ANTIBODIES DIRECTED AGAINST THE EXTRACELLULAR REGION OF THE EPIDERMAL GROWTH FACTOR RECEPTOR ADOPT DISTINCT MODES OF BINDING AND INHIBITION

Karl R. Schmitz

A DISSERTATION

in

Biochemistry and Molecular Biophysics

Presented to the Faculties of the University of Pennsylvania

in

Partial Fulfillment of the Requirements for the Degree of Doctor of Philosophy

2010

Supervisor of Dissertation

_____

Kathryn M. Ferguson, Associate Professor of Physiology

Graduate Group Chairperson

_____

Kathryn M. Ferguson, Associate Professor of Physiology

Dissertation Committee

Ronen Marmorstein, Professor, The Wistar Institute

Mark A. Lemmon, Professor of Biochemistry and Biophysics

Christopher G. Burd, Associate Professor of Cell and Developmental Biology

James Shorter, Assistant Professor of Biochemistry and Biophysics

Gregory Van Duyne, Professor of Biochemistry and Biophysics

Gregory P. Adams, Associate Professor, Fox Chase Cancer Center

Antibodies directed against the extracellular region of the epidermal growth factor receptor adopt distinct modes of binding and inhibition

Acknowledgement

I have had the benefit of much support, guidance, and advice over the years, and this seems like the right place to acknowledge some of the people who have helped me get this far. First, I have to thank Kate for being an excellent mentor and advisor. I have been lucky enough to work on a diverse set of projects with a great balance of independence and support. Kate's lab has been a fun place to do science.

I have to thank Shawn for encouraging me to join the lab and for fanning my interest in structures and crystallography; Yu-San for being a fun bay mate and a productive sounding board for data and ideas; Tom for his endless enthusiasm, even in the face of a daunting project; Thanuja for skilled help on many projects; Kat for making me laugh and bringing me chocolate; and Diego for an amalgam of intelligence, sarcasm, and Simpsons quotes. Thanks to everyone else from Kate and Mark's labs, former and present, who have made bench science, synchrotron trips, and lab meetings interesting and engaging. It has been a pleasure to work with such talented scientists.

I have to thank the many teachers who, over the years, helped lay the foundation for a career in science. I have always felt well prepared for the challenges ahead. I have to thank my family, and especially my parents for encouraging my fascination with science, nature, and everything that wriggles and crawls (perhaps against their better judgment – but I didn't study herpetology, so it all turned out ok). And lastly, but above all, I have to thank my wife, who has shared this scientific journey with me across many cities and many years. She has celebrated with me when the data were good and consoled me when they weren't. Thank you, Jess, for everything.

ABSTRACT

ANTIBODIES DIRECTED AGAINST THE EXTRACELLULAR REGION OF THE EPIDERMAL GROWTH FACTOR RECEPTOR ADOPT DISTINCT MODES OF BINDING AND INHIBITION

Karl R. Schmitz

Kathryn M. Ferguson

The work described in this dissertation comprises two distinct projects. In the first, we describe the structural and functional characterization of a family of Golgi associated cytosolic proteins, represented by Vps74 in fungi and GOLPH3 in animals, by X-ray crystallography, biophysical assays, and cellular techniques. We find that Vps74 is required for the proper steady state localization of a subset of Golgi enzymes in yeast, and that disruption of *vps74* results in incomplete protein glycosylation. We further describe the crystal structures of Vps74 and GOLPH3, identifying structural motifs required both for oligomer formation and protein function. Finally, we find that both Vps74 and GOLPH3 specifically bind the Golgi enriched phospholipid, PtdIns4*P*. These results suggest a role for Vps74 and GOLPH3 in retrograde trafficking of components to the Golgi apparatus.

In a separate and unrelated project, we characterize several inhibitory antibodies directed against the extracellular region of the epidermal growth factor receptor (EGFR). Aberrant activation of EGFR occurs in large proportion of epithelial cancers. Consequently, this receptor is a target for anti-cancer therapeutics that inhibit its activation, including antibodies and antibody-derived molecules. We have biochemically characterized a panel of conventional inhibitory antibodies with unique properties, and have identified approximate epitopes for these antibodies on domain 3 of EGFR. Additionally, we describe the crystal structures of three unconventional single chain antibody fragments in complex with the EGFR extracellular region. These single chain antibodies bind to novel epitopes on the receptor but share key characteristics with conventional inhibitory antibodies. Our findings highlight the diversity of binding modes among anti-EGFR antibodies, and suggest opportunities for novel therapeutics.

Table of Contents

List of Tables

List of Illustrations

1 Structure and function of Vps74 and GOLPH3

1.1 Introduction

1.1.1 Trafficking and the Golgi apparatus

One of the hallmarks of eukaryotic life is the presence of specialized intracellular membrane bound compartments. Distinct complements of protein, carbohydrate, and lipid enable these organelles to carry out specific cellular functions: DNA transcription occurs in the nucleus; the citric acid cycle and oxidative phosphorylation occur in the mitochondria; etc. The resulting compartmentalization of cellular processes is a key factor in the success and diversity of eukaryotic organisms. Consequently, eukaryotes possess well conserved and highly regulated mechanisms to enforce the proper spatial localization of cellular material.

The importance of this machinery is particularly apparent when considering the endomembrane system, and especially the Golgi apparatus. Much of the material that comprises the organelles, the plasma membrane, and matter secreted by the cell is processed, sorted, and packaged by the Golgi. Cargo molecules enter the Golgi from the endoplasmic reticulum (ER) and traverse the Golgi apparatus in the anterograde direction, crossing distinct c*is, medial,* and *trans* Golgi compartments in the process. (Multiple models of Golgi function have been proposed [32-36]; current data suggest that both cisternal maturation and vesicular transport occur to some degree [37-39].) Each sub-compartment is enriched with a subset of resident Golgi enzymes. Anterograde transport in conjunction with spatial organization of ER and Golgi enzymes gives rise to temporal organization in the processing of cargo, as in the stepwise modification of glycoproteins (see Section 1.1.2).

The steady state localization of resident ER and Golgi components must be constitutively enforced, as a portion of these molecules is continually mislocalized downstream along with anterograde cargo. Well conserved retrograde trafficking complexes identify mislocalized molecules and return them to their resident compartments. COPI (coat protein complex I) coated vesicles mediate intra-Golgi and Golgi-to-ER retrograde transport [40], while the retromer complex directs retrieval of proteins from endosomes to the *trans*-Golgi network

[41]. The molecular interactions that allow retrograde trafficking complexes to directly or indirectly recognize cargo and destination compartments are not thoroughly understood. Cytosolic KKXX and luminal KDEL motifs target cargo to the COPI complex, and thus to the ER [40]. Additionally, short retention motifs have been identified in the cytosolic tails of some retromer cargo proteins [41]. However, other classes of resident Golgi proteins exist for which no retention motif has been described.

1.1.2 Protein glycosylation is a primary function of the Golgi apparatus

Glycoproteins are a major constituent of Golgi anterograde cargo. A large fraction of proteins processed through the endomembrane system is modified by N- or O-linked glycosylation. The details of protein glycosylation vary somewhat among eukaryotes, but a general outline is conserved [42].

Core glycans are transferred to nascent polypeptides in the ER. In the Golgi apparatus, these glycans are modified and elaborated by the stepwise action of various glycosyltransferase enzymes. These enzymes adopt specific compartmental distributions within the Golgi (*cis, medial, or trans*) [42], and this spatial distribution roughly reflects the order in which these enzymes act on immature glycoproteins [43]. Disruption of the steady-state distribution of glycotransferases may alter the extent of protein glycosylation and manifest a variety of downstream effects in the cell; indeed, defective glycosylation is associated with several human diseases [44].

The majority of glycosyltransferases are type II transmembrane proteins, bearing a short N-terminal cytosolic tails [42, 43]. Studies have highlighted the importance these cytosolic regions in maintaining the proper localization of Golgi glycosyltransferases [43], but only recently have retention motifs been identified for a subset of these enzymes [19, 45, 46] (see Section 1.2).

