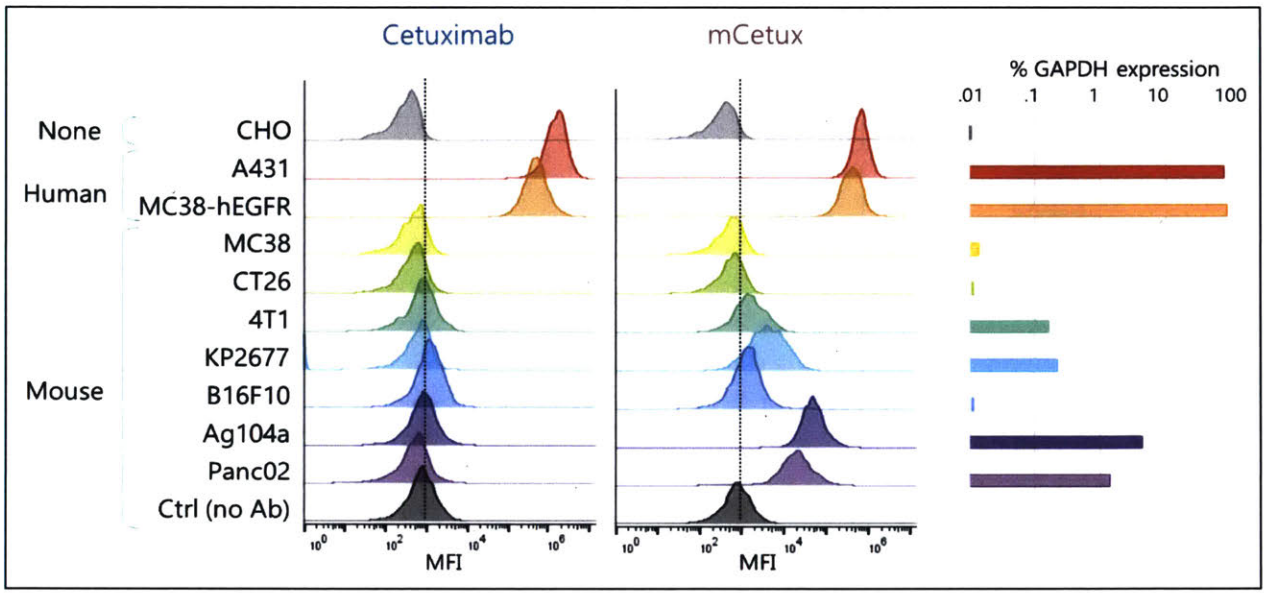


Figure 3.8 – Cell surface binding of mCetux. Cells were trypsinized, washed and exposed to 20nM antibody (cetuximab, left; mCetux, center) for 1 hr at room temperature. Binding was detected with an anti-human-488 antibody and samples were read by FACS. qPCR for murine EGFR RNA is shown, right, normalized to expression of housekeeping gene GAPDH.

The main molecular mechanism of action of cetuximab is to antagonize ligand-mediated EGFR signaling by direct competition. To determine if mCetux could similarly inhibit EGFR signaling on murine cell lines, in vitro signaling assays were performed. Cell lines expressing human or murine EGFR were stimulated with EGF in the presence or absence of antibody, and lysates were assayed for phosphorylated ERK by ELISA as a measure of downstream signaling. Both cetuximab and mCetux inhibit signaling on human EGFR-expressing B16F10 cells, but only mCetux is active against murine cells expressing endogenous EGFR (Figure 3.9). Further, because multiple distinct ligands are capable of binding EGFR and initiating signaling, the inhibition assay was also run using amphiregulin, heparin-binding EGF and TGFα as ligands. Again, mCetux

blocked signaling in both murine and human cells (Figure 3.9). The potency of mCetux in the human A431 cell line was comparable to cetuximab.

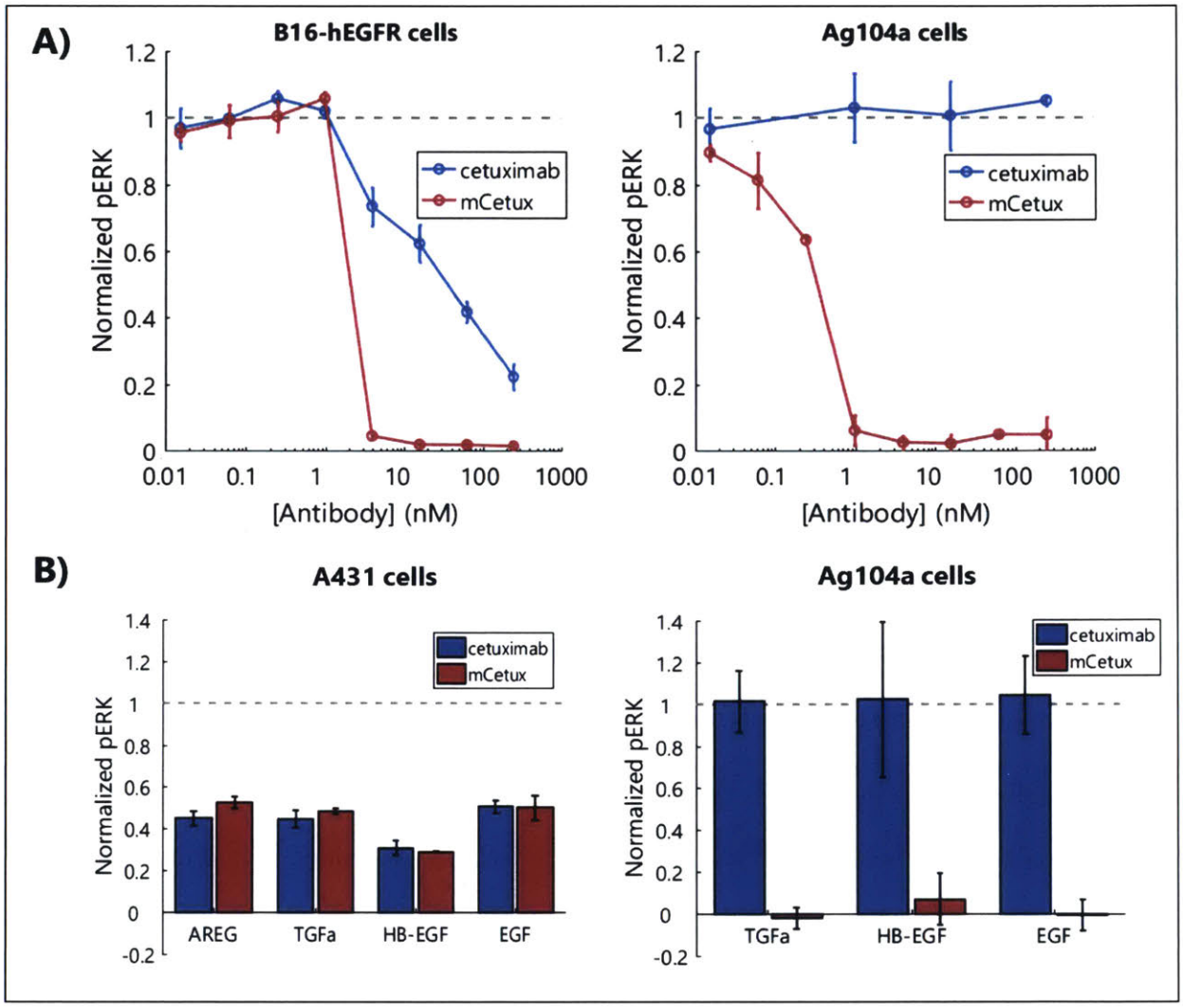


Figure 3.9 – Inhibition of EGFR signaling by mCetux. Adherent cells in confluent 12-well plates were serum starved, then incubated with antibody at indicated concentrations for 1hr at 37°C. Cells were then stimulated with ligand for 5min, immediately washed and lysed with RIPA buffer. Lysates were analyzed for pERK by ELISA. **A** – B16-hEGFR are cells stably transfected with a mutant EGFR that binds both mCetux and cetuximab, Ag104a murine cells express endogenous murine EGFR. Ligand is EGF at 8nM. **B** – A431 human cells and Ag104a murine cells both express endogenous EGFR. Ligand concentrations are 8nM EGF, 15nM AREG, 10nM HB-EGF, and 5nM TGFα.

Evaluation of mCetux pharmacokinetics in tumor-bearing mice

In preparation for in vivo administration of mCetux for therapeutic studies, the pharmacokinetics of the engineered antibody were investigated. Studies of cetuximab pharmacokinetics in humans have shown that the antibody exhibits dose-dependent clearance rates and undergoes target-mediated drug disposition (TMDD). The human clinical dosing regimen for cetuximab is 250 mg/m^2 weekly, equivalent to approximately 7mg/kg. Converting this dose to an appropriate injection in mice based on weight would indicate an equivalent dose in mice to be less than 150ug, in line with what is typically used for tumor targeting antibodies[141,142] or in studies with cetuximab in nude mice[58]. But allometric scaling according to body surface area[155] indicates an order of magnitude more protein is required – 1.67mg, or 83mg/kg. To assess whether mCetux exhibited this same dose-dependent pharmacokinetic behavior, and determine the appropriate dosing in mice, pharmacokinetic and tumor uptake studies were conducted. First, mice bearing subcutaneous B16-hEGFR tumors were given a single intraperitoneal dose of fluorescently labeled mCetux or cetuximab and serum concentrations were measured over time. For all four doses of cetuximab tested serum concentration peaked by 8 hours post-injection at which point nearly all of the injected dose had reached systemic circulation. Clearance was slow, with a terminal $t_{1/2}$ of approximately 200hrs, and independent of dose (Figure 3.10 A). In contrast, peak systemic mCetux concentrations were always well below the full administered dose, while the antibody half-life was much shorter ($t_{1/2}$ ranging from 20-52 hrs) and increased with dose, consistent with a molecule undergoing TMDD. When clearance rates were plotted against administered dose the trend closely mimicked that previously documented for cetuximab (Figure 3.10 B).

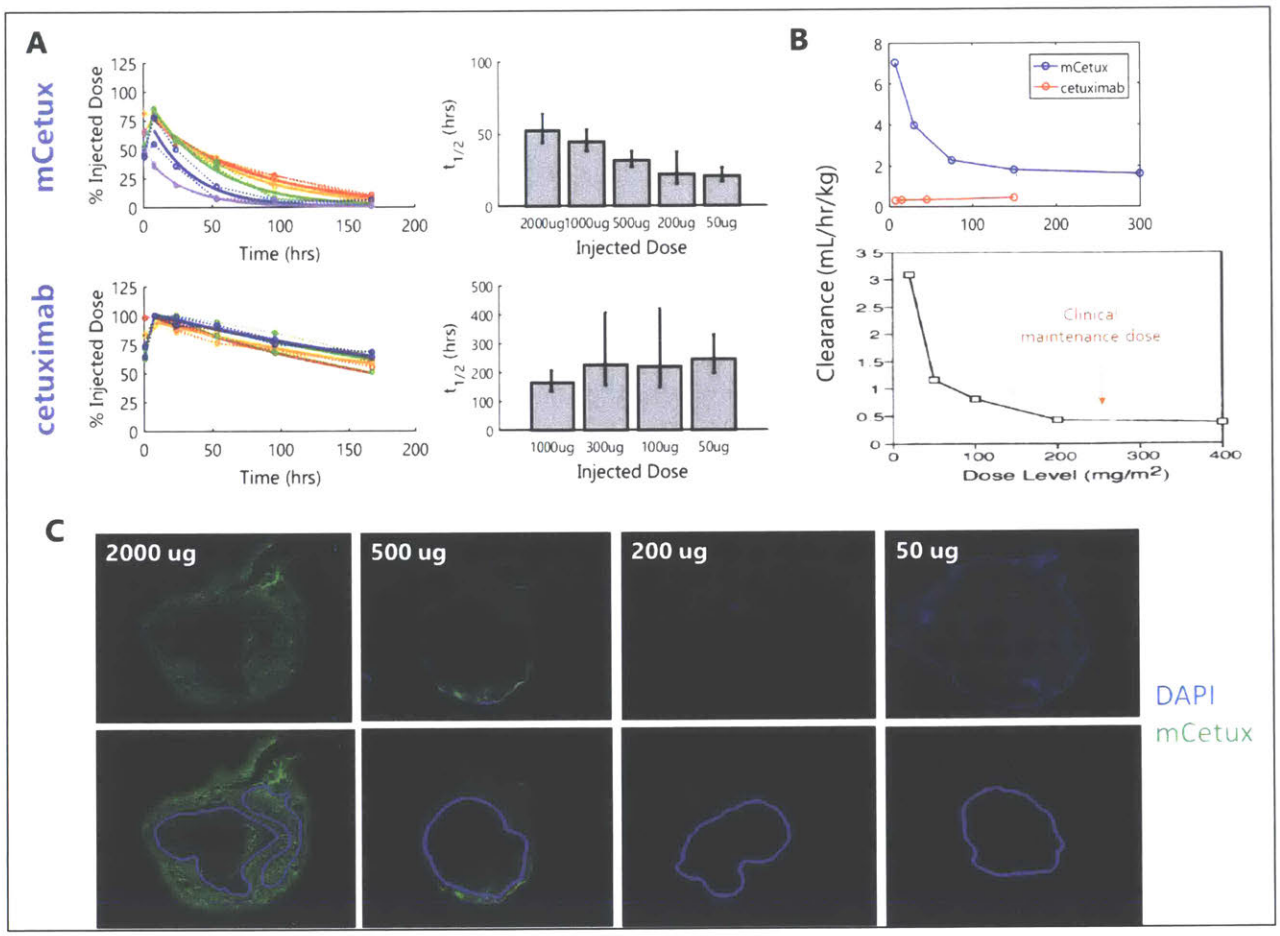


Figure 3.10 – Pharmacokinetic study of mCetux in tumor-bearing mice. Mice were inoculated with B16-hEGFR tumors and injected i.p. with a single dose of Alexa Fluor 647-labeled antibody on day 7 (2 mice / group). Blood was collected from the tail vein at indicated time points. **A** – Serum antibody concentrations (circles, dashed lines) over time for all mice along with non-compartmental fit (solid lines). Terminal half life is calculated from the non-compartmental fit of each group. **B** – Calculated clearance plotted against dose for each antibody (top), compared to historical data from Baselga, et al[134] of cetuximab clearance in humans (bottom). The clinical dose of 250 mg/m^2 is labeled for reference. **C** – Microscopy showing distribution of mCetux in B16-hEGFR tumor sections. Mice were inoculated with 1E6 B16-hEGFR cells at day 0, and at day 10 were given a single i.p. injection of the indicated amount of mCetux-AF647. Tumors were harvested 24hr later, and frozen sections were imaged by microscopy. Top panel is composite image, lower panel is mCetux-AF647 channel only (green), with outline shown to differentiate tumor from surrounding skin. Tumor area was based on H&E staining.

To qualitatively assess tumor distribution of mCetux tumors were harvested 24 hours following a single intraperitoneal injection of fluorescently-labeled antibody and visualized by microscopy (Figure 3.10 C). Antibody was barely visible in tumor sections at doses below 500ug,

which is in stark contrast to the saturation of tumors observed by Rhoden, et al[86] for tumor targeting antibodies at doses of greater than 150ug. Strikingly, fluorescence intensities were even higher in surrounding skin than in tumor at all doses tested. This poor tumor distribution is in line with some previous reports for other anti-EGFR antibodies[136,156]. Other studies have indicated that uniform distributions of cetuximab can be achieved at high doses[157], but it is notable that these analyses were done in xenografts which lack the complexity of endogenous, non-tumor target expression. The ability to account for and capture these types of complicating but physiologically relevant factors is what underlies the value of mCetux.

Overall these data demonstrate that mCetux pharmacokinetics closely mimic those of cetuximab, and suggest a large and frequent dose of 1-2mg is most appropriate for studying cetuximab in vivo to gain biological and therapeutic insight. The high doses required are inconvenient from a practical protein production standpoint, but ultimately the fidelity of mCetux behavior in mice to that of cetuximab in humans inspires confidence in its usage and suggests that in vivo experiments will yield relevant results.

Application of mCetux in murine colorectal organoids model system

For a model system in which to apply mCetux, a series of genetically-defined murine colorectal cancer organoids was chosen. This system, developed within the Yilmaz Lab at MIT, was particularly of interest since colorectal cancer is one of the approved indications for cetuximab use in humans. The organoids were derived from mouse colon stem cells and then edited to introduce known CRC driver mutations either singly or in combinations which occur in the context of human CRC patients[158]. The organoids can be implanted and studied orthotopically, making them a particularly relevant model for therapeutic studies.

mCetux was first tested for its effect on organoid growth in vitro. Organoids harboring combinations of driver mutations in APC (A), K-Ras (K), p53 (P), or Smad4 (S) were cultured in the presence of varying concentrations of mCetux or cetuximab in a minimal media lacking any growth factors. In this context the only source of EGFR ligands would be autocrine production by organoid cells. When growth was quantified after 3 days using resazurin dye, mCetux antibody clearly had a dose-dependent effect on both A and AP organoids. This effect is also visible in microscopy images taken of the A organoids at this same day 3 time point (Figure 3.11 B). Cetuximab, however – which should not bind the EGFR expressed on these murine cells – caused no change in growth for A organoids at even the highest concentrations tested. The effect of mCetux was much weaker on AKP organoids which also harbor the oncogenic K-Ras G12D mutation; since K-Ras is a downstream effector of EGFR and this mutation should result in constitutive signaling, this data is consistent with the mCetux mechanism of action being EGFR antagonism. Interestingly, however, at the highest concentrations tested there was a small but reproducible inhibition of growth for the AKP organoids, suggesting that mCetux exerts an effect on other EGFR signaling pathways, such as the PI3K pathway, even when ineffective against Ras/Raf/MEK/ERK signaling.

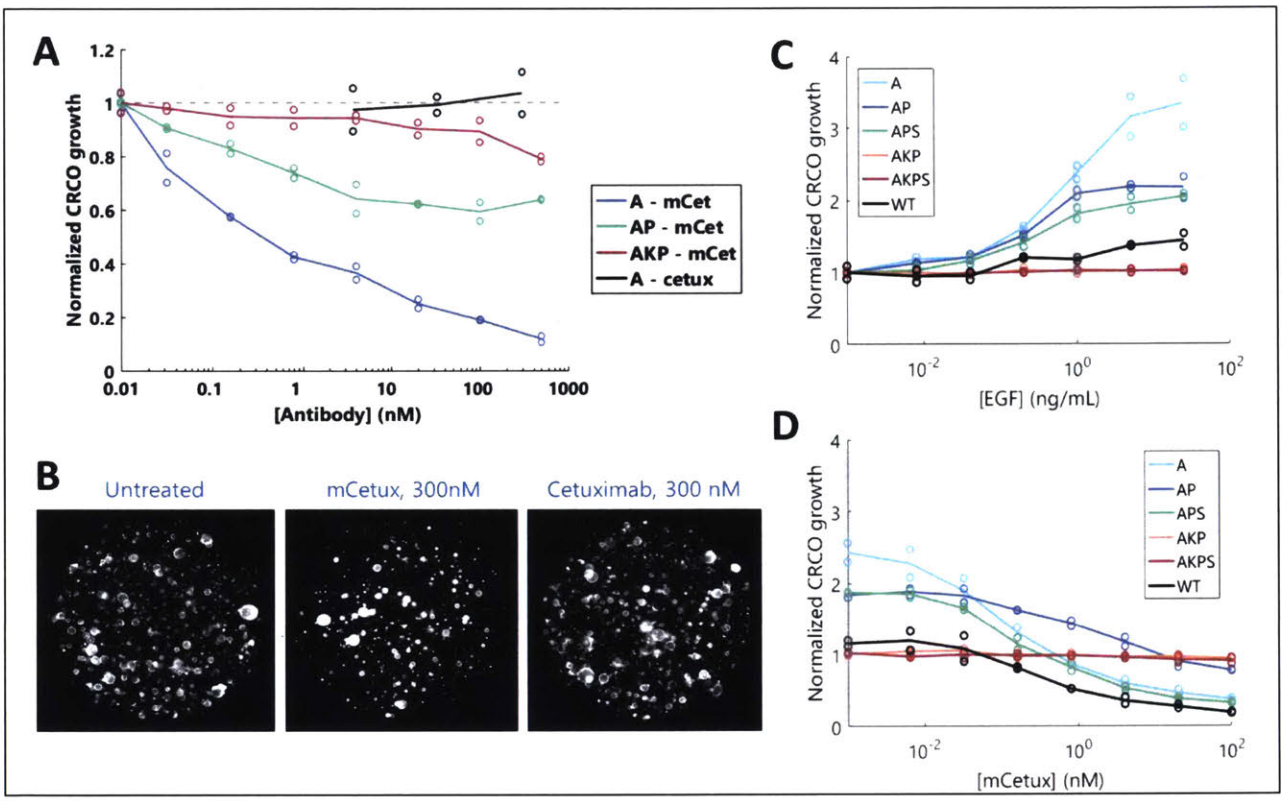


Figure 3.11 – In vitro effect of mCetux on CRC organoids. **A** – mCetux and cetuximab were incubated with organoids harboring oncogenic mutations (A: APC, P: p53, K: K-Ras, S: Smad4) for 3 days in minimal growth media. After 3 days growth was quantified by resazurin dye and data were normalized to growth without antibody. **B** – Fluorescence microscopy of tdTomato positive "A" organoids after incubation with antibody at indicated concentration for 3 days. **C** – Effect of exogenous EGF on organoid growth, quantified by resazurin after 3 days of incubation and normalized to growth in the absence of EGFR. 1ng/mL concentration indicated by gray bar denotes EGF concentration used for testing mCetux inhibition in panel D. **D** – Effect of mCetux on indicated organoid growth in the presence of 1ng/mL EGF. Growth is quantified by resazurin after 3 days and normalized to controls which lacked EGF or antibody. For panels A, C and D the dots are raw data and the solid lines connect the means of the two repeats. Data are representative of multiple experiments.

Next, the effect of exogenously added EGFR ligand on organoid growth was tested. Addition of EGF to organoid cultures had a dose-dependent effect on growth that in all cases plateaued at roughly 1-10ng/mL (Figure 3.11 C). Growth potentiation was also genotype-dependent; high EGF concentrations resulted in a three-fold increase in growth for the A organoids, moderate advantage in AP and APS organoids, and an even less pronounced effect for

wild type organoids. Both organoids harboring the K-Ras mutation, AKP and AKPS, were insensitive to the added EGF, confirming that the growth advantage was mediated by signaling through EGFR.

Addition of mCetux abrogated the EGF-mediated growth advantage entirely, again in both a dose- and genotype-dependent manner (Figure 3.11 D). For three of the four EGF-responsive genotypes the highest concentrations of mCetux not only overcame the effect of added ligand but also further reduced organoid growth below that of the non-EGF treated controls. An exception was the AP organoids, which were less sensitive to mCetux inhibition than APS, despite a similar dose-dependent response to added EGF. And again, as expected, both K-Ras mutant organoids were insensitive to the effects of mCetux. Overall this data demonstrates the ability of the mCetux antibody to inhibit EGFR signaling in a physiologically and translationally relevant model system. And excitingly, the ability to discern genotype-specific responses to what is effectively cetuximab therapy in a murine system presents the possibility of exploring the biology that underlies the lack of response to cetuximab therapy in subsets of CRC patients.

RNAseq of colorectal organonids inhibited by mCetux

To explore whether mCetux could be used to better understand or exploit the apparent genotype-specific responses of the organoids to EGFR inhibition, RNAseq was performed on organoids treated with EGF and in the presence or absence of 100nM mCetux, a concentration at which growth was strongly inhibited in vitro. Unsupervised clustering of the organoid expression data for 11,454 genes shows distinct profiles based on organoid type and the effect of treatment (Figure 3.12). Treated and untreated A and APS organoids could all be distinguished from each other, and interestingly the A and APS treated conditions were more similar to each other than to

their respective untreated controls. These data mirror the observed growth inhibition patterns, where A and APS were most the susceptible organoids to cetuximab treatment. All organoids with a K-Ras mutation (AKP and AKPS, treated and untreated) clustered together and were not clearly distinguishable, implying that the effect of the K-Ras mutation on RNA expression dominated over the effect of mCetux. This corresponds with the in vitro finding in Figure 3.11 that all K-Ras mutated organoids were indifferent both to EGF stimulation and mCetux inhibition. Interestingly, the expression profile of AP organoids was very different from any other genotypes and also unchanged in the presence of mCetux. The distinct pattern of expression was partly expected, since it was clear from Figure 3.11 that this genotype responded differently, and less overall, to mCetux treatment than the A and APS organoids, but the fact that the treated and untreated samples cluster together was surprising, since EGF clearly potentiates AP growth and mCetux can inhibit it.

GSEA analysis of differentially expressed genes in treated vs. untreated organoids showed that gene sets related to cell cycle progression were highly enriched for decreased expression in the presence of mCetux (Figure 3.13). These genes included classical Myc and E2F targets as well as genes related to progression through the G2/M checkpoint. These findings correlate with the observed decrease in cell growth observed for these organoids in vitro. Interestingly, these same gene sets were significantly negatively enriched for AKP organoids, despite the presence of the K-Ras mutation. This implies that the effect of EGFR inhibition by mCetux can still be felt by cells with mutated K-Ras, despite the constitutive activity through the Ras pathway. The data in Figure 3.11 A, which show small but reproducible decrease in growth of AKP organoids at high mCetux concentrations and in the absence of exogenous EGF, support this hypothesis. Only AKPS organoids did not have transcriptional profiles characteristic of

decreased growth or cell cycle progression, indicating that the addition of the Smad4 mutation confers increased mCetux resistance over the AKP combination.

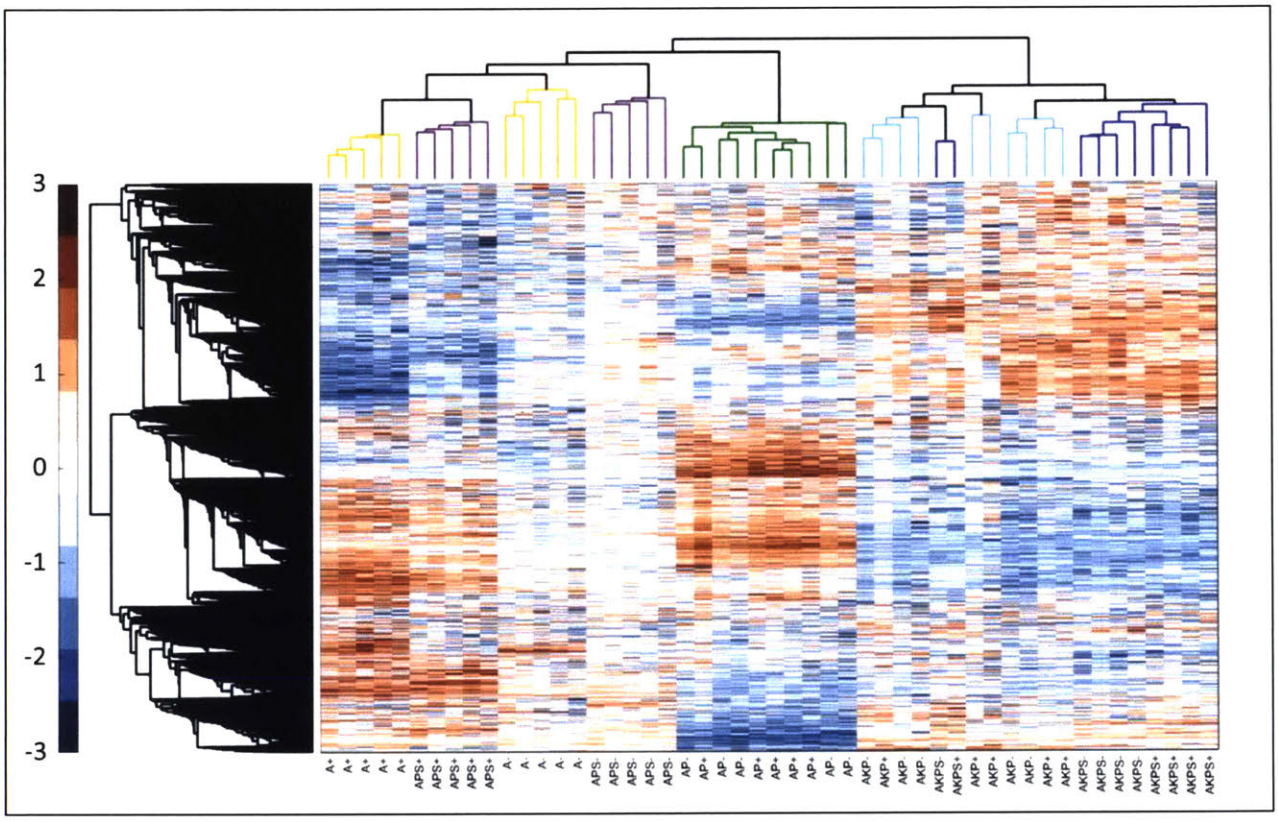


Figure 3.12 – RNAseq expression profiles of CRC organoids with and without mCetux treatment. Organoids (A – yellow; AP – green; APS – purple; AKP – cyan; AKPS – blue) were incubated with 1ng/mL EGF and with or without 100nM mCetux for 48hr. RNA was prepared from organoids after 24hr and analyzed by RNAseq. n=5 for each treatment condition. Expression data was log2 normalized and unsupervised clustering of 11,454 genes is shown.

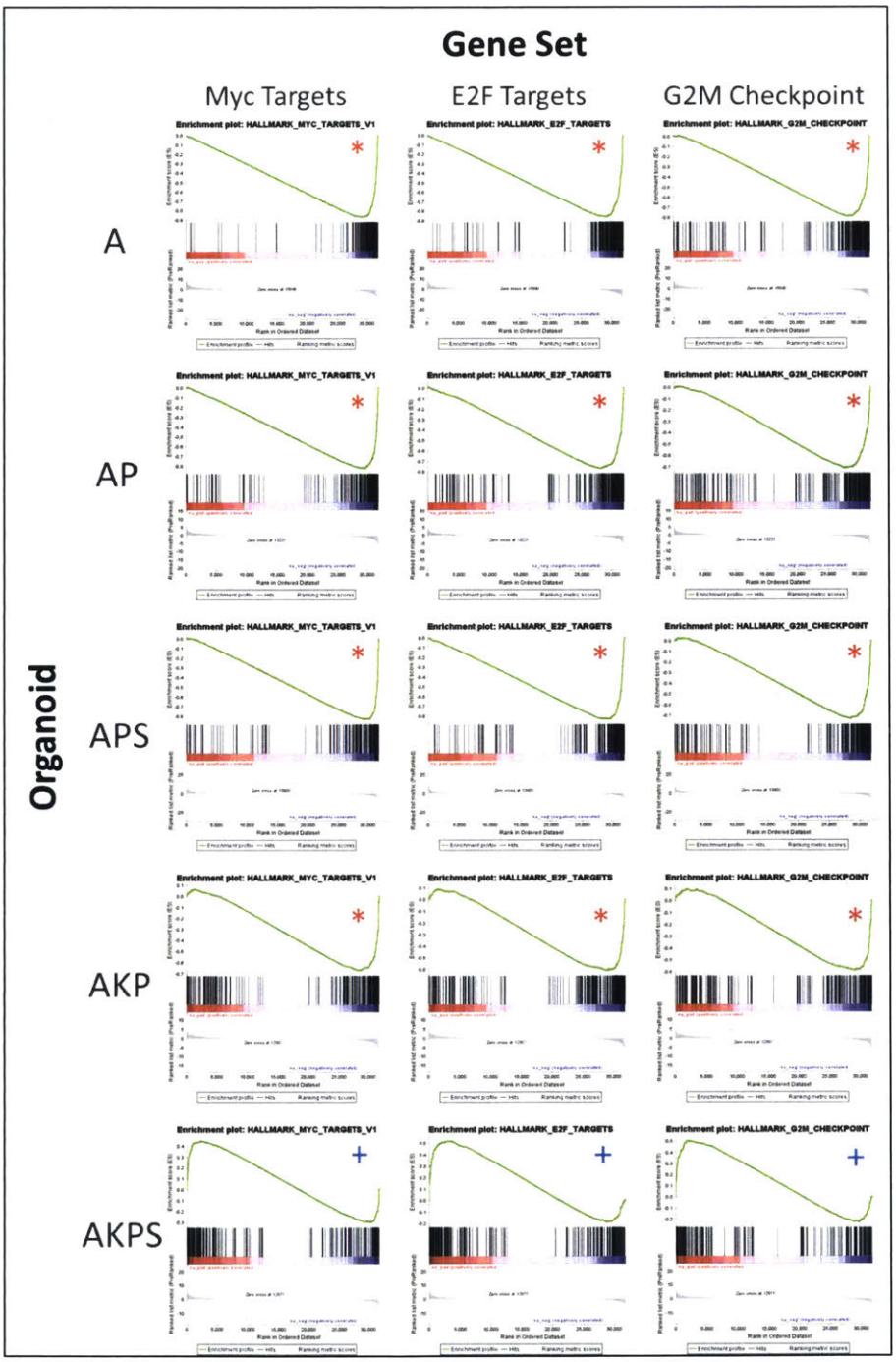


Figure 3.13 – GSEA results for three gene sets related to cell-cycle progression. GSEA analysis was run on a pre-ranked list of differentially expressed genes comparing mCetux treatment vs. no treatment in all of the genotypes. Red (*) indicates enrichment for decreased expression in the presence of mCetux with FDR < 0.0005. Blue (+) indicates enrichment for increased expression with FDR < 0.05.