1.1.3 Phospholipids function as Golgi markers and regulators

Phosphoinositide (PI) lipids comprise a small molar fraction of eukaryotic membranes overall, but play a large role in cellular signaling and membrane recognition [47, 48]. These phospholipids are enriched in specific membranes and organelles, and their levels are strictly

regulated. PIs serve as second messengers [49] (as in PI3K signaling [50]), as ligands for several classes of membrane binding domains (PH, PX, FYVE, etc.) [47], and as membrane markers (PtdIns(4,5)P_2 is enriched in the plasma membrane [48]; PtdIns4*P* is enriched in the Golgi [48]).

PtdIns4*P* has been found to be a major regulator of Golgi function. Steady state levels of PtdIns4*P* are governed by the action of phosphatidylinositol 4-kinases (PtdIns 4-kinases) and phosphatidylinositol phosphatases. Yeast express a relatively small number of PtdIns phosphatases and kinases, and serve as a convenient model for PI function. In quiescent yeast the Sac1 PtdIns phosphatase localizes to the Golgi and depletes Golgi PtdIns4*P*, which results in the downregulation of secretion [51]. Similarly, under starving conditions in yeast, PtdIns4*P* levels are diminished by the dissociation of Pik1/Frq1, the sole Golgi PtdIns 4-kinase, from the Golgi apparatus [52]; disruption of Pik1/Frq1 impairs secretion [53, 54], suggesting a reciprocal effect from that of Sac1. The molecular basis for the requirement of PtdIns4*P* is presumably multifaceted. It has been noted that vesicle trafficking from the *trans*-Golgi requires recognition of PtdIns4*P* by clathrin adapters [55]. Since PtdIns4*P* is the major PI marker in Golgi membranes [48, 56], additional links between PtdIns4*P* binding and trafficking are likely (see Section 1.3).

1.1.4 Golgi associated proteins Vps74 and GOLPH3 are implicated in trafficking

Vps74 (vacuolar protein sorting 74) and GOLPH3 (Golgi phosphoprotein 3; also called GMx33/GPP34) are fungal and animal orthologs, respectively, of a family of Golgi-associated proteins. (The sub-cellular localization of Vps74 in yeast has been reported as cytosolic [57]; however, this is likely an artifact induced by a C-terminal GFP tag. See section 1.2.6.) These proteins were the subject of few studies prior to 2008, yet several early reports link Vps74/GOLPH3 to retrograde trafficking. Rat GOLPH3 is a Golgi localized peripheral membrane protein [58]. Deletion of *VPS74* in yeast results in defective sorting of carboxypeptidase Y [59]. Deletion of *VPS74* along with either *YPT6* (a Rab-like GTPase required for fusion of endosome-derived vesicles to the *trans*-Golgi) or *RIC1* (a subunit of the GTP exchange factor that acts on Ypt6) results in synthetic lethality [60].

More recent studies, including the work discussed in the following sections (published in 2008 and 2009 [19, 28]), have identified several important functions for these proteins. Vps74 interacts with subunits of the COPI coat [45] and enforces the steady state localization of a subset of Golgi resident proteins in fungi [19, 45] (see Section 1.2); GOLPH3 links myosin to the Golgi membrane in animal cells [61]; and both Vps74 and GOLPH3 are PtdIns4*P* binding domains [28, 61] (see Section 1.3). Moreover, GOLPH3 is a potent and prevalent oncogene in human cancers [62].

1.2 Vps74 is required for the proper steady-state localization of Golgi mannosyltransferas enzymes in yeast

1.2.1 Deletion of *VPS74* in yeast results in a partial defect in glycosylation of secreted proteins

To begin to understand the cellular role of Vps74, we investigated the effects caused by deletion of the Vps74 genomic locus in *Saccharomyces cerevisiae*. We find that *vps74Δ* cells are viable, but exhibit deficient glycosylation of secreted proteins. In pulse chase experiments, this glycosylation defect is indicated by the higher electrophoretic mobility of ^{35}S-labeled secreted glycoproteins analyzed by SDS-PAGE (Fig 1A).

We next asked whether both N-linked and O-linked carbohydrate is altered in *vps74Δ* cells. The transfer of core N- and O-linked glycans is catalyzed by distinct enzymes in the ER, but the same Golgi glycosyltransferase may modify both classes polysaccharide [63]. We observe glycosylation defects in both an N-linked glycoprotein (carboxypeptidase Y, CPY; Fig. 1B) and an O-linked glycoprotein (Hsp150; Fig 1C).

Since deletion of *VPS74* does not entirely abolish protein glycosylation (an essential cellular function), we hypothesized that Vps74 is required for complete elaboration or modification of the immature oligosaccharide chain, subsequent to the addition of the core glycan in the ER. This is supported by the observation that the ER-modified precursor form of

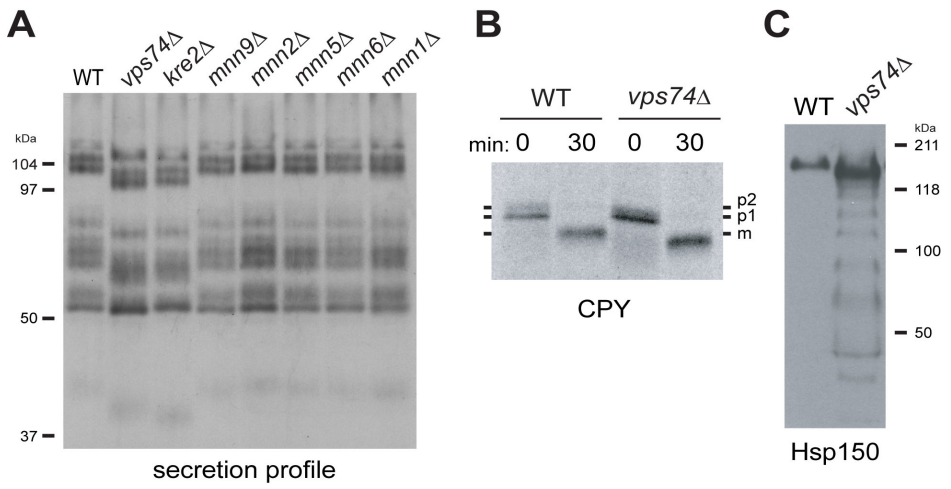

Figure 1. Vps74 is required for proper protein glycosylation in yeast.

(A) Pulse-chase secretion assay. The indicated strains were pulse labeled with ^{35}S-Met/Cys and chased with unlabeled Met/Cys. Secreted proteins were precipitated from growth medium, separated by SDS-PAGE, and visualized by fluorography. (B) The indicated strains were ^{35}S-Met/Cys pulse-labeled. The N-linked glycoprotein carboxypeptidase Y (CPY) was immunoprecipitated and visualized as in (A). The positions of the ER modified form (p1), Golgi modified form (p2), and mature CPY are indicated. (C) The glycosylation of HA-tagged Hsp150 was analyzed in wild type and *vps74Δ* strains. Secreted protein was precipitated from growth media, and exogenous Hsp150 was analyzed by western blot probed with anti-HA antibody.

(*From experiments done by J. Liu, C.S. Wood, C.G. Burd; adapted from [19].*)

CPY is unaffected in *vps74Δ* cells (Fig. 1B). This suggests that Vps74 is essential for the function of one or more of Golgi glycosyltransferase, which corresponds well with its established Golgi distribution.

A genome wide screen of *S. cerevisiae* identified yeast lacking *VPS74* as severely deficient in the binding of the cationic dye, Alcian Blue [64]. In wild type yeast, this dye binds to the negatively charged cell wall and indicates the presence of α-1,2-linked mannose phosphate. Since incorporation of mannose phosphate is accomplished by a family of Golgi-localized α-1,2-mannosyltransferases, we hypothesized that the observed glycosylation defect occurs through loss of function of these enzymes. To test which of these α-1,2-mannosyltransferases might be affected by loss of *VPS74*, we compared the glycosylation of ^{35}S-labeled secreted proteins in *vps74Δ* cells to that in strains carrying deletions of mannosyltransferase genes. Individual deletion of several mannosyltransferases partially phenocopies the *vps74Δ* phenotype, with *kre2Δ* resulting in the greatest decrease in glycosylation (Fig. 1A).

1.2.2 Vps74 is required for Golgi localization of yeast mannosyltransferases

Deletion of VPS74 has previously been noted to affect the sorting of carboxypeptidase Y [59], therefore we asked whether VPS74 deletion may affect mannosyltransferase function by altering the normal Golgi localization of these enzymes. To test this, the subcellular localization of C-terminally GFP tagged mannosyltransferases was determined in wild type and *vps74Δ* yeast. Mannosyltransferases adopt a punctuate distribution in wild type cells (Fig. 2), correlated with Golgi localization, and similar to the localization of N-terminally GFP tagged Vps4 (Fig. 2) [45]. However, in *vps74Δ* cells, mannosyltransferases localize to the lysozome-like yeast vacuole, downstream of the Golgi. This implicates Vps74 in retrograde trafficking of these enzymes to the Golgi apparatus.

1.2.3 Vps74 directly interacts with yeast Golgi mannosyltransferases

Given that Vps74 co-localizes with mannosyltransferase enzymes in wild type cells, we next asked whether Vps74 directly binds these enzymes. The α-1,2-mannosyltransferases are

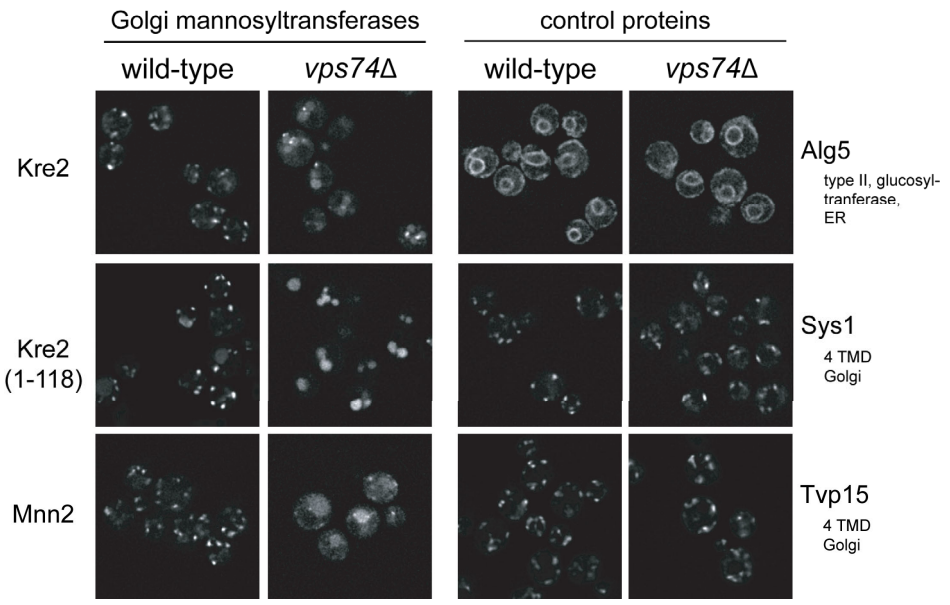

Figure 2. Golgi mannosyltransferases are mislocalized in *vps74Δ*.

The subcellular distribution is shown of C-terminally GFP-tagged versions of Golgi mannosyltransferases Kre2 and Mnn2, a truncated form of Kre2 lacking the luminal catalytic domain, and control proteins Alg5, Sys1, and Tvp15. Golgi distribution of Kre2, Mnn2, and truncated Kre2 is apparent by the punctuate distribution in wild type yeast. In the *vps74Δ* strain these proteins localize instead to diffuse bodies consistent with the lysozome-like yeast vacuole. The distribution of control proteins is unaffected in *vps74.*

(*From experiments done by J. Liu, C.S. Wood, C.G. Burd; adapted from [19]. Mannosyltransferase enzymes Mnn5 and Och1 are similarly mislocalized; see [19, 23]*.)

type II transmembrane proteins, composed of a short ≈ 10 amino acid cytosolic region, a transmembrane α-helix, a luminal juxtamembrane stalk region, and a luminal catalytic domain [65]. If cytosolic Vps74 directly interacts with mannosyltransferases, it must do so via their short cytosolic N-terminal regions. To test if this is the case, we generated glutathione S-transferase (GST) constructs with the N-terminal cytosolic peptides of mannosyltransferase enzymes fused to the N-terminus of GST, preserving the native topology of the peptide. (This is the reverse of the usual topology of GST fusion constructs.) We find that mannosyltransferase-GST fusions immobilized to glutathione (GSH) agarose beads pull down exogenously expressed Vps74 from *E. coli* lysates (Fig. 3B), as well as purified Vps74 in solution (Fig. 3A).

The N-terminal cytosolic peptides of yeast mannosyltransferases are typically composed of hydrophobic and basic amino acids. To address the concern that the presence of these sequences on the N-terminus of GST may have destabilized or partially unfolded the GST domain, giving rise to nonspecific hydrophobic interactions with Vps74, we also assayed the interaction between Vps74 and immobilized synthetic peptide. A peptide comprising the first 11 amino acids of the cytosolic region of mature Kre2, followed by cysteine, was synthesized and immobilized on agarose beads by thiol coupling. Free cysteine was immobilized on control beads. Immobilized Kre2 N-terminal peptide specifically pulls down purified Vps74 (Fig. 3C), confirming that Vps74 directly binds the mannosyltransferase cytosolic N-terminal region.

1.2.4 Crystal structure of Vps74 at 2.8 Å resolution

Cellular studies suggest a role for Vps74 in establishing interactions between retrograde transport components and specific Golgi enzymes. However, the lack of sequence homology between Vps74 and any structurally characterized motif led us to question the structural basis for these interactions. To address this and provide structural insight into the function of Vps74, we solved its crystal structure to a resolution of 2.8 Å. Full length *S. cerevisiae* Vps74 was expressed in *E. coli* and purified to homogeneity. Small needle shaped crystals with hexagonal cross section initially formed by hanging drop vapor diffusion in a number of crystallization conditions containing both polyethylene glycol (PEG), as the principle precipitant, and $CaCl_2$.

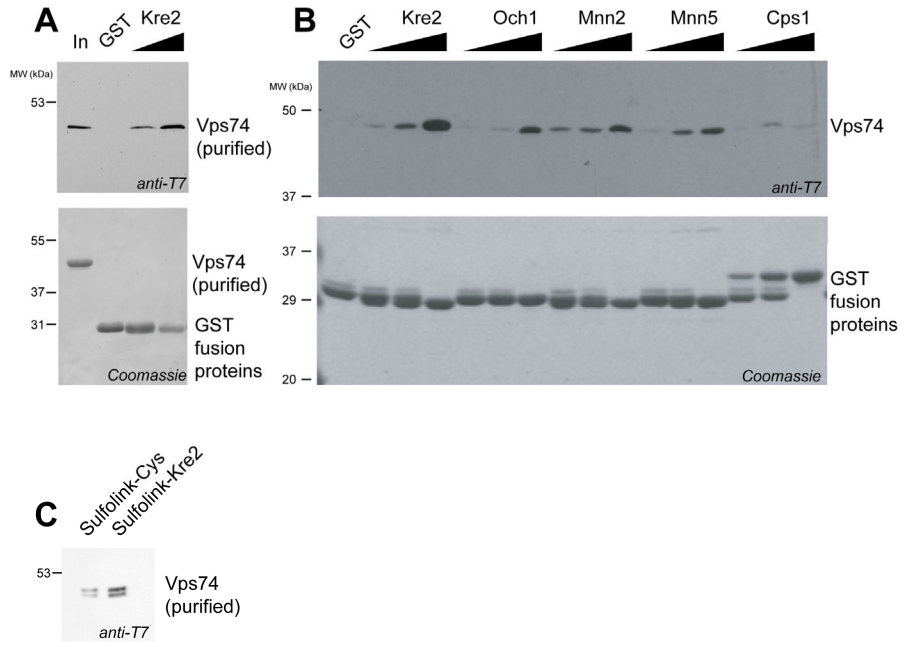

Figure 3. Vps74 interacts with the cytosolic N-terminal region of Golgi mannosyltransferases.