Because one of the intended applications of mCetux is helping to design rational treatment combinations with immunotherapies, the RNAseq data was analyzed for effects on secreted immune factors such as cytokines and chemokines and their receptors. Changes in expression patterns within this data set could be useful for generating hypotheses about which immunotherapeutic strategies hold the most promise and which genotypes are best poised to respond. Unsupervised clustering of expression data for the subsets of genes corresponding to interleukins (Figure 3.14 A), other cytokines (Figure 3.14 B) and chemokines (Figure 3.14 C) are shown. Overall a striking feature is the low level or total lack of expression for most of the genes queried. However the expression patterns for the various genotypes were distinct enough to generally result in clustering of organoids by type, though the effect of treatment is harder to discern bioinformatically. Particular genes of interest were identified; the focus was on genes in which mCetux treatment caused a significant change in sensitive organoids (A and APS), and where the direction of the effect was opposite that of a K-Ras mutation; the hypothesis was that this expression pattern would be more likely to indicate the change was a function of EGFR signaling. Four such genes are shown in Figure 3.14 D, including three soluble factors: IL-18, CCL2, CCL6, and one receptor: IL-17RE. Further experiments are necessary to determine what, if any, biological underpinnings or therapeutic value these differentially expressed genes represent, but nevertheless this data illustrates the ways in which mCetux can be used to help better understand and guide use of anti-EGFR therapy such as cetuximab.

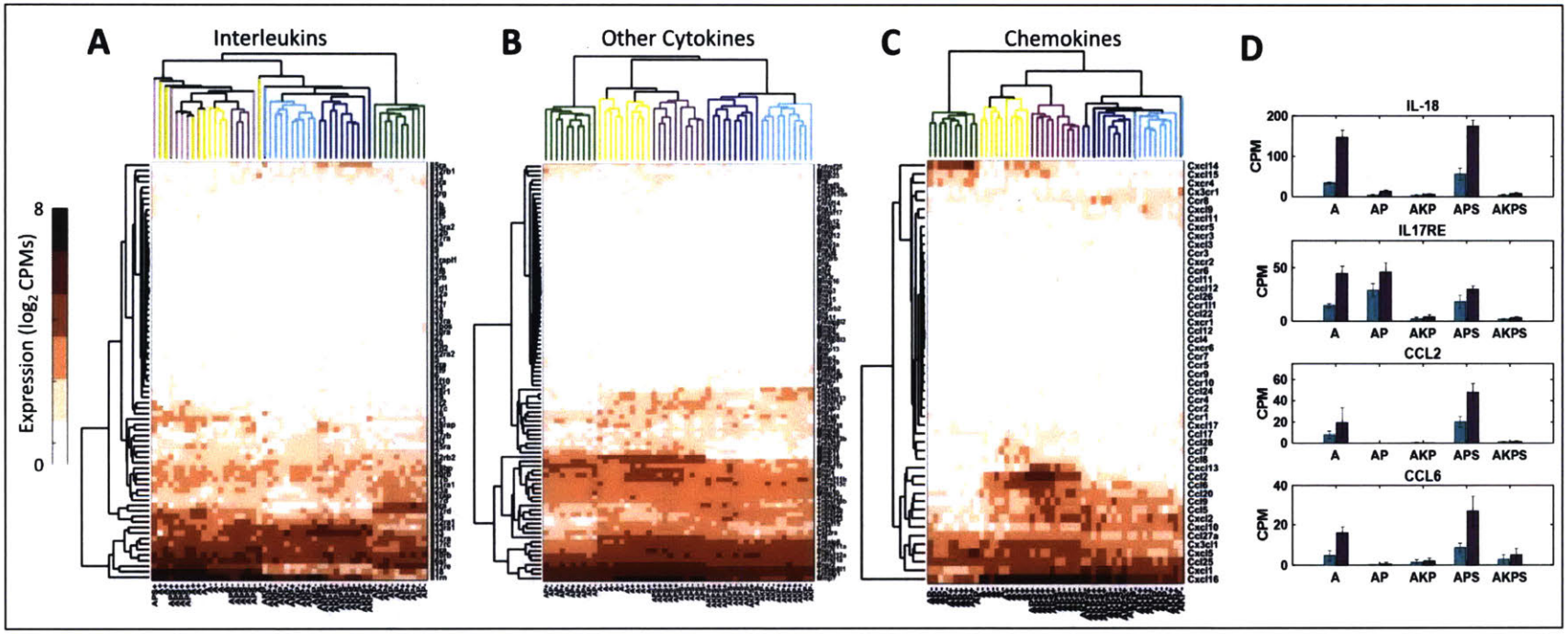


Figure 3.14 – Clustering of subsets of genes corresponding to interleukins (**A**), other cytokines (**B**) and chemokines (**C**). Log2 normalized CPMs were analyzed using unsupervised clustering in Matlab. Color along the top row corresponds to genotype: A – yellow; AP – green; APS – purple; AKP – cyan; AKPS – blue.

In vivo efficacy of mCetux against APS-tdT colorectal organoids

Finally, in order to demonstrate the activity of the mCetux antibody in vivo we tested its ability to control growth of APS organoids. The organoids were inoculated subcutaneously on the flank so that growth would be easier to follow over time in this small pilot treatment study. Two mCetux doses were tested, a 100ug dose and 1000ug dose, each given three times per week by intraperitoneal injection. As discussed above, the unusually high dose of 1000ug was chosen based on the allometric scaling argument and the supporting results of pharmacokinetics and tumor distribution studies.

On day 0, 20 mice were inoculated with 5,000 APS-tdT organoids (roughly equivalent to 3-5E5 cells), and after 2 weeks only very small nodules were visible in 16 of the mice. These growth kinetics are very different than those observed with classical tumor cell lines, where visible masses are found after 4-6 days. The organoid growth continued to remain slow throughout the study, with the PBS group tumors barely increasing in size over the course of the five weeks of treatment. However it was still clear that the mCetux antibody had a dose-dependent effect on organoid growth, with four of the five mice clearing their tumors in the high dose group (Figure 3.15). None of the mice exhibited obvious signs of toxicity as measured by weight (data not shown).

While the use of APS-tdT organoids in a mouse flank was clearly not the most relevant model system, this simple study illustrates that mCetux can control EGFR-driven tumors in vivo, and validates the use of the allometrically-scaled, if unusually high, antibody dose in animal studies. Further, the studies shown here lay the groundwork for expanding the scope of investigation to combination immunotherapies and the oncogenic mutational landscapes in which they might be most effective.

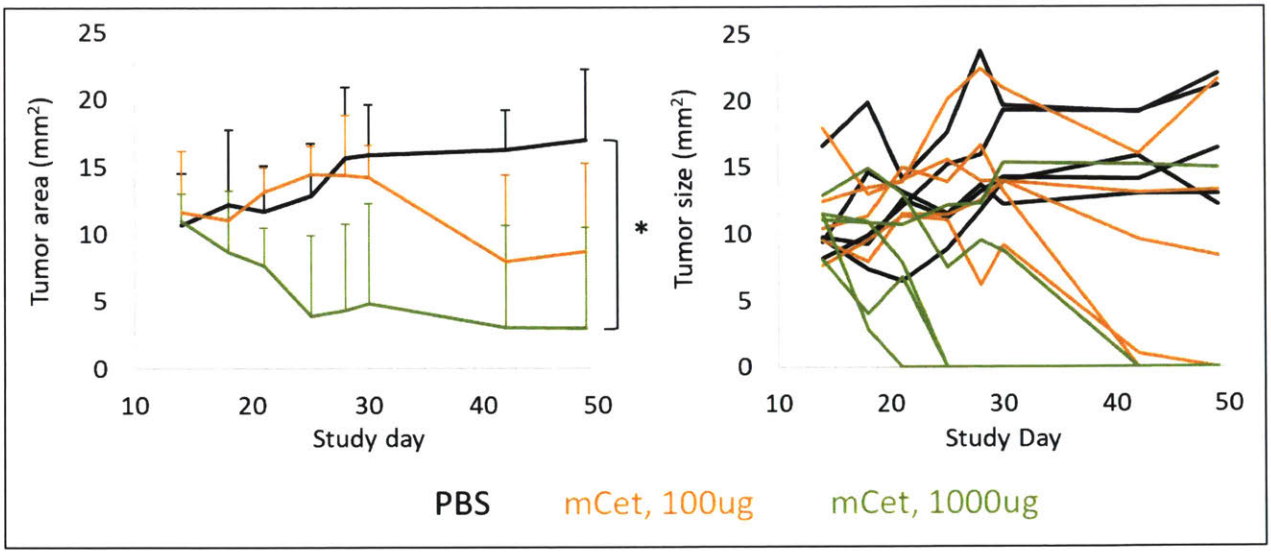


Figure 3.15 – In vivo efficacy of mCetux against APS organoids. Mice were inoculated in the right flank with 5,000 orgnaoids (approximately 500,000 cells) on day 0. Treatments were started when tumor nodules were barely visible, on day 14. Treatments were administered i.p. 3x weekly at indicated doses of antibody. Individual mouse growth curves are shown at right and averages on left. Asterisk indicates $p<0.05$.

Discussion

Cetuximab remains a critical component of standard-of-care therapy in colorectal and head and neck cancers, yet treatment typically only results in benefit for a minority of patients. Even those who initially respond often develop acquired mutations which inactivate therapy. But increased understanding of both the limitations of cetuximab antibody treatment specifically and of the immune-related implications of EGFR signaling in general has opened the door for considering new ways to potentiate cetuximab therapy. In this work we sought to address a significant limitation impeding progress toward new advances in this space: the inability to study fundamental biology or therapy in a native signaling context in an immune proficient mouse. To that end we endeavored to create a mouse cross-reactive version of cetuximab.

The engineering approach chosen was yeast surface display-based screening of a cetuximab mutant library. In particular, a structure-guided saturation mutagenesis approach was used as the initial approach to confer mouse binding, followed by randomized mutations through error-prone PCR to enhance affinity. An alternative method for generating a ligand-blocking murine EGFR antibody would be screening a naïve library or immunizing an animal, as was done for 7A7, the only other reported cetuximab mimetic. But in these cases the selection output, and in particular the binding epitope can not be easily controlled. EGFR ligands all bind generally the same epitope but stimulate divergent signaling through different downstream pathways, a phenomenon that is attributed to subtle conformational changes. If this is true then it could also be that two antibodies binding nominally the same site on the receptor could exert different biological functions; our protein engineering strategy minimizes this risk by producing an antibody highly similar in sequence to cetuximab.

The mCetux molecule created was only four point mutations different from cetuximab. The four amino acids were not tested for their individual contributions to binding, so it is possible that even fewer than all four are required for mouse cross-reactivity; indeed one mutation was all that was required to confer some initial mouse EGFR binding. Fortunately, though not an explicit selection criteria, human EGFR binding was maintained by mCetux, likely because none of the antibody contact residues implicated in binding based on the Fab/EGFR crystal structure were altered. This feature of mCetux should increase its translational relevance, since it means that results obtained from a mouse study can be tested or confirmed in a human context without changing reagents.

The mCetux antibody clearly recapitulates all of the critical functions and properties of cetuximab as measured by in vitro and in vivo results. Receptor binding, ligand blocking and signal antagonism were all preserved, and the antibody was subject to dose-dependent pharmacokinetics and saturable clearance rates that closely mimicked those of cetuximab in humans. Allometric scaling calculations indicated a very high dosing regimen would be required to overcome the pharmacokinetic challenge presented by endogenously expressed murine EGFR, and this hypothesis was validated in a study of mCetux against oncogenic colorectal organoids, where the best tumor control was achieved by systemically administering 1mg of antibody three times per week.

The potential applications for mCetux are numerous. For one, the antibody can be utilized in vitro in any murine system analogous to that of cetuximab in a human system. One example demonstrated here is the use of mCetux for understanding the effect of different oncogenic driver mutations on EGFR signaling in colorectal cancer. Probing a genetically-defined mouse system with cetuximab-like inhibition and examining phenotypic response and

RNA expression levels revealed genotype-specific differences. Particularly interesting was the unusual behavior of the organoids harboring APC and p53 mutations, an oncogenic combination harbored by approximately 30% of CRC patients (Figure 3.16). These were clearly sensitive to EGFR-inhibition, but to a very different extent than the other K-Ras wild type organoids tested. Their RNA expression profiles were also highly distinct, both over the entire RNAseq expression dataset as well among the cytokine and chemokine gene subsets. Further experiments are required to understand the biological meaning of these findings, but importantly, any derived hypotheses are straightforward to test in vivo since mCetux can bind EGFR in a wild type mouse. Overall, one hope is that these types of studies will provide valuable insight into human patient subsets that might be best poised to respond to cetuximab therapy.

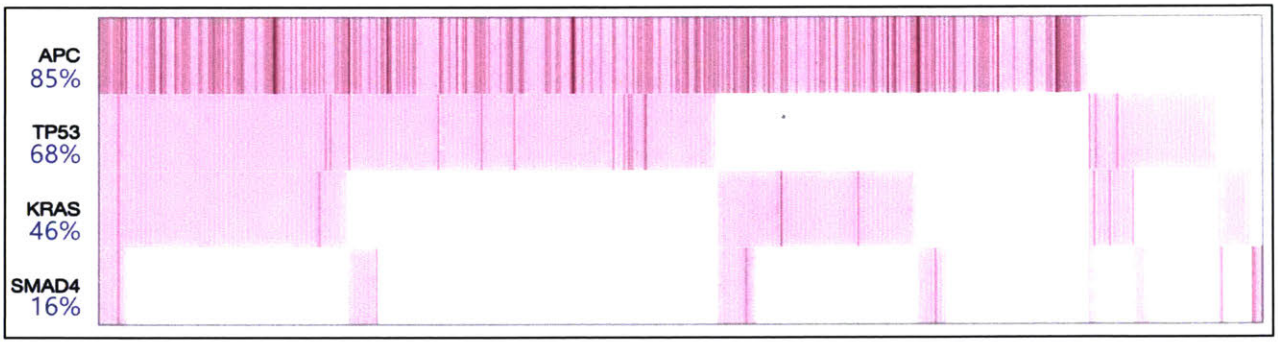


Figure 3.16 – TCGA data showing the prevalence of APC, p53, K-Ras and Smad4 mutations in colorectal cancer patients. TCGA data accessed 6/1/18.

Combinations of cetuximab with immunotherapies such as checkpoint inhibitors[64], systemic cytokines[159], and NK cell agonists[160] are currently being investigated; but these studies all rely on either in vitro experiments or clinical trials. Another obvious, straightforward and exciting application of the mCetux antibody is to make the middle ground – pre-clinical investigations in mice – accessible. Essentially, the antibody could be used to predict the outcome of a human clinical trial, or at minimum inform its design. Of course, the inherent shortcomings of animal models and their sometimes poor translation to humans would still

remain as potential pitfalls for any of these experiments, but in as much as mouse studies have the potential to yield important results, mCetux enables those results' discovery.