(A). A protein fusions incorporating the first 15 amino acids of Kre2 followed by GST was immobilized on GSH beads. Beads were incubated in binding buffer with purified T7-tagged Vps74. 0:1, 1:1, and 1:0 ratios of GST-fusion:GST were used. A constant total amount of immobilized protein was used. "In" represents 0.5% Vps74 input. (B) GST fusions incorporating the first 15 amino acids of the indicated proteins were immobilized on GSH beads and incubated with *E. coli* lysates containing T7 tagged Vps74. 1:5, 1:1, and 1:0 GST-fusion:GST ratios were used. (C) A peptide consisting of the first 11 amino acids of mature Kre2, followed by cysteine, was immobilized to Sulfolink beads by thiol coupling. Beads were incubated in binding buffer with purified T7-tagged Vps74. The lower band is C-terminal degradation product of Vps74.

(*Panel B from experiments done by J. Liu, C.S. Wood, C.G. Burd; adapted from [19].*)

Growth conditions were subsequently optimized, in part by exploring the effects of additive compounds. It was found that the chelating agents ethylenediaminetetraacetic acid (EDTA) and the chemically related molecule ethylene glycol tetraacetic acid (EGTA) substantially improve crystal growth and quality. Diffraction quality crystals of native (Fig. 4) and selenomethionine-containing Vps74 were ultimately obtained from reservoir conditions of 8% PEG 3350, 5% ethylene glycol, 50 mM MES, 50 mM NaCl, 50 mM $CaCl_2$, 10 mM EGTA, pH 6.0. Both EDTA and EGTA strongly chelate calcium anions with 1:1 stoichiometry. In the Vps74 crystallization condition, the concentration of free calcium is approximately 40 mM [66]. However, we note that crystal growth is markedly different in conditions of 40 mM $CaCl_2$, without EGTA, compared to the above condition. The effects of calcium and chelator on Vps74 are discussed below (Section 1.2.7).

The crystal structure of Vps74 was determined by selenium multi-wavelength anomalous dispersion (MAD) methods. Experimental phases were calculated from 3.5 Å resolution anomalous data, and were applied to a 2.8 Å native dataset (Table 1). Clear electron density for Vps74 was observed in experimentally phased maps, with the exception of the first 61 amino acids. MALDI (matrix-assisted laser desorption ionization) mass spectrometry of purified Vps74 gave no evidence of proteolysis of these amino acids, indicating that the N-terminal region is intact in the purified protein but disordered in the crystal. This region is poorly conserved throughout Vps74 and GOLPH3 proteins, and is predicted to be mostly disordered by secondary structure prediction algorithms Jpred [67] and Prof [68]. Additionally, it has been reported that deletion of the first 53 amino acids of GOLPH3 does not affect its targeting to the Golgi [69, 70], suggesting that this region of the protein may not be critical for at least this aspect of the protein's function. A variant lacking the first 59 amino acids of Vps74 (Vps74Δ59) crystallized under similar conditions, in the same crystal form, and with the same arrangement of protein in the asymmetric unit (Table 1). Crystallization of Vps74Δ59 exhibited less dependence on the concentration or presence of calcium and EGTA.

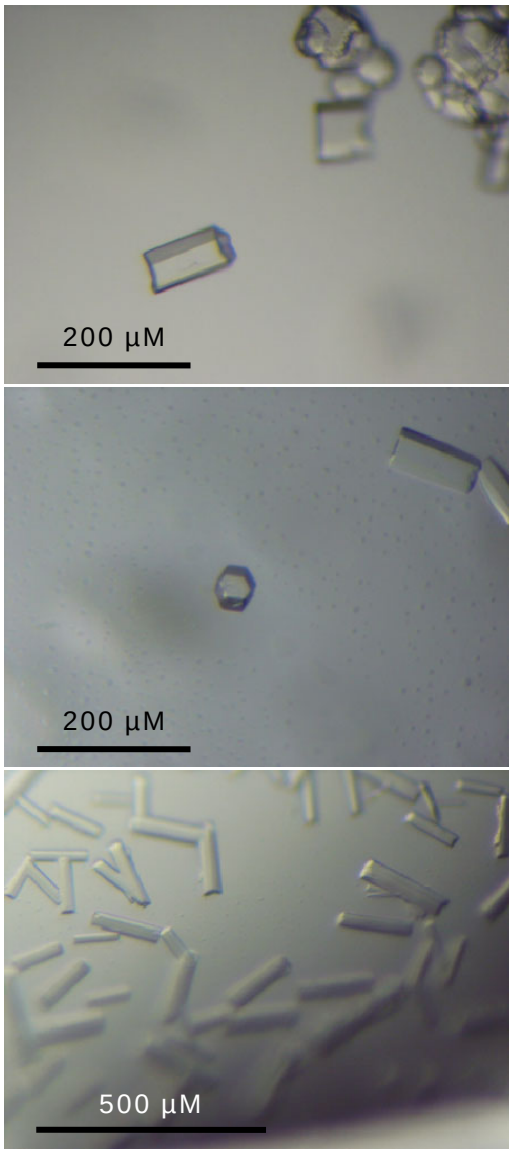

Figure 4. Crystals of full length Vps74.

Vps74 forms rod-shaped crystals with hexagonal cross section. Crystals with maximum dimensions of 75 × 75 × 300 µm formed in 8% PEG3350, 5% ethylene glycol, 50 mM MES, 50 mM NaCl, 50 mM $CaCl_2$, 10mM EGTA, pH 6.0.

Table 1. Vps74 Crystallographic Statistics

	Vps74 (native)	Se-Vps74 (peak)[a]	Se-Vps74 (inflection)[a]	Se-Vps74 (remote)[a]	Vps74Δ59
Data Collection Statistics[b]					
Space group	$P3_121$	$P3_121$	$P3_121$	$P3_121$	$P3_121$
Unique cell dimensions	a, b = 102.7 Å, c = 292.3 Å	a, b = 101.5 Å, c = 293.1 Å	a, b = 101.9 Å, c = 293.7 Å	a, b = 101.4 Å, c = 290.8 Å	a, b = 104.1 Å, c = 292.8 Å
X-ray source	APS 23-ID-D	APS 23-ID-D	APS 23-ID-D	APS 23-ID-D	CHESS F1
Wavelength	1.03602 Å	0.97926 Å	0.97942 A	0.94927 Å	0.918 A
Resolution limit	2.8 Å	3.4 Å	3.55 Å	2.9 Å	3.05 Å
Observed/Unique	228,024/43,558	254,363/24,850	214,868/21,767	388,891/38,973	218,302/35,872
Redundancy	5.2 fold	10.2 fold	9.9 fold	10.0 fold	6.1 fold
Completeness (%)	97.4 (83.8)	99.0 (95.4)	98.6 (91.8)	98.9 (89.6)	99.8 (100)
R_{sym}[c]	0.093 (0.342)	0.140 (0.456)	0.141 (0.584)	0.116 (0.459)	0.142 (0.529)
$\langle I/\sigma \rangle$	23.7 (2.9)	21.3 (3.3)	18.1 (2.3)	26.2 (2.6)	14.9 (3.25)
Refinement Statistics					
Resolution limits	38-2.8 Å				49-3.05 Å
No. reflections/No. test set	41,312/2,087				34,003/1,700
R factor (R_{free})[d]	0.24 (0.30)				0.22 (0.28)
Model					
Protein	4x Vps74				4x Vps74Δ59
	aa 61-343[e]				aa 61-343[e]
	5 water molecules				2 water molecules
Total number of atoms	9,079				9,068
RMSD bond lengths (Å)	0.013				0.013
RMSD bond angles (°)	1.499				1.651

a Se-Vps74 peak and remote datasets were collected on the same crystal; inflection was collected on a second crystal.

b Numbers in parenthesis refer to last resolution shell.

c $R_{sym} = |I_h - \langle I_h \rangle| / I_h$, where $\langle I_h \rangle$ = average intensity over symmetry equivalent measurements.

d R factor = $|F_o - F_c| / F_o$, where summation is over data used in the refinement; R_{free} includes only 5% of the data excluded from the refinement.

e Number of missing amino acids varies by chain.

1.2.5 Vps74 adopts a novel α-helical fold, which includes a prominent β-hairpin motif

Vps74 forms a single globular domain that is predominantly α-helical (Fig. 5). A central hydrophobic four-helix bundle (α1, α2, α8, and α12) lies at the core of the protein. A series of solvent-exposed loops and eight amphipathic helices lie roughly orthogonal to the core. Strands β3 and β4 form a β-hairpin (comprising amino acids 197-208), which projects away from the bulk of the helical domain and exposes hydrophobic residues at its side (M198, F201, F204, M206) and tip (F202, L203). It seems unlikely that this conformation of the β-hairpin is stable independently in solution, given the energetically unfavorable exposure of hydrophobic surface area that would result. Indeed, as discussed in detail below (Section 1.2.6), this region is involved in an intermolecular interaction with a non-crystallographically related Vps74 molecule in the asymmetric unit.

The Vps74 structure appears to be novel, and bears no detectible homology to known protein folds as assessed by the DALI server [71], automated comparison against the CATH structural database [72], and SSM comparison against the SCOP database [73]. DALI searches and sequence-based threading algorithms [74] do highlight similarity between the β-hairpin region of Vps74 (α1, β3, β4, and α6) and the β-hairpin "wing" of winged helix proteins [75, 76]. As highlighted in green in Fig. 6A & B, both structures contain a β-hairpin projection on a helical base. Though most winged helix (WH) domains are found in DNA-binding proteins, the fold also occurs in protein-protein interfaces [75]. Most interestingly, Vps22, Vps25, and Vps36, subunits in the ESCRT-II endosomal trafficking complex, all contain two repeats of a WH motif [77, 78]. Vps22 and Vps36 both use their C-terminal WH motifs to interact with Vps25. We suggest that the Vps74 WH-like region might also play an important role in protein-protein interactions (see Section 1.2.6). The Vps74 β-hairpin region is not directly homologous to the winged helix proteins, however, as it exhibits no significant sequence similarity and it differs from WH motifs in its topology (Fig. 6). Helices α1 and α6 replace H1 and H3 in the WH motif (H2 is missing). In

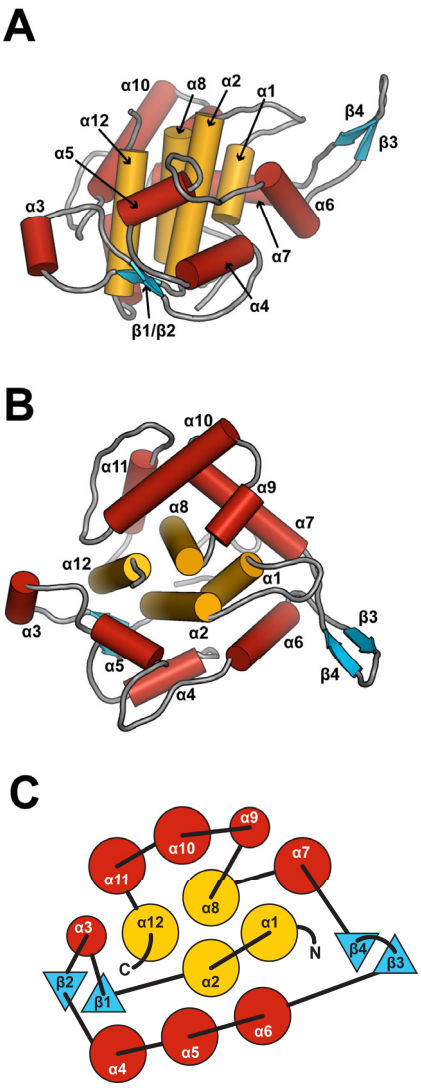

Figure 5. The Vps74 monomer adopts a novel α-helical fold.

(A) The X-ray crystal structure of a Vps74 monomer is shown as a cartoon. A network of amphipathic helices (red) surrounds a core of four hydrophobic helices (gold). A hydrophobic β-hairpin (cyan, β3/β4) projects outward from this helical domain. (B) An orthogonal view of Vps74, generated by an approximate 90° rotation of (A) about a horizontal axis. (C) A TOPS style [16] diagram shows the Vps74 fold topology. Circles represent helices, and triangles denote strands, colored as in (A) and (B).

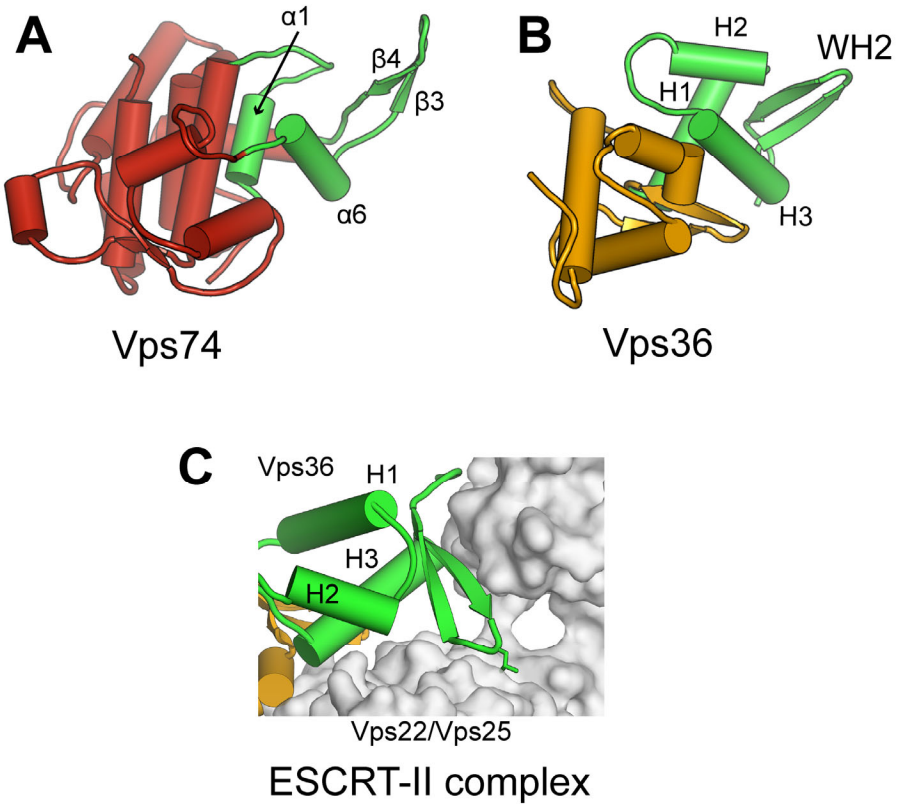

Figure 6. The Vps74 β-hairpin exhibits similarities to WH motifs.

The Vps74 β-hairpin region (A) exhibits structural similarities to the "wing" of winged-helix domains, such as Vps36 (B). β-hairpins and neighboring α-helices are indicated in green. Vps74 helices α1 and α6 adopt similar positions with respect to the β-hairpin as helices H1 and H3 in Vps36. (There is no analog to H2 in Vps74.) In both cases, the hairpin extends from the C-terminus of the last helix (α6 and H3). (C) The WH-domain of Vps36 mediates an interaction with protein partners in the ESCRT-II complex. The Vps74 β-hairpin is also involved in protein/protein interactions.

both cases, the hairpin extends from the C-terminus of the last helix (α6 and H3). Intriguingly, the WH-like motif of Vps74 mediates two intermolecular interactions in the crystal structure.

1.2.6 Vps74 forms a tetramer in the crystal structure

Vps74 crystallized with four molecules in the asymmetric unit. These subunits associate to form a tetramer with two-fold, non-crystallographic symmetry. As shown in Fig. 7A & B, the Vps74 tetramer comprises two central (yellow and cyan) and two peripheral (red and blue) subunits, stabilized by three distinct types of interaction that we refer to as the "inner" (Fig. 7A), "β-hairpin" (Fig. 7B), and "outer" (Fig. 7B) interactions.