Materials and Methods

Reagents

Chicken anti-cmyc antibody was purchased from Exalpha Biologics, Inc. (Shirley, MA). Streptavidin-APC, goat anti-human-AF647, and anti-chicken IgY-AF488 were purchased from ThermoFisher Scientific (Carlsbad, CA).

Protein production

The human and murine EGFR extracellular domains were expressed as fusions to human IgG1 Fc to make bivalent antigen, or to an engineered monomeric human Fc domain as previously described[118]. mCetux and cetuximab were expressed either as human IgG1 antibodies or as mouse IgG2c antibodies as indicated in text. All recombinant proteins were expressed by transient transfection of HEK293F cells and purified by affinity chromatography over rProtein A Sepharose Fast Flow resin (GE Healthcare), then validated by SDS-PAGE and size-exclusion chromatography. When needed, proteins were biotinylated using EZ-Link Sulfo-NHS-Biotin reagent (ThermoFisher Scientific) according to manufacturer's instructions.

Yeast display library design and selections

The cetuximab variable domain sequences were cloned into the pCTCon2 vector as an scFv between the N-terminal HA and C-terminal cmyc tags. Mutagenized scFv fragments were created using primers with degenerate codons NDT / VHG / TGG at a 12:9:1 mixture at each mutation site. Full length inserts were assembled via overlap/extension PCR before transformation into yeast with digested pCTCon2 backbone by electroporation as previously described[120]. Libraries all had >1E8 transformants. All yeast culture protocols were as previously described[119,120]. Sorts were done by labeling with chicken anti-cmyc antibody and Fc-containing EGFR antigen for >1hr at

room temperature, followed by washing and incubating with anti-chicken-488 and anti-human-AF647 secondary antibodies for 30min at 4°C. FACS selections were done using a BD FACS Aria sorter (BD Biosciences).

Mammalian cell culture

A431, CHO, CT26, 4T1 and B16F10 cells were all purchased from ATCC and cultured in media as recommended by ATCC. All media and supplements were purchased from either ATCC or ThermoFisher Scinetific. B16-hEGFR and MC38-hEGFR cells were provided by Dr. Yang-Xin Fu (University of Texas, Southwestern). These cells express stably transfected version of murine EGFR which has six point mutations: K353R, G418S, R443K, M467I, N468S, and K473N, all of which are reversions to human sequence and confer cetuximab binding. Ag104a cells[161] were a gift from Dr. Hans Schreiber (University of Chicago).

EGFR binding and competition assays

For assays on yeast cells, yeast transformed with the mCetux or cetuximab scFv were induced and incubated at 20°C overnight for expression, then washed in PBS+0.1% BSA (PBSA). Cells were incubated in PBSA in 96-well plates in the presence or absence of biotinylated EGFR and/or competitor at indicated concentrations, along with chicken anti-cmyc antibody for >1 hr at room temperature. After washing, cells were incubated for 30min at 4°C with anti-human-AF647 anti-chicken-AF488 for secondary labeling. Cells were washed again before reading on BD Accuri C6 flow cytometer (BD Biosciences).

For mammalian cell binding cells were trypsinized, washed and resuspended in PBSA with biotinylated antibodies at indicated concentrations for >1hr at room temperature. After washing,

cells were incubated with streptavidin-APC for 30min at 4°C. Cells were washed once more before resuspension in PBSA and reading on a BD Accuri flow cytometer.

RT-qPCR

RNA was isolated from cultured cells using NucleoSpin RNA kit (Clontech). RT-PCR was done using QuantiTect SYBR Green RT-PCR Kit 200 (Qiagen) according to manufacturer's instructions and with 200ng template RNA per reaction. Reactions were run in triplicate on a Roche LightCycler 480 Instrument. Forward and reverse primers for human and murine EGFR and GAPDH are listed in Appendix A.

Pharmacokinetic and imaging studies

C57Bl/6 mice were inoculated with 1E6 B16-hEGFR cells subcutaneously on the left flank. On day 7 the mice were given a single i.p. injection of antibody. Each animal received an amount of Alexa Fluor-647 labeled antibody equal to 0.7nmol of fluorophore, with the remainder of the antibody consisting of unlabeled antibody. Antibodies were produced with mouse IgG2c Fc domains (sequences in Appendix A). The dose groups were 2000, 1000, 500, 200 or 50ug mCetux, or 1000, 300, 100 or 50ug cetuximab, 2 mice per group. For PK blood was collected from the tail vein into capillary tubes at indicated time points, sealed with parafilm and stored at 4°C until samples were read at the end of the study. Capillary fluorescence was quantified on a Typhoon imaging system (GE Healthcare) and analyzed using ImageJ software. Samples were normalized against a standard curve of labeled antibody diluted in PBS.

For microscopy studies, tumors were inoculated as above and antibody was injected on day 7. Doses were 2000, 500, 100 or 50ug Alexa Fluor-647 labeled mCetux as above. At 24 hours, animals were euthanized, tumors harvested and fixed. Frozen sections 7 um thick were cut and slides were mounted with VectaShield containing DAPI. Images were captured on a Nikon A1R

Ultra-Fast Spectral Scanning Confocal microscope. All imaging parameters and exposure time were identical for all images to enable direct comparison.

Tumor studies

All animal work was conducted under the approval of the Massachusetts Institute of Technology (MIT) Division of Comparative Medicine in accordance with federal, state, and local guidelines. Male C57/Bl6 mice were obtained from Jackson Labs. At 7-8 weeks of age. 5,000 APS-tdT organoids in X% matrigel were injected subcutaneously in the right flank. Treatment with mCetux at either 1mg or 100ug was administered 3x/week i.p. Treatments were started on day 14 and continued for the length of the study, totaling 5 weeks. Tumors were measured with digital calipers and areas were calculated as LxW.

Chapter 4 – Conclusions and Future Outloook

In this thesis two projects were presented which leveraged the protein engineering expertise of the Wittrup lab against the deficiencies of EGFR-targeted antibody therapy in cancer. In the first project a novel therapeutic fusion protein was generated which combined the cetuximab antibody and engineered EGFR-targeting Sso7d binding moieties. This fusion was a more potent antagonist in a number of in vitro assays, but failed to inhibit tumor growth in vivo (detailed discussion in Chapter 2). In the second project, the cetuximab antibody was "murinized" to enable its use as a pre-clinical tool in murine models. The potential applications of mCetux were demonstrated in a mouse colorectal organoid model system.

A few key themes emerged from these studies. The first of these is the difficulty of employing an EGFR-targeted therapy in the confounding context of normal, endogenous target expression. EGFR was originally designated as an attractive therapeutic target because of its frequent overexpression on cancerous cells, but its presence on healthy tissues throughout the body presents considerable therapeutic challenges. One manifestation of this is an unfavorable pharmacokinetic profile, owing to the presence of the target both within and outside of the tumor, as was encountered in each of the two projects detailed in this thesis. In both cases, the ability of the administered protein to cross-react with native (murine) EGFR dramatically reduced its circulating half-life compared to the non-cross reactive cetuximab antibody alone. And this problem was even further confounded when the engineered agent – the antibody-Sso7d fusion – induced receptor downregulation and resulted in rapid turnover of the drug.

Overcoming this problem requires substantially increasing dosage (a feasible but undesirable solution since production scale-up corresponds with increased technical difficulty and cost), but even then the presence of endogenous EGFR expression poses a second problem –

on-target toxicity. Dysregulated EGFR signaling is oncogenic, but native EGFR signaling is critical for epithelial homeostasis and blockade of these pathways impacts normal tissue function; indeed, skin toxicities are a noted common side-effect of EGFR antibody treatment in humans. Systemically administered agents which target wild-type proteins will access targets both within and outside the tumor, and any potentiation of activity at the tumor site will correspond with increased activity in the surrounding healthy tissues. Overall then, this work drives home the idea that targeting wild-type EGFR by relying on simple overexpression at the tumor site is not a promising strategy. As an alternative, targeting aberrant forms of the receptor, as is achieved by the mAb 806 antibody against the EGFRviii mutant, for example, may provide more promise.

Another related theme of this work was the importance of testing EGFR-targeted therapies in murine models which capture these intuitively straightforward but nevertheless critical aspects of therapeutic development. By all in vitro measures, the antibody-Sso7d fusion was a more potent EGFR antagonist, and the incorporation of receptor downregulating capability was an asset. If this therapeutic had lacked mouse cross-reactivity and had been tested in a human xenograft model it likely would have inhibited tumor growth; but this result would mask potential impending translational difficulty, since in a human context a human EGFR-targeted antibody-Sso7d fusion would encounter target both inside and outside the tumor, likely leading to the difficulties described above. Though it was accidental and led to a therapeutic "failure", the fact that the cetuximab-Sso7d fusion tested in this work was murine cross-reactive allowed for identification of this context-dependent limitation early. Instead of reinforcing the belief that EGFR downregulation is advantageous because it decreases oncogenic EGFR signaling, these studies demonstrated some of the other challenges to be considered if downregulation is to be

pursued. Indeed, the necessity for more faithful models of EGFR biology and therapy is what underlies the significance of the mCetux project. This minimally-modified antibody is the closest mimic of cetuximab function in a mouse that exists, and it should be useful in helping to elucidate all of the complexities of EGFR targeting in a pre-clinical setting.

Despite the failure of the work described here to yield promising new cancer therapies, these projects were highly successful on a protein engineering level. As detailed in Chapter 2 the application of simple in silico methods to the discovery of Sso7d binders was highly fruitful, and lessons were learned about the intrinsic benefits and limitations of the Sso7d scaffold. Further, an intuitive but non-standard, structure-based saturation mutagenesis approach was useful for producing a minimally-modified cetuximab mimic antibody by yeast display. This reagent should have numerous fundamental and therapeutic applications in the field of EGFR inhibition.

Works Cited

1. Krall, J. A., Beyer, E. M. & MacBeath, G. High- and low-affinity epidermal growth factor receptor-ligand interactions activate distinct signaling pathways. *PLoS One* **6,** 1–10 (2011).

2. Harris, R. C., Chung, E. & Coffey, R. J. EGF receptor ligands. *Exp. Cell Res.* **284,** 2–13 (2003).

3. Kaszuba, K. *et al. N* -Glycosylation as determinant of epidermal growth factor receptor conformation in membranes. *Proc. Natl. Acad. Sci.* **112,** 4334–4339 (2015).

4. Azimzadeh Irani, M., Kannan, S. & Verma, C. Role of N-glycosylation in EGFR ectodomain ligand binding. *Proteins Struct. Funct. Bioinforma.* **85,** 1529–1549 (2017).

5. Whitson, K. B. *et al.* Functional effects of glycosylation at Asn-579 of the epidermal growth factor receptor. *Biochemistry* **44,** 14920–14931 (2005).

6. Takahashi, M. *et al.* N-glycan of ErbB family plays a crucial role in dimer formation and tumor promotion. *Biochim. Biophys. Acta* **1780,** 520–524 (2008).

7. Liu, Y.-C. *et al.* Sialylation and fucosylation of epidermal growth factor receptor suppress its dimerization and activation in lung cancer cells. *Proc. Natl. Acad. Sci.* **108,** 11332–11337 (2011).

8. Ferguson, K. M. A structure-based view of Epidermal Growth Factor Receptor regulation. *Annu. Rev. Biophys.* **37,** 353–373 (2008).

9. Ogiso, H. *et al.* Crystal structure of the complex of human epidermal growth factor and receptor extracellular domains. *Cell* **110,** 775–787 (2002).

10. Ferguson, K. M. Active and inactive conformations of the epidermal growth factor receptor. *Biochem. Soc. Trans.* **32,** 742–745 (2004).

11. Wilson, K. J., Gilmore, J. L., Foley, J., Lemmon, M. A. & Ii, D. J. R. Pharmacology & Therapeutics Functional selectivity of EGF family peptide growth factors : Implications for cancer. *Pharmacol. Ther.* **122,** 1–8 (2009).

12. Bessman, N. J., Bagchi, A., Ferguson, K. M. & Lemmon, M. A. Complex relationship between ligand binding and dimerization in the epidermal growth factor receptor. **9,** 1306–1317 (2015).

13. Lemmon, M. A. & Schlessinger, J. Cell signaling by receptor tyrosine kinases. *Cell* **141,** 1117–1134 (2010).

14. Chung, I. *et al.* Spatial control of EGF receptor activation by reversible dimerization on living cells. *Nature* **464,** 783–787 (2010).

15. Lemmon, M. A., Schlessinger, J. & Ferguson, K. M. The EGFR Family: Not So Prototypical Receptor Tyrosine Kinases. *Cold Spring Harb. Perspect. Biol.* **6,** (2014).

16. Endres, N. F. *et al.* Conformational coupling across the plasma membrane in activation of

the EGF receptor. *Cell* **152,** 543–556 (2013).

17. Erba, E. B. *et al.* Quantitation of multisite EGF receptor phosphorylation using mass spectrometry and a novel normalization approach. *J. Proteome Res.* **6,** 2768–2785 (2007).

18. Yarden, Y. & Sliwkowski, M. X. Untangling the ErbB signalling network. *Nat. Rev. Mol. Cell Biol.* **2,** 127–137 (2001).

19. Carpenter, G. The EGF receptor: A nexus for trafficking and signaling. *BioEssays* **22,** 697–707 (2000).

20. Herbst, R. S. & Shin, D. M. Monoclonal antibodies to target epidermal growth factor receptor-positive tumors a new paradigm for cancer therapy. *Cancer* **94,** 1593–1611 (2002).

21. Kumar, V. L., Majumder, P. K., Gujral, S. & Kumar, V. Comparative analysis of epidermal growth factor receptor mRNA levels in normal, benign hyperplastic and carcinomatous prostate. *Cancer Lett.* **134,** 177–180 (1998).

22. Lorenzo, G. Di *et al.* Expression of Epidermal Growth Factor Receptor Correlates with Disease Relapse and Progression to Androgen-independence in Human Prostate Cancer Expression of Epidermal Growth Factor Receptor Correlates with Disease Relapse and Progression to Androgen-ind. *Clin. Cancer Res.* **8,** 3438–3444 (2002).

23. Citri, A. & Yarden, Y. EGF-ERBB signalling: Towards the systems level. *Nat. Rev. Mol. Cell Biol.* **7,** 505–516 (2006).

24. Irmer, D., Funk, J. O. & Blaukat, A. EGFR kinase domain mutations - Functional impact and relevance for lung cancer therapy. *Oncogene* **26,** 5693–5701 (2007).

25. Ji, H. *et al.* The impact of human EGFR kinase domain mutations on lung tumorigenesis and in vivo sensitivity to EGFR-targeted therapies. *Cancer Cell* **9,** 485–495 (2006).

26. Shigematsu, H. *et al.* Clinical and biological features associated with epidermal growth factor receptor gene mutations in lung cancers. *J. Natl. Cancer Inst.* **97,** 339–346 (2005).

27. Pao, W. & Miller, V. A. Epidermal growth factor receptor mutations, small-molecule kinase inhibitors, and non-small-cell lung cancer: Current knowledge and future directions. *J. Clin. Oncol.* **23,** 2556–2568 (2005).

28. Sugawa, N., Ekstrand, A. J., James, C. D. & Collins, V. P. Identical splicing of aberrant epidermal growth factor receptor transcripts from amplified rearranged genes in human glioblastomas. *Proc. Natl. Acad. Sci.* **87,** 8602–8606 (1990).

29. Wong, A. J. *et al.* Structural alterations of the epidermal growth factor receptor gene in human gliomas. *Proc. Natl. Acad. Sci.* **89,** 2965–2969 (1992).

30. Frederick, L., Wang, X. Y., Eley, G. & James, C. D. Diversity and frequency of epidermal growth factor receptor mutations in human glioblastomas. *Cancer Res.* **60,** 1383–1387 (2000).

31. Garcia de Palazzo, I. E. *et al.* Expression of mutated epidermal growth factor receptor by non-small cell lung carcinomas. *Cancer Res.* **53,** 3217–3220 (1993).

32. Moscatello, D. K. *et al.* Frequent Expression of a Mutant Epidermal Growth Factor Receptor in Multiple Human Tumors. *Cancer Res.* **55,** 5536–5539 (1995).

33. Mendelsohn, J. & Baselga, J. Epidermal Growth Factor Receptor Targeting in Cancer. *Semin. Oncol.* **33,** 369–385 (2006).

34. Chong, C. R. & Jänne, P. a. The quest to overcome resistance to EGFR-targeted therapies in cancer. *Nat. Med.* **19,** 1389–400 (2013).

35. Cross, D. A. E. *et al.* AZD9291, an irreversible EGFR TKI, overcomes T790M-mediated resistance to EGFR inhibitors in lung cancer. *Cancer Discov.* **4,** 1046–1061 (2014).

36. Sunada, H., Magunt, B. E., Mendelsohn, J. & Macleod, C. L. Monoclonal antibody against epidermal growth factor receptor is internalized without stimulating receptor phosphorylation. *Proc. Natl. Acad. Sci.* **83,** 3825–3829 (1986).

37. Sunada, H., Yu, P., Peacock, J. S. & Mendelsohn, J. Modulation of Tyrosine, Serine, and Threonine Phosphorylation and lntracellular Processing of the Epidermal Growth Factor Receptor by Antireceptor Monoclonal Antibody. *J. Cell. Physiol.* **142,** 284–292 (1990).

38. Doody, J. F. *et al.* Inhibitory activity of cetuximab on epidermal growth factor receptor mutations in non – small cell lung cancers. *Mol. Cancer Ther.* **6,** 2642–2652 (2007).

39. Montagut, C. *et al.* Identification of a mutation in the extracellular domain of the Epidermal Growth Factor Receptor conferring cetuximab resistance in colorectal cancer. *Nat. Med.* **18,** 221–224 (2012).

40. Bonner, J. A. *et al.* Radiotherapy plus Cetuximab for Squamous-Cell Carcinoma of the Head and Neck. *N. Engl. J. Med.* **354,** 567–578 (2006).

41. Cunningham, D. *et al.* Cetuximab monotherapy and cetuximab plus irinotecan in irinotecan-refractory metastatic colorectal cancer. *N. Engl. J. Med.* **351,** 337–345 (2004).

42. Vermorken, J. B. *et al.* Platinum-based chemotherapy plus cetuximab in head and neck cancer. *N. Engl. J. Med.* **359,** 1116–1127 (2008).

43. Lynch, T. J. *et al.* Cetuximab and first-line taxane/carboplatin chemotherapy in advanced non-small-cell lung cancer: Results of the randomized multicenter phase III trial BMS099. *J. Clin. Oncol.* **28,** 911–917 (2010).

44. Pirker, R. *et al.* Cetuximab plus chemotherapy in patients with advanced non-small-cell lung cancer (FLEX): an open-label randomised phase III trial. *Lancet* **373,** 1525–1531 (2009).

45. Fakih, M. & Vincent, M. Adverse events associated with anti-EGFR therapies for the treatment of metastatic colorectal cancer. *Curr. Oncol.* **17,** S18–S30 (2010).

46. Hecht, J. R. *et al.* Panitumumab monotherapy in patients with previously treated metastatic colorectal cancer. *Cancer* **110,** 980–988 (2007).

47. Saltz, L. B. *et al.* Phase II trial of cetuximab in patients with refractory colorectal cancer that expresses the epidermal growth factor receptor. *J. Clin. Oncol.* **22,** 1201–1208 (2004).

48. Li, S. *et al.* Structural basis for inhibition of the epidermal growth factor receptor by cetuximab. *Cancer Cell* **7,** 301–311 (2005).

49. Carter, P. *et al.* Humanization of an anti-p185HER2 antibody for human cancer therapy. *Proc. Natl. Acad. Sci.* **89,** 4285–4289 (1992).

50. Reff, B. M. E. *et al.* Depletion of B Cells In Vivo by a Chimeric Mouse Human Monoclonal Antibody to CD20. *Blood* **83,** 435–445 (1994).

51. Fan, Z., Masui, H., Altas, I. & Mendelsohn, J. Blockade of Epidermal Growth Factor Receptor Function by Bivalent and Monovalent Fragments of 225 Anti-Epidermal Growth Factor Receptor Monoclonal Antibodies. *Cancer Res.* **53,** 4322–4329 (1993).

52. Kurai, J. *et al.* Antibody-dependent cellular cytotoxicity mediated by cetuximab against lung cancer cell lines. *Clin. Cancer Res.* **13,** 1552–1561 (2007).

53. Wilson, N. S. *et al.* Article An Fc gamma Receptor-Dependent Mechanism Drives in Cancer Cells. *Cancer Cell* **19,** 101–113 (2011).

54. Bibeau, F. *et al.* Impact of Fc gamma RIIa-Fc gamma RIIIa Polymorphisms and KRAS Mutations on the Clinical Outcome of Patients With Metastatic Colorectal Cancer Treated With Cetuximab Plus Irinotecan. *J. Clin. Oncol.* **27,** 1122–1129 (2009).

55. Zhang, W. *et al.* FCGR2A and FCGR3A Polymorphisms Associated With Clinical Outcome of Epidermal Growth Factor Receptor – Expressing Metastatic Colorectal Cancer Patients Treated With Single-Agent Cetuximab. *J. Clin. Oncol.* **25,** 3712–3718 (2007).

56. Trivedi, S. *et al.* Anti-EGFR targeted monoclonal antibody isotype influences antitumor cellular immunity in head and neck cancer patients. *Clin. Cancer Res.* **22,** 5229–5237 (2016).

57. Garrido, G. *et al.* T cells are crucial for the anti-metastatic effect of anti-epidermal growth factor receptor antibodies. *Cancer Immunol. Immunother.* **56,** 1701–1710 (2007).

58. Yang, X. *et al.* Cetuximab-mediated tumor regression depends on innate and adaptive immune responses. *Mol. Ther.* **21,** 91–100 (2013).

59. Pozzi, C. *et al.* The EGFR-specific antibody cetuximab combined with chemotherapy triggers immunogenic cell death. *Nat. Med.* **22,** 624–631 (2016).

60. Pollack, B. P., Sapkota, B. & Cartee, T. V. Epidermal Growth Factor Receptor Inhibition Augments the Expression of MHC Class I and II Genes. *Clin. Cancer Res.* **17,** 4400–4414 (2011).

61. Akbay, E. A. *et al.* Activation of the PD-1 Pathway Contributes to Immune Escape in EGFR-Driven Lung Tumors. *Cancer Discov.* **3,** 1355–1364 (2013).

62. Li, C. *et al.* Glycosylation and stabilization of programmed death ligand-1 suppresses T-cell activity. *Nat. Commun.* **7,** 1–11 (2016).

63. Hippe, A. *et al.* Tumor immune escape by the loss of homeostatic chemokine expression. *Proc. Natl. Acad. Sci.* **104,** 19055–19060 (2007).

64. Ferris, R. L. *et al.* Rationale for combination of therapeutic antibodies targeting tumor cells and immune checkpoint receptors : Harnessing innate and adaptive immunity through IgG1 isotype immune effector stimulation. *Cancer Treat. Rev.* **63,** 48–60 (2018).

65. Phillips, A. C. *et al.* ABT-414, an Antibody-Drug Conjugate Targeting a Tumor-Selective EGFR Epitope. *Mol. Cancer Ther.* **128,** 438–45 (2016).

66. Johnson, L. A. *et al.* Rational development and characterization of humanized anti-EGFR variant III chimeric antigen receptor T cells for glioblastoma. *Sci. Transl. Med.* **7,** 275ra22 (2015).

67. Fajardo, C. A. *et al.* Oncolytic adenoviral delivery of an EGFR-targeting t-cell engager improves antitumor efficacy. *Cancer Res.* **77,** 2052–2063 (2017).

68. Nagane, M., Lin, H., Cavenee, W. K. & Huang, H. J. Aberrant receptor signaling in human malignant gliomas: mechanisms and therapeutic implications. *Cancer Lett.* **162 Suppl,** S17–S21 (2001).

69. Spangler, J. B. *et al.* Combination antibody treatment down-regulates epidermal growth factor receptor by inhibiting endosomal recycling. *Proc. Natl. Acad. Sci.* **107,** 13252–13257 (2010).

70. Spangler, J. B., Manzari, M. T., Rosalia, E. K., Chen, T. F. & Wittrup, K. D. Triepitopic antibody fusions inhibit cetuximab-resistant BRAF and KRAS mutant tumors via EGFR signal repression. *J. Mol. Biol.* **422,** 532–544 (2012).

71. Chen, W. W. *et al.* Input-output behavior of ErbB signaling pathways as revealed by a mass action model trained against dynamic data. *Mol. Syst. Biol.* **5,** 239 (2009).

72. Schoeberl, B., Eichler-Jonsson, C., Gilles, E. D. & Müller, G. Computational modeling of the dynamics of the MAP kinase cascade activated by surface and internalized EGF receptors. *Nat. Biotechnol.* **20,** 370–375 (2002).

73. Yarden, Y. & Pines, G. The ERBB network: at last, cancer therapy meets systems biology. *Nat. Rev. Cancer* **12,** 553–563 (2012).

74. Kearns, J. D. *et al.* Enhanced Targeting of the EGFR Network with MM-151, an Oligoclonal Anti-EGFR Antibody Therapeutic. *Mol. Cancer Ther.* **14,** 1625–36 (2015).

75. Monast, C. S. & Lazzara, M. J. Identifying Determinants of EGFR-Targeted Therapeutic Biochemical Efficacy Using Computational Modeling. *CPT pharmacometrics Syst. Pharmacol.* **3,** e141 (2014).

76. Hackel, B. J., Kapila, A. & Dane Wittrup, K. Picomolar Affinity Fibronectin Domains Engineered Utilizing Loop Length Diversity, Recursive Mutagenesis, and Loop Shuffling. *J. Mol. Biol.* **381,** 1238–1252 (2008).

77. Parker, M. H. *et al.* Antibody mimics based on human fibronectin type three domain engineered for thermostability and high-affinity binding to vascular endothelial growth factor receptor two. *Protein Eng. Des. Sel.* **18,** 435–444 (2005).

78. Pankov, R. & Yamada, K. M. Fibronectin at a glance. *J. Cell Sci.* **115,** 3861–3863 (2002).

79. Singh, P., Carraher, C. & Schwarzbauer, J. E. Assembly of fibronectin into extracellular matrix. *Annu. Rev. Cell Dev. Biol.* **26,** 397–419 (2010).

80. Knapp, S. *et al.* Thermal unfolding of the DNA-binding protein Sso7d from the hyperthermophile Sulfolobus solfataricus. *J. Mol. Biol.* **264,** 1132–1144 (1996).

81. Gera, N., Hussain, M., Wright, R. C. & Rao, B. M. Highly stable binding proteins derived from the hyperthermophilic Sso7d scaffold. *J. Mol. Biol.* **409,** 601–616 (2011).

82. Traxlmayr, M. W. *et al.* Strong Enrichment of Aromatic Residues in Binding Sites from a Charge-Neutralized Hyperthermostable Sso7d Scaffold Library. *Revis.* 130 (2016). doi:10.1016/j.immuni.2009.01.001

83. Kauke, M. J. *et al.* An engineered protein antagonist of K-Ras / B-Raf interaction. *Sci. Rep.* **7,** 5831–5839 (2017).

84. Thurber, G. M., Zajic, S. C. & Wittrup, K. D. Theoretic Criteria for Antibody Penetration into Solid Tumors and Micrometastases. 995–1000 (2007). doi:10.2967/jnumed.106.037069

85. Schmidt, M. M. & Wittrup, K. D. A modeling analysis of the effects of molecular size and binding affinity on tumor targeting. *Mol. Cancer Ther.* **8,** 2861–71 (2009).

86. Rhoden, J. J. & Wittrup, K. D. Dose Dependence of Intratumoral Perivascular Distribution of Monoclonal Antibodies. *J. Pharm. Sci.* **101,** 860–867 (2012).

87. Hackel, B. J., Neil, J. R., White, F. M. & Wittrup, K. D. Epidermal growth factor receptor downregulation by small heterodimeric binding proteins. *Protein Eng. Des. Sel.* **25,** 47–57 (2012).

88. Li, J. Y. *et al.* Article A Biparatopic HER2-Targeting Antibody-Drug Conjugate Induces Tumor Regression in Primary Models Refractory to or Ineligible for HER2-Targeted Therapy. *Cancer Cell* **29,** 117–129 (2016).

89. Oliveira, S. *et al.* Downregulation of EGFR by a novel multivalent nanobody-liposome platform. *J. Control. Release* **145,** 165–175 (2010).

90. LaFleur, D. W. *et al.* Monoclonal antibody therapeutics with up to five specificities. *MAbs* **5,** 208–218 (2013).

91. Machiels, J. P. *et al.* A proof of concept trial of the anti - EGFR antibody mixture Sym004 in patients with squamous cell carcinoma of the head and neck. *Cancer Chemother. Pharmacol.* **76,** 13–20 (2015).

92. Marabelle, A., Tselikas, L., Baere, T. De & Houot, R. Intratumoral immunotherapy : using the tumor as the remedy. *Ann. Oncol.* **28,** 33–43 (2017).

93. Miller, E. A., Traxlmayr, M. W., Shen, J. & Sikes, H. D. Activity-based assessment of an engineered hyperthermophilic protein as a capture agent in paper-based diagnostic tests. *Mol. Syst. Des. Eng.* **1,** 377–381 (2016).

94. Bogan, A. A. & Thorn, K. S. Anatomy of hot spots in protein interfaces. *J. Mol. Biol.* **280,** 1–9 (1998).

95. Moreira, I. S., Fernandes, P. A. & Ramos, M. J. Hot spots—A review of the protein – protein interface determinant amino-acid residues. *Proteins Struct. Funct. Bioinforma.* **68,** 803–812 (2007).

96. Ma, B., Wolfson, H. J. & Nussinov, R. Protein functional epitopes : hot spots , dynamics and combinatorial libraries. *Curr. Opin. Struct. Biol.* **11,** 364–369 (2001).

97. Fleury, D., Daniels, R. S., Skehel, J. J., Knossow, M. & Bizebard, T. Structural Evidence for Recognition of a Single Epitope by Two Distinct Antibodies. *Proteins Struct. Funct. Genet.* **40,** 572–578 (2000).

98. Wind, T., Jensen, M. A. & Andreasen, P. A. Epitope mapping for four monoclonal antibodies against human plasminogen activator inhibitor type-1. *Eur. J. Biochem.* **268,** 1095–1106 (2001).

99. Scott, J. K. *et al.* Evidence that a protein-protein interaction 'hot spot' on heterotrimeric G protein beta-gamma subunits is used for recognition of a subclass of effectors. *EMBO J.* **20,** 767–776 (2001).

100. Nakamura, G. R., Starovasnik, M. A., Reynolds, M. E. & Lowman, H. B. A Novel Family of Hairpin Peptides That Inhibit IgE Activity by Binding to the High-Affinity IgE Receptor. *Biochemistry* **40,** 9828–9835 (2001).

101. Skelton, N. J. *et al.* Structure - Function Analysis of a Phage Display-Derived Peptide That Binds to Insulin-like Growth Factor Binding Protein 1. *Biochemistry* **40,** 8487–8498 (2001).

102. Tikhomirov, O. Y. & Thomas, J. W. Alanine Scanning Mutants of Rat Proinsulin I Show Functional Diversity of Anti-Insulin Monoclonal Antibodies. *J. Immunol.* **165,** 3876–3882 (2000).

103. Schmiedel, J., Blaukat, A., Li, S., Knoechel, T. & Ferguson, K. M. Matuzumab binding to EGFR prevents the conformational rearrangement required for dimerization. *Cancer Cell* **13,** 365–373 (2008).

104. Roovers, R. C. *et al.* A bi-paratopic anti-EGFR nanobody efficiently inhibits solid tumor growth. *Int. J. Cancer* **129,** 2013–2024 (2011).

105. Boersma, Y. L., Chao, G., Steiner, D., Wittrup, K. D. & Plückthun, A. Bispecific Designed Ankyrin Repeat Proteins (DARPins) Targeting Epidermal Growth Factor Receptor Inhibit A431 Cell Proliferation and Receptor Recycling. *J. Biol. Chem.* **286,** 41273–41285 (2011).