The inner interaction involves a symmetric association of the two central Vps74 molecules through an interface consisting of their α1-α2 and α9-α10 loops (Fig. 8A & B). This interface is stabilized predominantly through electrostatic and hydrogen-bonds (Table 2) mediated by the α1- α2 loop (particularly R79 and symmetric van der Waals ring packing of F87), α7, and the α9- α10 region (notably T271 and R277). The inner interaction has the largest associated buried surface area of the three interaction types (burying 1255 $Å^2$ on each molecule). The β-hairpin interaction links a peripheral molecule (e.g. the red molecule in Fig. 7A & B) to one of its two central interaction partners (yellow in Fig. 7A & B). This interaction involves the mutual burial of a hydrophobic side chain (M206) between the β3/β4 hairpins of two Vps74 molecules (Fig. 8C), and the projection of an FFLF motif at the tip of each hairpin (residues 201-204) into a pocket formed by α6 and the α1-α2 loop on the partner (Fig. 8C). A salt bridge between the hairpin (D205) and a side chain on loop α6 (R181) further stabilizes the interaction. The projection of hydrophobic side chains from the tip of a β-hairpin into the pocket of a binding partner is a feature observed in WH motif-driven protein-protein interactions [77-79], including those that stabilize the ESCRT-II complex (Fig. 6). The β-hairpin interaction buries ≈ 1000 $Å^2$ on each subunit, of which a high proportion (70%) is hydrophobic (Table 2) [80].

Finally, the outer interaction connects each peripheral molecule to its other central interaction partner (yellow and blue subunits, Fig. 7B & 8A). This is an asymmetric interaction in

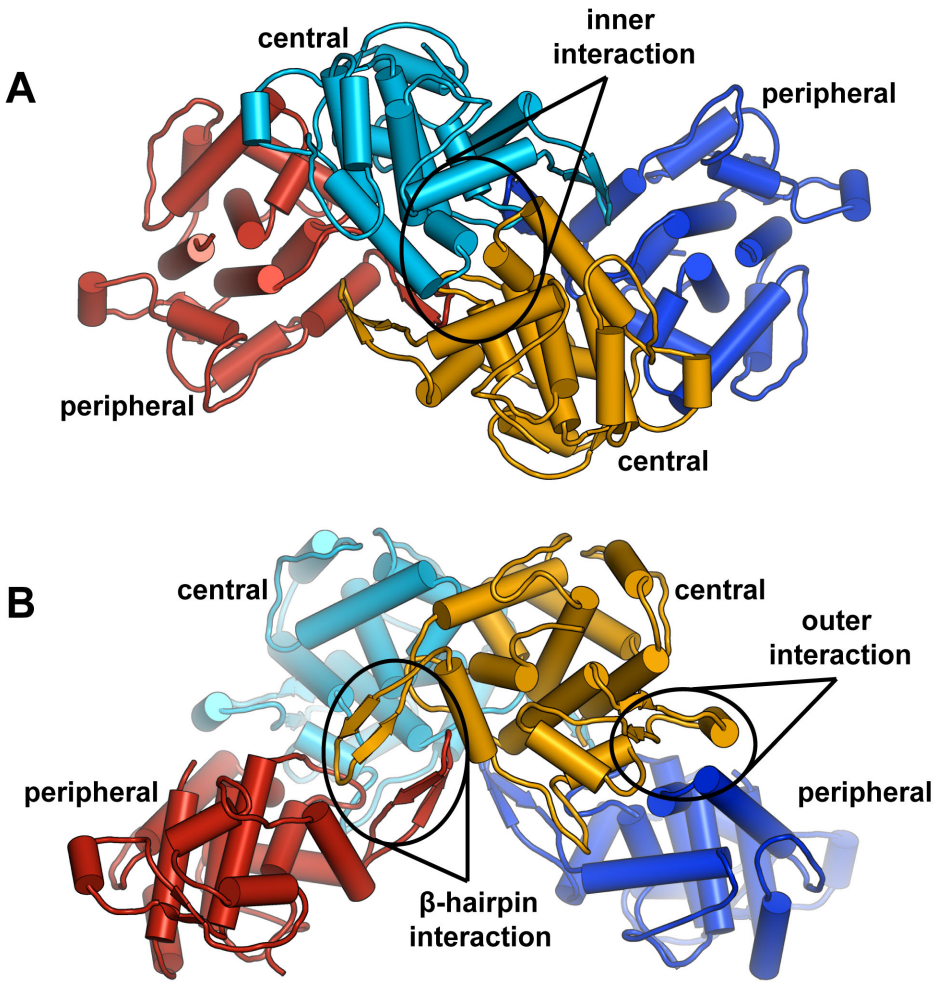

Figure 7. Vps74 forms a tetramer in the crystal structure.

A cartoon of the Vps74 tetramer is shown in "top" orientation (A) showing the overall two-fold symmetry, and in "side" orientation (B). Monomers occupy two distinct positions: central (gold and cyan) and peripheral (blue and red). The central subunits associate via a two-fold symmetric "inner interaction," highlighted in (A). A central and peripheral subunit associate via their hydrophobic β-hairpins in a "β-hairpin interaction," indicated in (B). Each central subunit also associates with a peripheral subunit in an "outer interaction." Loops have been smoothed for clarity. Cartoons rendered in Pymol [12].

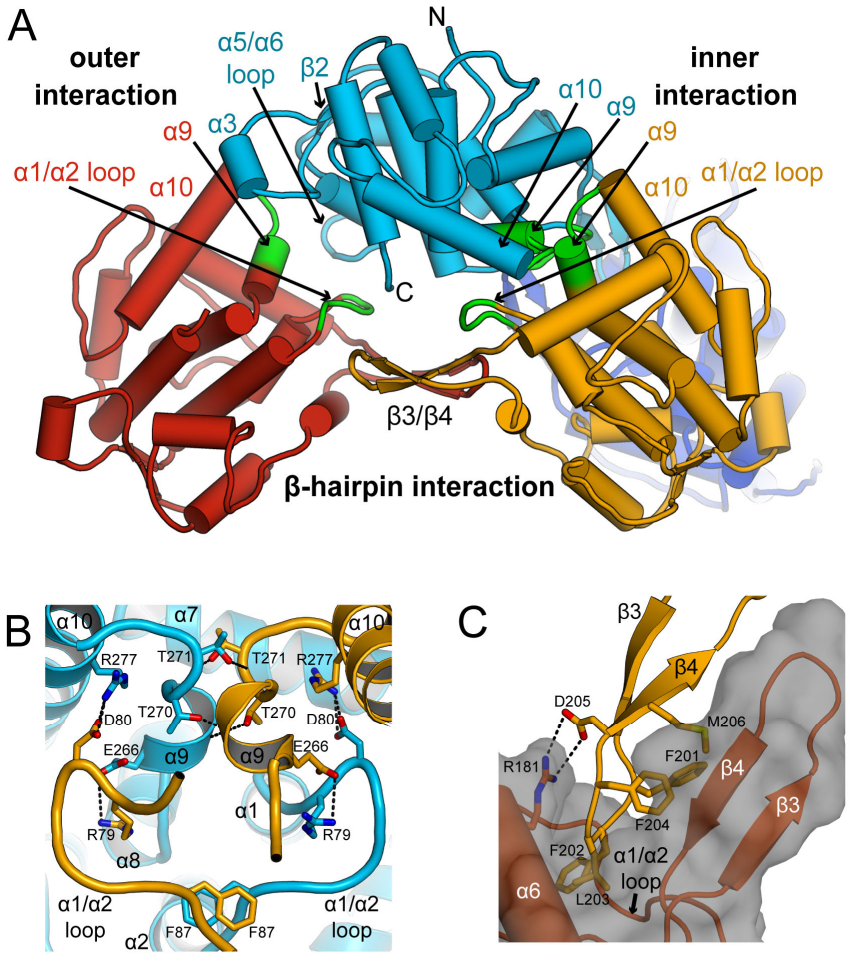

Figure 8. The Vps74 tetramer is stabilized by three distinct interactions.

(A) Cartoon of Vps74 tetramer, showing the approximate two-fold symmetry of the β-hairpin interaction. In the three foreground subunits, the α1/α2 loop, α9, and the α9/α10 loop are highlighted in green. Loops have been smoothed for clarity. (B) Detailed view of the inner interaction. Side chains that make contacts across the interface are shown in stick representation. Hydrogen bond/salt bridge interactions are shown. (C) Close up view of the interaction of the tip of the b-hairpin with the pocket formed by α6 and the α1/α2 loop. The side chains at the tip of the β-hairpin from the gold subunit (amino acids 202–206) are shown in stick representation. A transparent surface is shown around the red subunit.

which amino acids from the α1-α2 and α9-α10 loops (notably D80, R79, Y274, R277, D278) of the peripheral subunit contact the α3-β2, α5-α6, and C-terminal regions (notably R120, F121, D161, E166) of the central subunit. This interaction comprises mostly polar and electrostatic contacts, and the regions involved are quite open to solvent, rather than forming a unified solvent excluded contact area (as in the inner and β-hairpin interactions). Although the outer interaction buries the smallest area (≈ 750 $Å^2$ per molecule), it contributes to the simultaneous binding of each peripheral subunit to two central subunits, and is likely to be critical for stabilizing tetrameric assembly.