106. Stave, J. W. & Lindpaintner, K. Antibody and Antigen Contact Residues Define Epitope and Paratope Size and Structure. *J. Immunol.* **191,** 1428–1435 (2013).

107. Peng, H., Hao, K. H., Jian, J. & Yang, A. Origins of specificity and affinity in antibody – protein interactions. *Proc. Natl. Acad. Sci.* **111,** E2656–E2665 (2014).

108. Boder, E. T., Midelfort, K. S. & Wittrup, K. D. Directed evolution of antibody fragments with monovalent femtomolar antigen-binding affinity. *Proc. Natl. Acad. Sci.* **97,** 10701–10705 (2000).

109. Orlova, A. *et al.* Tumor Imaging Using a Picomolar Affinity HER2 Binding Affibody Molecule. *Cancer Res.* **66,** 4339–4349 (2006).

110. Ravn, U. *et al.* Deep sequencing of phage display libraries to support antibody discovery. *Methods* **60,** 99–110 (2013).

111. Fantini, M. *et al.* Assessment of antibody library diversity through next generation sequencing and technical error compensation. *PLoS One* **12,** e0177574 (2017).

112. Hu, D. *et al.* Effective optimization of antibody affinity by phage display integrated with high-throughput DNA synthesis and sequencing technologies. *PLoS One* **10,** 1–17 (2015).

113. Van Blarcom, T. *et al.* Precise and efficient antibody epitope determination through library design, yeast display and next-generation sequencing. *J. Mol. Biol.* **427,** 1513–1534 (2015).

114. Kowalsky, C. A. *et al.* Rapid fine conformational epitope mapping using comprehensive mutagenesis and deep sequencing. *J. Biol. Chem.* **290,** 26457–26470 (2015).

115. Domina, M. *et al.* Rapid profiling of the antigen regions recognized by serum antibodies using massively parallel sequencing of antigen-specific libraries. *PLoS One* **9,** 1–18 (2014).

116. Yang, W. *et al.* Next-generation sequencing enables the discovery of more diverse positive clones from a phage-displayed antibody library. *J. Exp. Mol. Med.* **49,** e308–e316 (2017).

117. Glanville, J. *et al.* Deep sequencing in library selection projects: what insight does it bring? *Curr. Opin. Struct. Biol.* **33,** 146–160 (2015).

118. Ishino, T. *et al.* Engineering a monomeric Fc domain modality by N-glycosylation for the half-life extension of biotherapeutics. *J. Biol. Chem.* **288,** 16529–16537 (2013).

119. Boder, E. T. & Wittrup, K. D. Yeast surface display for screening combinatorial polypeptide libraries. *Nat. Biotechnol.* **15,** 553–557 (1997).

120. Chao, G. *et al.* Isolating and engineering human antibodies using yeast surface display. *Nat. Protoc.* **1,** 755–68 (2006).

121. Harms, B. D., Bassi, G. M., Horwitz, A. R. & Lauffenburger, D. A. Directional Persistence of EGF-Induced Cell Migration Is Associated with Stabilization of Lamellipodial Protrusions. *Biophys. J.* **88,** 1479–1488 (2005).

122. Douillard, J.-Y. *et al.* Panitumumab–FOLFOX4 Treatment and RAS Mutations in Colorectal Cancer. *N. Engl. J. Med.* **369,** 1023–1034 (2013).

123. De Roock, W. *et al.* on the effi cacy of cetuximab plus chemotherapy in chemotherapy-refractory metastatic colorectal cancer : a retrospective consortium analysis. *Lancet Oncol.* **11,** 753–762 (2010).

124. Allegra, C. J. *et al.* American Society of Clinical Oncology Provisional Clinical Opinion : Testing for KRAS Gene Mutations in Patients With Metastatic Colorectal Carcinoma to Predict Response to Anti – Epidermal Growth Factor Receptor Monoclonal Antibody

Therapy. *J. Clin. Oncol.* **27,** 2091–2096 (2009).

125. Bertotti, A. *et al.* The Genomic Landscape of Response to EGFR Blockade in Colorectal Cancer. *Nature* **8,** 263–267 (2015).

126. Chen, D., Huang, X., Wery, J. & Li, Q. A set of defined oncogenic mutation alleles seems to better predict the response to cetuximab in CRC patient-derived xenograft than KRAS 12 / 13 mutations. **6,**

127. Song, J. I. & Grandis, J. R. STAT signaling in head and neck cancer. *Oncogene* **19,** 2489–2495 (2000).

128. Bardelli, A. & Siena, S. Molecular mechanisms of resistance to cetuximab and panitumumab in colorectal cancer. *J. Clin. Oncol.* **28,** 1254–1261 (2010).

129. Bardelli, A. & Jänne, P. A. The road to resistance: EGFR mutation and cetuximab. *Nat. Med.* **18,** 199–200 (2012).

130. Therkildsen, C., Bergmann, T. K., Henrichsen-schnack, T. & Nilbert, M. The predictive value of KRAS, NRAS, BRAF, PIK3CA and PTEN for anti-EGFR treatment in metastatic colorectal cancer : A systematic review and meta- analysis. *Acta Oncol. (Madr).* **53,** 852–864 (2014).

131. Bronte, G. *et al.* New findings on primary and acquired resistance to anti-EGFR therapy in metastatic colorectal cancer : do all roads lead to RAS? *Oncotarget* **6,** (2015).

132. Xu, J. *et al.* PIK3CA mutations contribute to acquired cetuximab resistance in metastatic colorectal cancer patients. *Clin. Cancer Res.* **23,** 4602–4616 (2017).

133. Yonesaka, K. *et al.* Activation of ERBB2 signaling causes resistance to the EGFR-directed therapeutic antibody cetuximab. *Sci. Transl. Med.* **3,** 99ra86 (2012).

134. Baselga, J. *et al.* Phase I studies of anti-epidermal growth factor receptor chimeric antibody C225 alone and in combination with cisplatin. *J. Clin. Oncol.* **18,** 904–914 (2000).

135. Tan, A. R. *et al.* Pharmacokinetics of cetuximab after administration of escalating single dosing and weekly fixed dosing in patients with solid tumors. *Clin. Cancer Res.* **12,** 6517–6522 (2006).

136. Baker, J. H. E. *et al.* Heterogeneous distribution of trastuzumab in HER2-positive xenografts and metastases : role of the tumor microenvironment. *Clin. Exp. Metastasis* **35,** 691–705 (2018).

137. de Boer, E. *et al.* In Vivo Fluorescence Immunohistochemistry: Localization of Fluorescently Labeled Cetuximab in Squamous Cell Carcinomas. *Sci. Rep.* **5,** 1–11 (2015).

138. Azzopardi, N. *et al.* Cetuximab pharmacokinetics influences progression-free survival of metastatic colorectal cancer patients. *Clin. Cancer Res.* **17,** 6329–6337 (2011).

139. Larkin, J. *et al.* Combined Nivolumab and Ipilimumab or Monotherapy in Untreated Melanoma. *N. Engl. J. Med.* **373,** 23–24 (2015).

140. Antonia, S. J. *et al.* Durvalumab after Chemoradiotherapy in Stage III Non–Small-Cell Lung Cancer. *N. Engl. J. Med.* **377,** 1919–1929 (2017).

141. Moynihan, K. D. *et al.* Eradication of large established tumors in mice by combination immunotherapy that engages innate and adaptive immune responses. *Nat. Med.* **22,** 1402–1410 (2016).

142. Zhu, E. F. *et al.* Synergistic innate and adaptive immune response to combination immunotherapy with anti-tumor antigen antibodies and extended serum half-life il-2. *Cancer Cell* **27,** 489–501 (2015).

143. Kimura, H. *et al.* Antibody-dependent cellular cytotoxicity of cetuximab against tumor cells with wild-type or mutant epidermal growth factor receptor. *Cancer Sci.* **98,** 1275–1280 (2007).

144. Jie, H. *et al.* Increased PD-1+ and TIM-3+ TILs during Cetuximab Therapy Inversely Correlate with Response in Head and Neck Cancer Patients. *Cancer Immunol. Res.* **5,** 1–10 (2017).

145. Macdonald, F. & Zaiss, D. M. W. The Immune System's Contribution to the Clinical Efficacy of EGFR Antagonist Treatment. *Front. Pharmacol.* **8,** 1–7 (2017).

146. Concha-Benavente, F. *et al.* Identification of the cell-intrinsic and -extrinsic pathways downstream of EGFR and IFNγ that induce PD-L1 expression in head and neck Cancer. *Cancer Res.* **76,** 1031–1043 (2016).

147. Pollack, B. P. EGFR inhibitors, MHC expression and immune responses. *Oncoimmunology* **1,** 71–74 (2012).

148. Tang, H. *et al.* Facilitating T Cell Infiltration in Tumor Microenvironment Overcomes Resistance to PD-L1 Blockade. *Cancer Cell* **29,** 285–296 (2016).

149. Concha-Benavente, F. & Ferris, R. L. Reversing EGFR mediated immunoescape by targeted monoclonal antibody therapy. *Front. Pharmacol.* **8,** 1–8 (2017).

150. Srivastava, R. M. *et al.* CD137 stimulation enhances cetuximab-induced natural killer: Dendritic cell priming of antitumor T-cell immunity in patients with head and neck cancer. *Clin. Cancer Res.* **23,** 707–716 (2017).

151. Veluchamy, J. P. *et al.* Combination of NK Cells and Cetuximab to Enhance Anti-Tumor Responses in RAS Mutant Metastatic Colorectal Cancer. *PLoS One* 1–16 (2016). doi:10.1371/journal.pone.0157830

152. Vanneman, M. & Dranoff, G. Combining immunotherapy and targeted therapies in cancer treatment. *Nat. Rev. Cancer* **12,** 237–251 (2012).

153. Garrido, G. *et al.* 7A7 MAb: a new tool for the pre-clinical evaluation of EGFR-based therapies. *Hybrid Hybridomics* **23,** 168–175 (2004).

154. He, X. *et al.* Characterization of 7A7, an anti-mouse EGFR monoclonal antibody proposed to be the mouse equivalent of cetuximab. *Oncotarget* **9,** 12250–12260 (2018).

155. Nair, A. & Jacob, S. A simple practice guide for dose conversion between animals and

human. *J. Basic Clin. Pharm.* **7,** 27 (2016).

156. Minchinton, A. I. & Tannock, I. F. Drug penetration in solid tumors. *Nat. Rev. Cancer* **6,** 583–592 (2006).

157. Lee, C. M. & Tannock, I. F. The distribution of the therapeutic monoclonal antibodies cetuximab and trastuzumab within solid tumors. *BMC Cancer* **10,** 255 (2010).

158. Walther, A. *et al.* Prognostic and predictive markers in colorectal cancer. *Nat. Rev. Cancer* **9,** 489–499 (2009).

159. Mcmichael, E. L. *et al.* IL-21 Enhances Natural Killer Cell Response to Cetuximab-Coated Pancreatic Tumor Cells. 489–503 (2017). doi:10.1158/1078-0432.CCR-16-0004

160. Kohrt, H. E. *et al.* Targeting CD137 enhances the efficacy of cetuximab. *J. Clin. Invest.* **124,** 2668–2682 (2014).

161. Ward, P. L., Koeppen, H., Hurteau, T. & Schreiber, H. Tumor Antigens Defined by Cloned Immunological Probes are Highly Polymorphic and are not Detected on Autologous Normal Cells. *J. Exp. Med.* **170,** 217–232 (1989).

162. Thurber, G. M., Schmidt, M. M. & Wittrup, K. D. Antibody tumor penetration : Transport opposed by systemic and antigen-mediated clearance ☆. **60,** 1421–1434 (2008).

163. Thurber, G. M., Schmidt, M. M. & Wittrup, K. D. Factors determining antibody distribution in tumors. **29,** (2010).

Appendix A – Plasmid and Primer DNA Sequences

EGFR-Fc in gWiz vector

…gWiz – Signal peptide – EGFR ECD – IEGRMD – Human IgG1 Fc – (Gly4Ser)2 – Biotin Acceptor Peptide – His Tag – STOP – gWiz …

TAATAGCTGACAGACTAACAGACTGTTCCTTTCCATGGGTCTTTTCTGCAGATGCGACCCT
CCGGGACGGCCGGGGCAGCGCTCCTGGCGCTACTGGCTGCGCTCTGCCCGGCTAGCCG
GGCTCTGGAGGAAAAGAAAGTTTGCCAAGGCACGAGTAACAAGCTCACGCAGTTGGGCAC
TTTTGAAGATCATTTTCTCAGCCTCCAGAGGATGTTCAATAACTGTGAGGTGGTCCTTGGG
AATTTGGAAATTACCTATGTGCAGAGGAATTATGATCTTTCCTTCTTAAAGACCATCCAGGA
GGTGGCTGGTTATGTCCTCATTGCCCTCAACACAGTGGAGCGAATTCCTTTGGAAAACCTG
CAGATCATCAGAGGAAATATGTACTACGAAAATTCCTATGCCTTAGCAGTCTTATCTAACTA
TGATGCAAATAAAACCGGACTGAAGGAGCTGCCCATGAGAAATTTACAGGAAATCCTGCAT
GGCGCCGTGCGGTTCAGCAACAACCCTGCCCTGTGCAACGTGGAGAGCATCCAGTGGCG
GGACATAGTCAGCAGTGACTTTCTCAGCAACATGTCGATGGACTTCCAGAACCACCTAGGC
AGCTGCCAAAAGTGTGATCCAAGCTGTCCCAATGGGAGCTGCTGGGGTGCAGGAGAGGA
GAACTGCCAGAAACTGACCAAAATCATCTGTGCCCAGCAGTGCTCCGGGCGCTGCCGTGG
CAAGTCCCCCAGTGACTGCTGCCACAACCAGTGTGCTGCAGGCTGCACAGGCCCCCGGG
AGAGCGACTGCCTGGTCTGCCGCAAATTCCGAGACGAAGCCACGTGCAAGGACACCTGC
CCCCCACTCATGCTCTACAACCCCACCACGTACCAGATGGATGTGAACCCCGAGGGCAAA
TACAGCTTTGGTGCCACCTGCGTGAAGAAGTGTCCCCGTAATTATGTGGTGACAGATCACG
GCTCGTGCGTCCGAGCCTGTGGGGCCGACAGCTATGAGATGGAGGAAGACGGCGTCCGC
AAGTGTAAGAAGTGCGAAGGGCCTTGCCGCAAAGTGTGTAACGGAATAGGTATTGGTGAA
TTTAAAGACTCACTCTCCATAAATGCTACGAATATTAAACACTTCAAAAACTGCACCTCCATC
AGTGGCGATCTCCACATCCTGCCGGTGGCATTTAGGGGTGACTCCTTCACACATACTCCTC
CTCTGGATCCACAGGAACTGGATATTCTGAAAACCGTAAAGGAAATCACAGGGTTTTTGCT
GATTCAGGCTTGGCCTGAAAACAGGACGGACCTCCATGCCTTTGAGAACCTAGAAATCATA
CGCGGCAGGACCAAGCAACATGGTCAGTTTTCTCTTGCAGTCGTCAGCCTGAACATAACAT
CCTTGGGATTACGCTCCCTCAAGGAGATAAGTGATGGAGATGTGATAATTTCAGGAAACAA
AAATTTGTGCTATGCAAATACAATAAACTGGAAAAAACTGTTTGGGACCTCCGGTCAGAAAA
CCAAAATTATAAGCAACAGAGGTGAAAACAGCTGCAAGGCCACAGGCCAGGTCTGCCATG
CCTTGTGCTCCCCCGAGGGCTGCTGGGGCCCGGAGCCCAGGGACTGCGTCTCTTGCCGG
AATGTCAGCCGAGGCAGGGAATGCGTGGACAAGTGCAAACTTCTGGAGGGTGAGCCAAG
GGAGTTTGTGGAGAACTCTGAGTGCATACAGTGCCACCCAGAGTGCCTGCCTCAGGCCAT
GAACATCACCTGCACAGGACGGGGACCAGACAACTGTATCCAGTGTGCCCACTACATTGA
CGGCCCCCACTGCGTCAAGACCTGCCCGGCAGGAGTCATGGGAGAAAACAACACCCTGG
TCTGGAAGTACGCAGACGCCGGCCATGTGTGCCACCTGTGCCATCCAAACTGCACCTACG
GATGCACTGGGCCAGGTCTTGAAGGCTGTCCAACGAATGGGCCTAAGATCCCGTCCATTG
AAGGCCGCATGGATAAGCCCAGCAACACCAAGGTGGACAAGAAAGTTGAGCCCAAATCT
TGTGACAAAACTCACACATGCCCACCGTGCCCAGCACCTGAACTCCTGGGGGGACCGTCA
GTCTTCCTCTTCCCCCCAAAACCCAAGGACACCCTCATGATCTCCCGGACCCCTGAGGTCA
CATGCGTGGTGGTGGACGTGAGCCACGAAGACCCTGAGGTCAAGTTCAACTGGTACGTGG
ACGGCGTGGAGGTGCATAATGCCAAGACAAAGCCGCGGGAGGAGCAGTACAACAGCACG
TACCGTGTGGTCAGCGTCCTCACCGTCCTGCACCAGGACTGGCTGAATGGCAAGGAGTAC
AAGTGCAAGGTCTCCAACAAAGCCCTCCCAGCCCCCATCGAGAAAACCATCTCCAAAGCC
AAAGGGCAGCCCCGAGAACCACAGGTGTACACCCTGCCCCCATCCCGGGATGAGCTGAC
CAAGAACCAGGTCAGCCTGACCTGCCTGGTCAAAGGCTTCTATCCCAGCGACATCGCCGT

GGAGTGGGAGAGCAATGGGCAGCCGGAGAACAACTACAAGACCACGCCTCCCGTGCTGG
ACTCCGACGGCTCCTTCTTCCTCTACAGCAAGCTCACCGTGGACAAGAGCAGGTGGCAGC
AGGGGAACGTCTTCTCATGCTCCGTGATGCATGAGGCTCTGCACAACCACTACACGCAGA
AGAGCCTCTCCCTGTCTCCGGGTAAAGGAGGCGGTGGGTCTGGCGGAGGTGGACAC
GTGCTTAATGACATTTTCGAGGCCCAAAAAATAGAGTGGCATGAAACCGGTCATCA
CCATCACCATCACTGATAAGTCGACACGTGTGATCAGATATCGCGGCCGCTCTAGACCA
MRPSGTAGAALLALLAALCPASRALEEKKVCQGTSNKLTQLGTFEDHFLSLQRMFNNC
EVVLGNLEITYVQRNYDLSFLKTIQEVAGYVLIALNTVERIPLENLQIIRGNMYYENSYA
LAVLSNYDANKTGLKELPMRNLQEILHGAVRFSNNPALCNVESIQWRDIVSSDFLSNMS
MDFQNHLGSCQKCDPSCPNGSCWGAGEENCQKLTKIICAQQCSGRCRGKSPSDCCHNQ
CAAGCTGPRESDCLVCRKFRDEATCKDTCPPLMLYNPTTYQMDVNPEGKYSFGATCVK
KCPRNYVVTDHGSCVRACGADSYEMEEDGVRKCKKCEGPCRKVCNGIGIGEFKDSLSI
NATNIKHFKNCTSISGDLHILPVAFRGDSFTHTPPLDPQELDILKTVKEITGFLLIQAWPEN
RTDLHAFENLEIIRGRTKQHGQFSLAVVSLNITSLGLRSLKEISDGDVIISGNKNLCYANTI
NWKKLFGTSGQKTKIISNRGENSCKATGQVCHALCSPEGCWGPEPRDCVSCRNVSRGR
ECVDKCKLLEGEPREFVENSECIQCHPECLPQAMNITCTGRGPDNCIQCAHYIDGPHCVK
TCPAGVMGENNTLVWKYADAGHVCHLCHPNCTYGCTGPGLEGCPTNGPKIPSIEGRMD
KPSNTKVDKKVEPKSCDKTHTCPPCPAPELLGGPSVFLFPPKPKDTLMISRTPEVTCVVV
DVSHEDPEVKFNWYVDGVEVHNAKTKPREEQYNSTYRVVSVLTVLHQDWLNGKEYK
CKVSNKALPAPIEKTISKAKGQPREPQVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVE
WESNGQPENNYKTTPPVLDSDGSFFLYSKLTVDKSRWQQGNVFSCSVMHEALHNHYT
QKSLSLSPGKGGGGSGGGGHVLNDIFEAQKIEWHETGHHHHHH**

EGFR-monoFc in gWiz vector

This construct introduces two extra glycosylation sites in the Fc domain to stabilize monomeric solubility. For more info, see Ishino, et al[118]. Highlighted locations are mutated from wt Fc.

…gWiz – EGFR signal peptide – EGFR ECD – $(G_4S)_2$ – monoFc (119-330) – AgeI – BAP – MluI – **Stop** – gWiz…

CAGACTAACAGACTGTTCCTTTCCATGGGTCTTTTCTGCAGATGCGACCCTCCGGGA
CGGCCGGGGCAGCGCTCCTGGCGCTACTGGCTGCGCTCTGCCCGGCTAGCCGGGCTC
TGGAGGAAAAGAAAGTTTGCCAAGGCACGAGTAACAAGCTCACGCAGTTGGGCACT
TTTGAAGATCATTTTCTCAGCCTCCAGAGGATGTTCAATAACTGTGAGGTGGTCCTT
GGGAATTTGGAAATTACCTATGTGCAGAGGAATTATGATCTTTCCTTCTTAAAGACC
ATCCAGGAGGTGGCTGGTTATGTCCTCATTGCCCTCAACACAGTGGAGCGAATTCCT
TTGGAAAACCTGCAGATCATCAGAGGAAATATGTACTACGAAAATTCCTATGCCTTA
GCAGTCTTATCTAACTATGATGCAAATAAAACCGGACTGAAGGAGCTGCCCATGAG
AAATTTACAGGAAATCCTGCATGGCGCCGTGCGGTTCAGCAACAACCCTGCCCTGTG
CAACGTGGAGAGCATCCAGTGGCGGGACATAGTCAGCAGTGACTTTCTCAGCAACA
TGTCGATGGACTTCCAGAACCACCTAGGCAGCTGCCAAAAGTGTGATCCAAGCTGTC
CCAATGGGAGCTGCTGGGGTGCAGGAGAGGAGAACTGCCAGAAACTGACCAAAATC
ATCTGTGCCCAGCAGTGCTCCGGGCGCTGCCGTGGCAAGTCCCCCAGTGACTGCTGC
CACAACCAGTGTGCTGCAGGCTGCACAGGCCCCCGGGAGAGCGACTGCCTGGTCTG
CCGCAAATTCCGAGACGAAGCCACGTGCAAGGACACCTGCCCCCCACTCATGCTCT
ACAACCCCACCACGTACCAGATGGATGTGAACCCCGAGGGCAAATACAGCTTTGGT
GCCACCTGCGTGAAGAAGTGTCCCCGTAATTATGTGGTGACAGATCACGGCTCGTGC
GTCCGAGCCTGTGGGGCCGACAGCTATGAGATGGAGGAAGACGGCGTCCGCAAGTG
TAAGAAGTGCGAAGGGCCTTGCCGCAAAGTGTGTAACGGAATAGGTATTGGTGAAT
TTAAAGACTCACTCTCCATAAATGCTACGAATATTAAACACTTCAAAAACTGCACCT
CCATCAGTGGCGATCTCCACATCCTGCCGGTGGCATTTAGGGGTGACTCCTTCACAC
ATACTCCTCCTCTGGATCCACAGGAACTGGATATTCTGAAAACCGTAAAGGAAATCA
CAGGGTTTTTGCTGATTCAGGCTTGGCCTGAAAACAGGACGGACCTCCATGCCTTTG
AGAACCTAGAAATCATACGCGGCAGGACCAAGCAACATGGTCAGTTTTCTCTTGCA
GTCGTCAGCCTGAACATAACATCCTTGGGATTACGCTCCCTCAAGGAGATAAGTGAT
GGAGATGTGATAATTTCAGGAAACAAAAATTTGTGCTATGCAAATACAATAAACTG
GAAAAAACTGTTTGGGACCTCCGGTCAGAAAACCAAAATTATAAGCAACAGAGGTG
AAAACAGCTGCAAGGCCACAGGCCAGGTCTGCCATGCCTTGTGCTCCCCCGAGGGC
TGCTGGGGCCCGGAGCCCAGGGACTGCGTCTCTTGCCGGAATGTCAGCCGAGGCAG
GGAATGCGTGGACAAGTGCAAACTTCTGGAGGGTGAGCCAAGGGAGTTTGTGGAGA
ACTCTGAGTGCATACAGTGCCACCCAGAGTGCCTGCCTCAGGCCATGAACATCACCT
GCACAGGACGGGGACCAGACAACTGTATCCAGTGTGCCCACTACATTGACGGCCCC
CACTGCGTCAAGACCTGCCCGGCAGGAGTCATGGGAGAAAACAACACCCTGGTCTG
GAAGTACGCAGACGCCGGCCATGTGTGCCACCTGTGCCATCCAAACTGCACCTACG
GATGCACTGGGCCAGGTCTTGAAGGCTGTCCAACGAATGGGCCTAAGATCCCGTCC
GGCGGAGGAGGTTCAGGAGGTGGTGGATCTGGGGGACCGTCAGTCTTCCTCTTCCCC
CCAAAACCCAAGGACACCCTCATGATCTCCCGGACCCCTGAGGTCACATGCGTGGT
GGTGGACGTGAGCCACGAAGACCCTGAGGTCAAGTTCAACTGGTACGTGGACGGCG

TGGAGGTGCATAATGCCAAGACAAAGCCGCGGGAGGAGCAGTACAACAGCACGTA
CCGTGTGGTCAGCGTCCTCACCGTCCTGCACCAGGACTGGCTGAATGGCAAGGAGT
ACAAGTGCAAGGTCTCCAACAAAGCCCTCCCAGCCCCCATCGAGAAAACCATCTCC
AAAGCCAAAGGGCAGCCCCGAGAACCACAGGTGTACACCCTGCCCCCATCCCGGGA
TGAGCTGACCAAGAACCAGGTCAACCTGACCTGCCTGGTCAAAGGCTTCTATCCCAG
CGACATCGCCGTGGAGTGGGAGAGCAATGGGCAGCCGGAGAACAACTACAAGACC
ACGCCTCCCGTGCTGGACTCCGACGGCTCCTTCTTCCTCAACAGCACGCTCACCGTG
GACAAGAGCAGGTGGCAGCAGGGGAACGTCTTCTCATGCTCCGTGATGCATGAGGC
TCTGCACAACCACTACACGCAGAAGAGCCTCTCCCTGTCTCCGGGTAAAACCGGTCT
TAATGACATTTTCGAGGCCCAAAAAATAGAGTGGCATGAAACGCGT**TGATAA**GTCG
ACACGTGTGATCAGATATCGCGGCCGCTC

MRPSGTAGAALLALLAALCPASRALEEKKVCQGTSNKLTQLGTFEDHFLSLQRMFNNC
EVVLGNLEITYVQRNYDLSFLKTIQEVAGYVLIALNTVERIPLENLQIIRGNMYYENSYA
LAVLSNYDANKTGLKELPMRNLQEILHGAVRFSNNPALCNVESIQWRDIVSSDFLSNMS
MDFQNHLGSCQKCDPSCPNGSCWGAGEENCQKLTKIICAQQCSGRCRGKSPSDCCHNQ
CAAGCTGPRESDCLVCRKFRDEATCKDTCPPLMLYNPTTYQMDVNPEGKYSFGATCVK
KCPRNYVVTDHGSCVRACGADSYEMEEDGVRKCKKCEGPCRKVCNGIGIGEFKDSLSI
NATNIKHFKNCTSISGDLHILPVAFRGDSFTHTPPLDPQELDILKTVKEITGFLLIQAWPEN
RTDLHAFENLEIIRGRTKQHGQFSLAVVSLNITSLGLRSLKEISDGDVIISGNKNLCYANTI
NWKKLFGTSGQKTKIISNRGENSCKATGQVCHALCSPEGCWGPEPRDCVSCRNVSRGR
ECVDKCKLLEGEPREFVENSECIQCHPECLPQAMNITCTGRGPDNCIQCAHYIDGPHCVK
TCPAGVMGENNTLVWKYADAGHVCHLCHPNCTYGCTGPGLEGCPTNGPKIPSGGGGS
GGGGSGGPSVFLFPPKPKDTLMISRTPEVTCVVVDVSHEDPEVKFNWYVDGVEVHNAK
TKPREEQYNSTYRVVSVLTVLHQDWLNGKEYKCKVSNKALPAPIEKTISKAKGQPREPQ
VYTLPPSRDELTKNQVNLTCLVKGFYPSDIAVEWESNGQPENNYKTTPPVLDSDGSFFL
NSTLTVDKSRWQQGNVFSCSVMHEALHNHYTQKSLSLSPGK**TG**LNDIFEAQKIEWHE**T**
R**

Sso7d clones

Sso7d sequences were expressed either in pE-SUMO or in pCTCon2.

Clone A

GCAACCGTGAAATTCACATACCAAGGCGAAGAAAAACAGGTGGATATTAGCAAAAT
CAAGATCGTGGACCGTTACGGCCAGTCTATTCATTTTAACTATGATGAAGGTGGTGG
TGCCTATAGTTGGGTTTGGATGAGCGAAAAAGATGCACCGAAAGAACTGCTGCTGA
TGCTGGAAAAGCGA

ATVKFTYQGEEKQVDISKIKIVDRYGQSIHFNYDEGGGAYSWVWMSEKDAPKELLLML
EKR

Clone B

GCAACCGTAAAATTCACATACCAAGGCGGAGAAAAACAGGTGGATATTAGCAAAAT
CATAAACGTGCTGCGTTACGGCCAGCTGGTTGTTTTTTCTTATGATGAAGGTGGTGG
TGCCGGCGGTAACGGTTTCGTGAGCGAAAAGGATGCGCCGAAAGAACTGCTGCAGA
TGCTGGAAAAGCGA

ATVKFTYQGGEKQVDISKIINVLRYGQLVVFSYDEGGGAGGNGFVSEKDAPKELLQML
EKR

Clone C

GCATCCGTGAAATTCGCATACCAAGGCGAAGAAAAACAGGTGGATATTAGCAAAAT
CAAGCATATGTGGCGTTGGGGCCAGTTTATTATGTTTGAATATGATGAAGGTGGTGG
TGCCTGGGGTTTGGGTCAAGTGAGCGAAAAAGATGCACCGAAAGAGCTGCTGCAGA
TGCTGTGGAGGCAA

ASVKFAYQGEEKQVDISKIKHMWRWGQFIMFEYDEGGGAWGLGQVSEKDAPKELLQM
LWRQ

Clone D

GCAACCGTGAAATTCACATACCAAGGCAAAGAAAGACAGGTGGATATTAGCAAAAT
CACGCGGGCGTACCGTATTGGCCAGTGGATTGATTTTGCGTATGATGAAGGTGGGGG
GTGGTGGGGTTGGGGTAATGTGAGCGAAAAAGATGCACCGAAAGAACTGCTGCAGA
TGCTGGAGGAGCGA

ATVKFTYQGKERQVDISKITRAYRIGQWIDFAYDEGGGWWGWGNVSEKDAPKELLQM
LEEK

Clone E

GTAACCGTGGGATTCACATACCAAGGCGGAGTAAAACAGGTGGATATTAGCAAAAT
CAAGATAGTGAGGCGTTGGGGCCAGTGGATTGGTTTTTACTATGATGAAGGTGGTGG
TGCCAAAGGTTGGGGTGATGTGAGGGAAAAAGATGCACCGAAAGAACTGCTGCAGA
TGCTGGCAAAGCAA

VTVGFTYQGGVKQVDISKIKIVRRWGQWIGFYYDEGGGAKGWGDVREKDAPKELLQM
LAKR

Clone F

GCAACCGTGAAATTCACATACCAAGGCGAAGAAAAACAGGTGGATATTAGCAAAAT
CAAGGGAGTGGCACGTTCCGGCCAGAACATTATCTTTTCTTATGATGAAGGTGGTGG
TGCCTATGGTAACGGTGACGTGAGCGAAAAAGATGCACCGAAAGAACTGCTGCAGA
TGCTGGAAAAGCAA

ATVKFTYQGEEKQVDISKIKGVARSGQNIIFSYDEGGGAYGNGDVSEKDAPKELLQMLE
KQ

DH-LD in gWiz

Sequences for DH-LD are representative of all cetuximab-Sso7d fusions.