As described, subunits in the Vps74 tetramer occupy either central or peripheral positions. Subunits in each position stabilize tetramer formation through a different complement of interactions. Amino acids of the α1/α2 loop participate in all three interactions (indicated in Fig. 8A), and this loop occupies distinct conformations in the central and peripheral subunits. Consequently, there are no "unused" interaction surfaces on the crystallographic Vps74 tetramer, and a higher order oligomer cannot form solely utilizing the observed inner, outer, and β-hairpin interactions.

It is notable that the C-termini of the central subunits are relatively occluded in the tetramer, by both an outer and a β-hairpin interaction (Fig. 8A). A bulky group (such as GFP) fused to the Vps74 C-terminus (Vps74-GFP) would be expected to hinder tetramer formation. This may account for the discrepancy in reported localization [57] and loss of *in vivo* function (Fig. 17A) of Vps74-GFP.

In addition to the extent of the protein-protein interfaces in the Vps74 tetramer, evidenced by their substantial buried surface areas and shape complementarities (Table 2), sequence conservation data suggest that the tetramer is physiologically significant. The regions of Vps74 that form interaction surfaces in the tetramer, particularly the β3/β4 hairpin, are strongly conserved across fungal and animal homologs (Figs. 9 & 10). Further, by examining the degree of conservation of solvent-exposed residues on the surface of the tetramer, distinct regions are apparent (Fig. 11A). The lower concave surface (Fig. 11A, bottom view) possesses a

Table 2. Details of the three protein-protein interfaces that stabilize the Vps74 tetramer

	β-hairpin[1]	Inner	Outer[1]
Number of Hydrogen bonds (≤ 3.5 Å)	6	20	6
Salt bridge interactions	2	4	3
Area occluded from solvent[2] (Å^2)	1004 Å^2	1256 Å^2	757 Å^2
Polar area (% of total)	295 Å^2 (29 %)	698 Å^2 (56 %)	362 Å^2 (48 %)
Apolar area (% of total)	709 Å^2 (71 %)	558 Å^2 (44 %)	395 Å^2 (52%)
Shape complementarity[3]	0.66	0.60	0.62

1. Average for the two representations of this interface in the asymmetric unit.
2. Average area occluded on each subunit.
3. Calculated with the SC module in CCP4 [8].

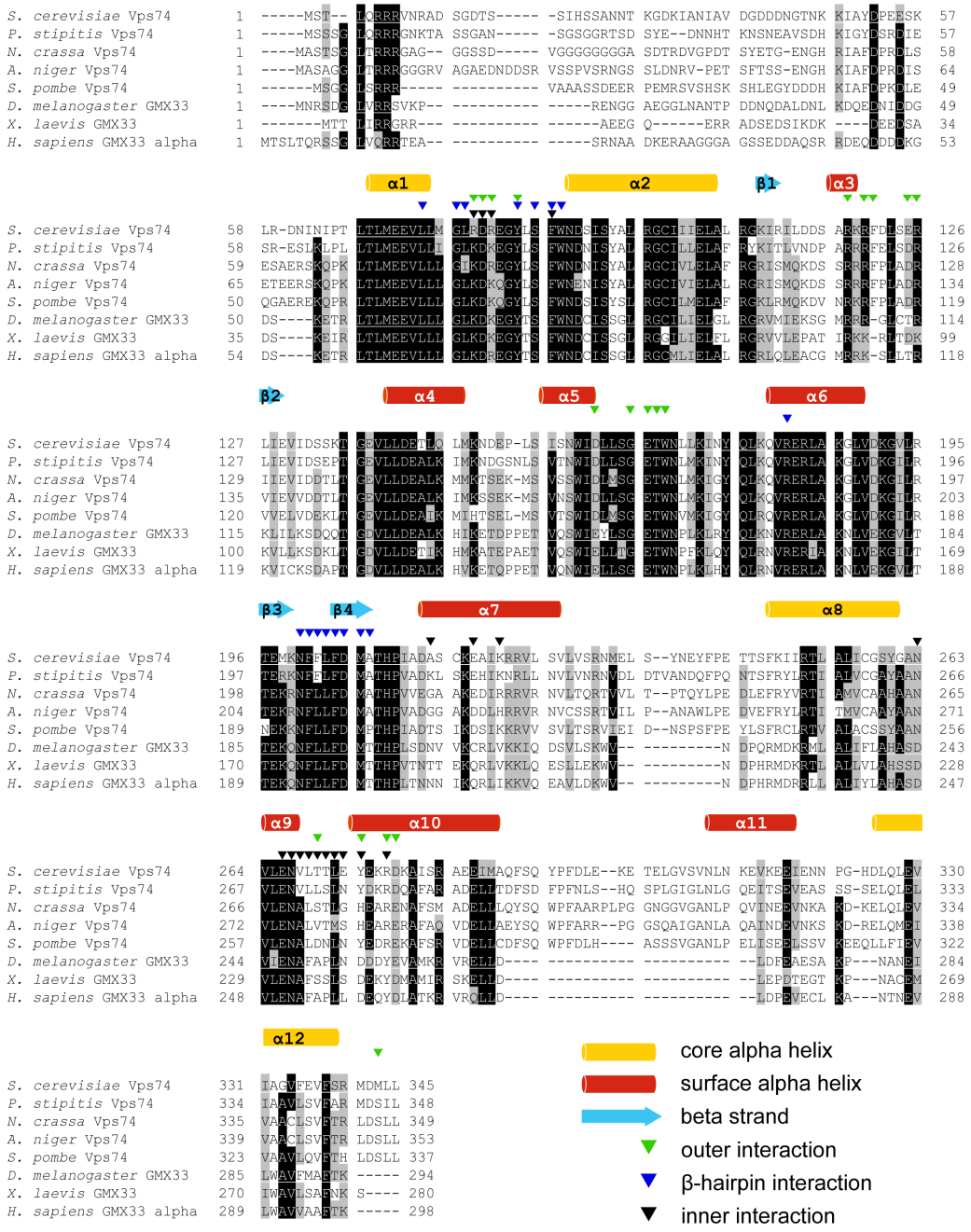

Figure 9. Multiple sequence alignment of Vps74 fungal and animal homologs.

A multiple sequence alignment shows sequence homology between fungal Vps74 and animal GOLPH3 homologs. Vps74 secondary structure and tetramer interface amino acids are indicated. The alignment was generated by ClustalW [4] and manually optimized.

high degree of sequence conservation, while the upper convex surface (Fig. 11A, top view) does not. This segregation of conserved surface-exposed amino acids is mirrored in the electrostatic potential of the tetramer (Fig. 11B), which may be relevant since the N-terminal tails of mannosyltransferases have substantial basic character. The conserved concave face of Vps74 is negatively-charged whereas the less conserved convex surface is more basic. Additional regions of high amino acid conservation, appreciated in the top view in Fig. 11A, also correlate with regions of strong negative electrostatic potential (Fig. 11B). In addition to supporting the relevance of the oligomer that we observe in the crystal, these observations may have implications for Vps74 function.

1.2.7 Vps74 forms an oligomer in certain solution conditions

To assess whether this tetramer may form outside of the crystallization conditions, we analyzed the oligomeric state of Vps74 in solution. In simple buffer conditions (25 mM HEPES, 150 mM NaCl, pH 7.5) both Vps74 and Vps74Δ59 appear monomeric by sedimentation velocity analytical ultracentrifugation (SV AUC; Fig. 12 & solid blue and black lines Fig. 13), size exclusion chromatography, and dynamic light scattering experiments (data not shown). Only at high protein concentrations (> 25 μM) are higher-order species detected (apparent as a faster sedimenting shoulder by SV AUC, red curve in Fig. 12), and data were consistent only with weak dimerization (K_D > 100 μM).