…gWiz – LC Leader – Sso7d – G4S – cetux LC – **Stop** – gWiz…

CAGACTGTTCCTTTCCATGGGTCTTTTCTGCAGGCCGCCACCATGAGGGTCCCCGCTC
AGCTCCTGGGGCTCCTGCTGCTCTGGCTCCCAGGTGCACGATGTGCAACCGTGAAAT
TCACATACCAAGGCGAAGAAAAACAGGTGGATATTAGCAAAATCAAGATCGTGGAC
CGTTACGGCCAGTCTATTCATTTTAACTATGATGAAGGTGGTGGTGCCTATAGTTGG
GTTTGGATGAGCGAAAAAGATGCACCGAAAGAACTGCTGCTGATGCTGGAAAAGCG
AGGAGGCGGAGGGTCGGACATCCTGCTGACCCAGTCTCCAGTCATCCTGTCTGTGAG
TCCAGGAGAAAGAGTCAGTTTCTCCTGCAGGGCCAGTCAGAGTATTGGCACAAACA
TACACTGGTATCAGCAAAGAACAAATGGTTCTCCAAGGCTTCTCATAAAGTATGCTT
CTGAGTCTATCTCTGGCATCCCTTCCAGGTTTAGTGGCAGTGGATCAGGGACAGATT
TTACTCTTAGCATCAACAGTGTGGAGTCTGAAGATATTGCAGATTATTACTGTCAAC
AAAATAATAACTGGCCAACCACGTTCGGTGCTGGGACCAAGCTGGAGCTCAAACGT
ACGGTGGCTGCACCATCTGTCTTCATCTTCCCGCCATCTGATGAGCAGTTGAAATCT
GGAACTGCCTCTGTTGTGTGCCTGCTGAATAACTTCTATCCCAGAGAGGCCAAAGTA
CAGTGGAAGGTGGATAACGCCCTCCAATCGGGTAACTCCCAGGAGAGTGTCACAGA
GCAGGACAGCAAGGACAGCACCTACAGCCTCAGCAGCACCCTGACGCTGAGCAAAG
CAGACTACGAGAAACACAAAGTCTACGCCTGCGAAGTCACCCATCAGGGCCTGAGC
TCGCCCGTCACAAAGAGCTTCAACAGGGGAGAGTGT**TAATAG**GTCGACACGTGTGA
TCAGATATCGCGG

QTVPFHGSFLQAATMRVPAQLLGLLLLWLPGARCATVKFTYQGEEKQVDISKIKIVDRY
GQSIHFNYDEGGGAYSWVWMSEKDAPKELLLMLEKRGGGGSDILLTQSPVILSVSPGER
VSFSCRASQSIGTNIHWYQQRTNGSPRLLIKYASESISGIPSRFSGSGSGTDFTLSINSVESE
DIADYYCQQNNNWPTTFGAGTKLELKRTVAAPSVFIFPPSDEQLKSGTASVVCLLNNFY
PREAKVQWKVDNALQSGNSQESVTEQDSKDSTYSLSSTLTLSKADYEKHKVYACEVTH
QGLSSPVTKSFNRGEC**VDTCDQISR

…gWiz – HC Leader – cetux HC – G4S – Sso7d– **Stop** – gWiz…

CCTTTCCATGGGTCTTTTCTGCAGGCCGCCACCATGGGTTGGAGCCTCATCTTGCTCT
TCCTTGTCGCTGTTGCTACGCGTGTACTGAGCCAGGTACAACTGAAGCAGTCAGGAC
CTGGCCTAGTGCAGCCCTCACAGAGCCTGTCCATCACCTGCACAGTCTCTGGTTTCT
CATTAACTAACTATGGTGTACACTGGGTTCGCCAGTCTCCAGGAAAGGGTCTGGAGT
GGCTGGGAGTGATATGGAGTGGTGGAAACACAGACTATAATACACCTTTCACATCC
AGACTGAGCATCAACAAGGACAATTCCAAGAGCCAAGTTTTCTTTAAAATGAACAG
TCTGCAATCTAATGACACAGCCATATATTACTGTGCCAGAGCCCTCACCTACTATGA
TTACGAGTTTGCTTACTGGGGCCAAGGGACCCTGGTCACCGTTTCCGCTGCTAGCAC
CAAGGGCCCATCGGTCTTCCCCCTGGCACCCTCCTCCAAGAGCACCTCTGGGGGCAC
AGCGGCCCTGGGCTGCCTGGTCAAGGACTACTTCCCCGAACCGGTGACGGTGTCGTG
GAACTCAGGCGCCCTGACCAGCGGCGTGCACACCTTCCCGGCTGTCCTACAGTCCTC
AGGACTCTACTCCCTCAGCAGCGTGGTGACCGTGCCCTCCAGCAGCTTGGGCACCCA

GACCTACATCTGCAACGTGAATCACAAGCCCAGCAACACCAAGGTGGACAAGAAAG
TTGAGCCCAAATCTTGTGACAAAACTCACACATGCCCACCGTGCCCAGCACCTGAAC
TCCTGGGGGGACCGTCAGTCTTCCTCTTCCCCCCAAAACCCAAGGACACCCTCATGA
TCTCCCGGACCCCTGAGGTCACATGCGTGGTGGTGGACGTGAGCCACGAAGACCCT
GAGGTCAAGTTCAACTGGTACGTGGACGGCGTGGAGGTGCATAATGCCAAGACAAA
GCCGCGGGAGGAGCAGTACAACAGCACGTACCGTGTGGTCAGCGTCCTCACCGTCC
TGCACCAGGACTGGCTGAATGGCAAGGAGTACAAGTGCAAGGTCTCCAACAAAGCC
CTCCCAGCCCCCATCGAGAAAACCATCTCCAAAGCCAAAGGGCAGCCCCGAGAACC
ACAGGTGTACACCCTGCCCCCATCCCGGGATGAGCTGACCAAGAACCAGGTCAGCC
TGACCTGCCTGGTCAAAGGCTTCTATCCCAGCGACATCGCCGTGGAGTGGGAGAGC
AATGGGCAGCCGGAGAACAACTACAAGACCACGCCTCCCGTGCTGGACTCCGACGG
CTCCTTCTTCCTCTACAGCAAGCTCACCGTGGACAAGAGCAGGTGGCAGCAGGGGA
ACGTCTTCTCATGCTCCGTGATGCATGAGGCTCTGCACAACCACTACACGCAGAAGA
GCCTCTCCCTGTCTCCGGGTAAAGGAGGCGGAGGGTCGGCAACCGTGAAATTCACA
TACCAAGGCAAAGAAAGACAGGTGGATATTAGCAAAATCACGCGGGCGTACCGTAT
TGGCCAGTGGATTGATTTTGCGTATGATGAAGGTGGGGGGTGGTGGGGTTGGGGTA
ATGTGAGCGAAAAAGATGCACCGAAAGAACTGCTGCAGATGCTGGAGGAGCAA**TA**
ATAGGTCGACACGTGTGATCAGATATCGCGGCCGCTC

PFHGSFLQAATMGWSLILLFLVAVATRVLSQVQLKQSGPGLVQPSQSLSITCTVSGFSLT
NYGVHWVRQSPGKGLEWLGVIWSGGNTDYNTPFTSRLSINKDNSKSQVFFKMNSLQSN
DTAIYYCARALTYYDYEFAYWGQGTLVTVSAASTKGPSVFPLAPSSKSTSGGTAALGCL
VKDYFPEPVTVSWNSGALTSGVHTFPAVLQSSGLYSLSSVVTVPSSSLGTQTYICNVNHK
PSNTKVDKKVEPKSCDKTHTCPPCPAPELLGGPSVFLFPPKPKDTLMISRTPEVTCVVVD
VSHEDPEVKFNWYVDGVEVHNAKTKPREEQYNSTYRVVSVLTVLHQDWLNGKEYKC
KVSNKALPAPIEKTISKAKGQPREPQVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEW
ESNGQPENNYKTTPPVLDSDGSFFLYSKLTVDKSRWQQGNVFSCSVMHEALHNHYTQK
SLSLSPGKGGGGSATVKFTYQGKERQVDISKITRAYRIGQWIDFAYDEGGGWWGWGNV
SEKDAPKELLQMLEEQ**VDTCDQISRPL

Cetuximab scFv in pCTCon2 vector

…(G4S)3 – BamHI – cetuximab scFv – NheI – cmyc – Stop…

GGTGGAGGAGGCTCTGGTGGAGGCGGTAGCGGAGGCGGAGGGTCGGCTAGCGACA
TCCTGCTGACCCAGTCTCCAGTCATCCTGTCTGTGAGTCCAGGAGAAAGAGTCAGTT
TCTCCTGCAGGGCCAGTCAGAGTATTGGCACAAACATACACTGGTATCAGCAAAGA
ACAAATGGTTCTCCAAGGCTTCTCATAAAGTATGCTTCTGAGTCTATCTCTGGCATCC
CTTCCAGGTTTAGTGGCAGTGGATCAGGGACAGATTTTACTCTTAGCATCAACAGTG
TGGAGTCTGAAGATATTGCAGATTATTACTGTCAACAAAATAATAACTGGCCAACCA
CGTTCGGTGCTGGGACCAAGCTGGAGCTCAAACGTACGGTGGCTGGTACTACTGCCG
CTAGTGGTAGTAGTGGTGGCAGTAGCAGTGGTGCCCAGGTACAACTGAAGCAGTCA
GGACCTGGCCTAGTGCAGCCCTCACAGAGCCTGTCCATCACCTGCACAGTCTCTGGT
TTCTCATTAACTAACTATGGTGTACACTGGGTTCGCCAGTCTCCAGGAAAGGGTCTG
GAGTGGCTGGGAGTGATATGGAGTGGTGGAAACACAGACTATAATACACCTTTCAC
ATCCAGACTGAGCATCAACAAGGACAATTCCAAGAGCCAAGTTTTCTTTAAAATGA
ACAGTCTGCAATCTAATGACACAGCCATATATTACTGTGCCAGAGCCCTCACCTACT
ATGATTACGAGTTTGCTTACTGGGGCCAAGGGACCCTGGTCACCGTTTCCGCTGGAT
CCGAACAAAAGCTTATTTCTGAAGAGGACTTGTAATAG

GGGGSGGGGSGGGGSASDILLTQSPVILSVSPGERVSFSCRASQSIGTNIHWYQQRTNGS
PRLLIKYASESISGIPSRFSGSGSGTDFTLSINSVESEDIADYYCQQNNNWPTTFGAGTKLE
LKRTVAGTTAASGSSGGSSSGAQVQLKQSGPGLVQPSQSLSITCTVSGFSLTNYGVHWV
RQSPGKGLEWLGVIWSGGNTDYNTPFTSRLSINKDNSKSQVFFKMNSLQSNDTAIYYCA
RALTYYDYEFAYWGQGTLVTVSAGSEQKLISEEDL**

mCetux variable domains

Antibodies were expressed in gWiz either as human IgG1 or murine IgG2c. Highlighting denotes mutations from cetuximab.

Heavy Chain variable domain

CAGGTACAACTGAAGCAGTCAGGACCTGGCCTAGTGCAGCCCTCACAGAGCCTGTC
CATCACCTGCACAGTCTCTGGTTTCTCATTAACTAACTATGGTGTACACTGGGTTCGC
CAGTCTCCAGGAAAGGGTCTGGAGTGGCTGGGAGTGATATGGAGTGGTGGAAACAC
AGACTATAATACACCTTTCGCATCCAGACTGAGCATCAACAAGGACAATTCCAAGA
GCCAAGTTTTCTTTAAAATGAACAGTCTGCAATCTAATGACACAGCCATATATTACT
GTGCCAGAGCCCTCACCTACTATGATTACGAGTTTGCTTACTGGGGCCAAGGGACCC
TGGTCACCGTTTCCGCT

QVQLKQSGPGLVQPSQSLSITCTVSGFSLTNYGVHWVRQSPGKGLEWLGVIWSGGNTD
YNTPFASRLSINKDNSKSQVFFKMNSLQSNDTAIYYCARALTYYDYEFAYWGQGTLVT
VSA

Light chain variable domain

GACACCCTGCTGACCCAGTCTCCAGTCATCCTGTCTGTGAGTCCAGGAGAAAGAGTC
AGTTTCTCCTGCAGGGCCAGTCAGAGTATTGGCACAAACATACACTGGTATCAGCAA
AGAACAAATGGTTCTCCAAGGCTTCTCATAAAGTATGCTTCTGAGTCTATCTCTGGC
ATCCCTTCCAGGTTTAGTGGCAGTGGATCAGGGACAGATTTTACTCTTAGCATCAAC
AGTGTGGAGTCTGAAGATATTGCAGATTATTACTGTCAACAAAATCTTAATTGGCCA
ATCACGTTCGGTGCTGGGACCAAGCTGGAGCTCAAACGG

DTLLTQSPVILSVSPGERVSFSCRASQSIGTNIHWYQQRTNGSPRLLIKYASESISGIPSRFS
GSGSGTDFTLSINSVESEDIADYYCQQNLNWPITFGAGTKLELKR

qPCR primers for human and mouse EGFR and GAPDH

qPCR-mEGFR-F	AGGCACAAGTAACAGGCTCAC
qPCR-mEGFR-R	AAGGTCGTAATTCCTTTGCAC
qPCR-hEGFR-F	GCGTCTCTTGCCGGAATGT
qPR-hEGFR-R	GCAGGTGATGTTCATGGCCT
qPR-mGAPDH-F	GTTGTCTCCTGCGACTTCA
qPCR-mGAPDH-R	GGTGGTCCAGGGTTTCTTA
qPCR-hGAPDH-F	AGGGCTGCTTTTAACTCTGGT
qPR-hGAPDH-R	CCCCACTTGATTTTGGAGGGA

Appendix B – Parameters for cetuximab-Sso7d fusion PK modeling

For further description of the model, see Thurber, et al[162,163] and Rhoden, et al[86].

Parameter	Symbol	Units	Values
Volume of plasma	Vplasma	mL	2
Volume of tumor	Vtumor	uL	100
Tumor void fraction	ε	--	0.24
Number of cells per tumor	Ncells	#	10E6
Vascular permeability	P	cm/s	2.8E-7
Capillary raidus	Rcap	um	8
Krogh cylinder radius	Rk	um	75
Antibody affinity	Kd	nM	0.5
Antibody binding on-rate	kon	1/(M*s)	2E5
EGFR receptors/cell at steady state	EGFRss	1/cell	5E5
EGFR synthesis rate	ksyn	#/cell/min	150
Downregulation half-life (=ln(2)/ke)	t1/2,down	hrs	2.5
Antibody serum half-life (=ln(2)/kclear)	t1/2,clear	hrs	38

Species definitions:

```
Ab_p = antibody in plasma
Ab_tf = antibody free in tumor
Ab_tb = antibody bound in tumor
Ag = free EGFR antigen
Ab_int = antibody internalized (by normal receptor
synthesis/degradation)
Ab_e = antibody downregulated
Ab_clear = antibody cleared from plasma
```

ODE equations:

```
dAb_p = -k_clear*Ab_p + 2*P*Rcap/Rk^2 * (Ab_tf/eps - Ab_p) *
(V_tumor/V_plasma*eps)

dAb_tf = 2*P*Rcap/Rk^2*(Ab_p*eps - Ab_tf) - k_on*Ab_tf*Ag/eps +
k_off*Ab_tb

dAb_tb = k_on*Ab_tf*Ag/eps - k_off*Ab_tb - k_e*Ab_tb - k_int*Ab_tb

dAg = - k_on*Ab_tf*Ag/eps + k_off*Ab_tb + k_syn - k_int*Ag

dAb_int = k_int*Ab_tb

dAb_e = k_e*Ab_tb

dAb_clear = k_clear*Ab_p
```