By contrast, in buffer containing 10 mM $CaCl_2$ plus 10 mM EGTA we observe clear evidence of tetramer formation by SV AUC by both Vps74 and Vps74Δ59 (dashed black and blue curves, Fig. 13). Based on the distribution of Vps74 into monomer and tetramer populations, we estimate a dissociation constant (K_D) for the monomer-tetramer equilibrium in the presence of Ca/EGTA in the range of 5-10 μM. The formation of some higher order species (Fig. 13A) prevented a more quantitative analysis of tetramer stability. Nevertheless, the appearance of tetramer both in the crystallographic asymmetric unit and under certain solution conditions suggests that the same interactions stabilize both oligomers.

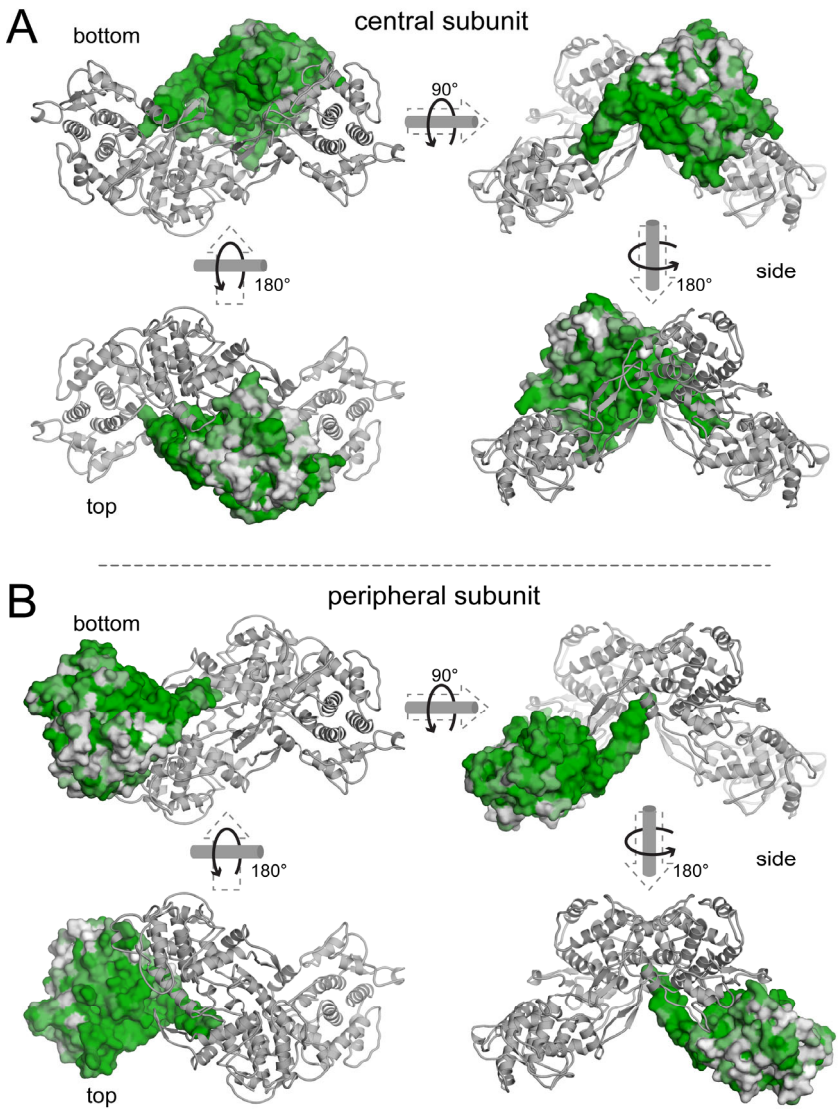

Figure 10. Vps74 surfaces involved in the tetramer interactions are well conserved. Amino acid conservation based on a multiple sequence alignment of fungal Vps74 proteins is projected on a surface representation of a central (A) or a peripheral (B) subunit of the Vps74 tetramer. Green indicates greatest conservation; white indicates least. Remaining subunits are shown as gray cartoons. Loops have been smoothed for clarity. Amino acids that form the interface between tetramer subunits are relatively well conserved. This pattern of conservation extends to animal GOLPH3 (not shown).

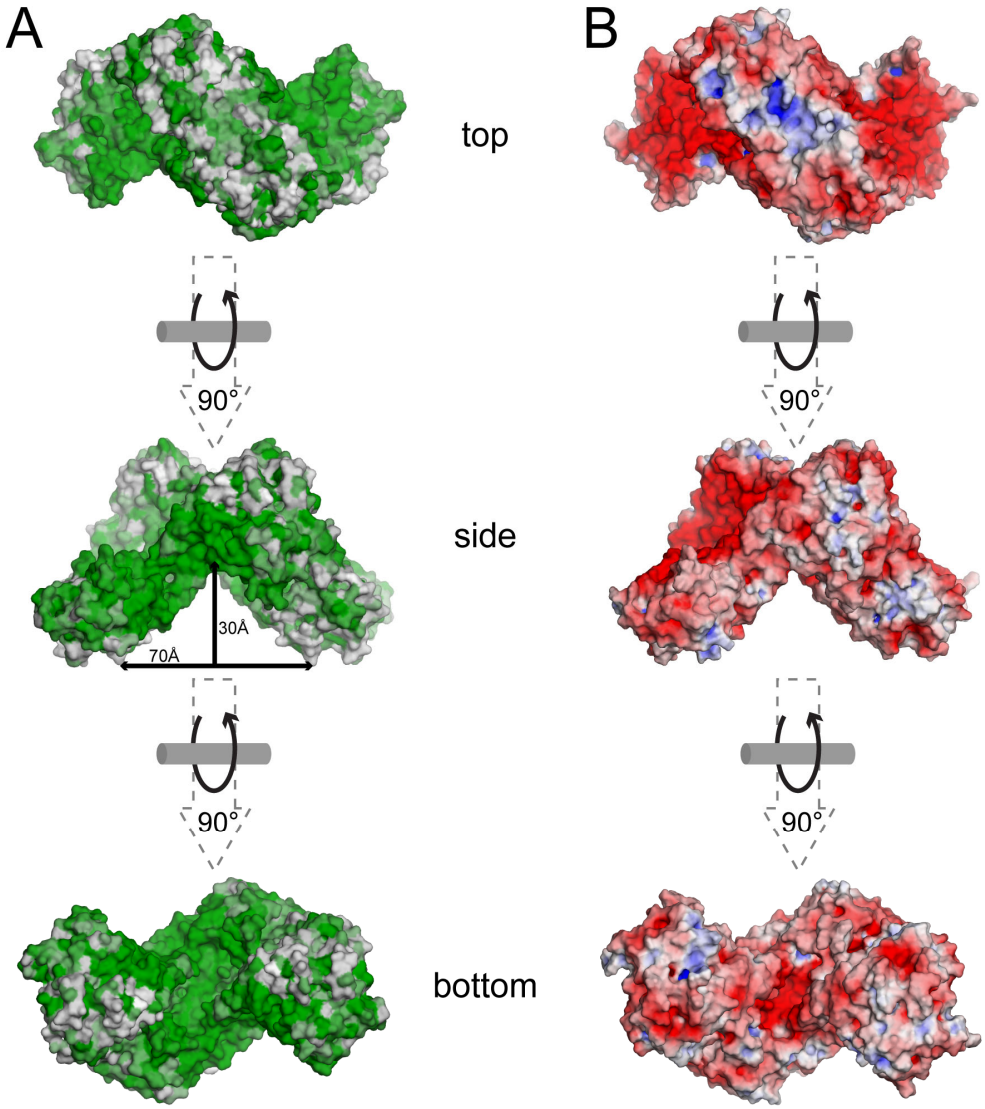

Figure 11. The distribution of conserved amino acids and surface charges reveal features on the surface of Vps74.

(A) Sequence conservation among fungal homologs of Vps74 is mapped onto the surface of the tetramer, as in Fig. 10. The top, convex surface exhibits low sequence conservation. Amino acids lining a 70 Å × 30 Å cavity in the bottom, concave surface are well conserved, and may indicate a protein interaction site. (B) Electrostatic potential (-8 kT, red, to +8 kT, blue) is shown on the tetramer surface. Patches of uniform charge occur on the top, sides, and bottom. Generated by APBS [13] in Pymol [12].

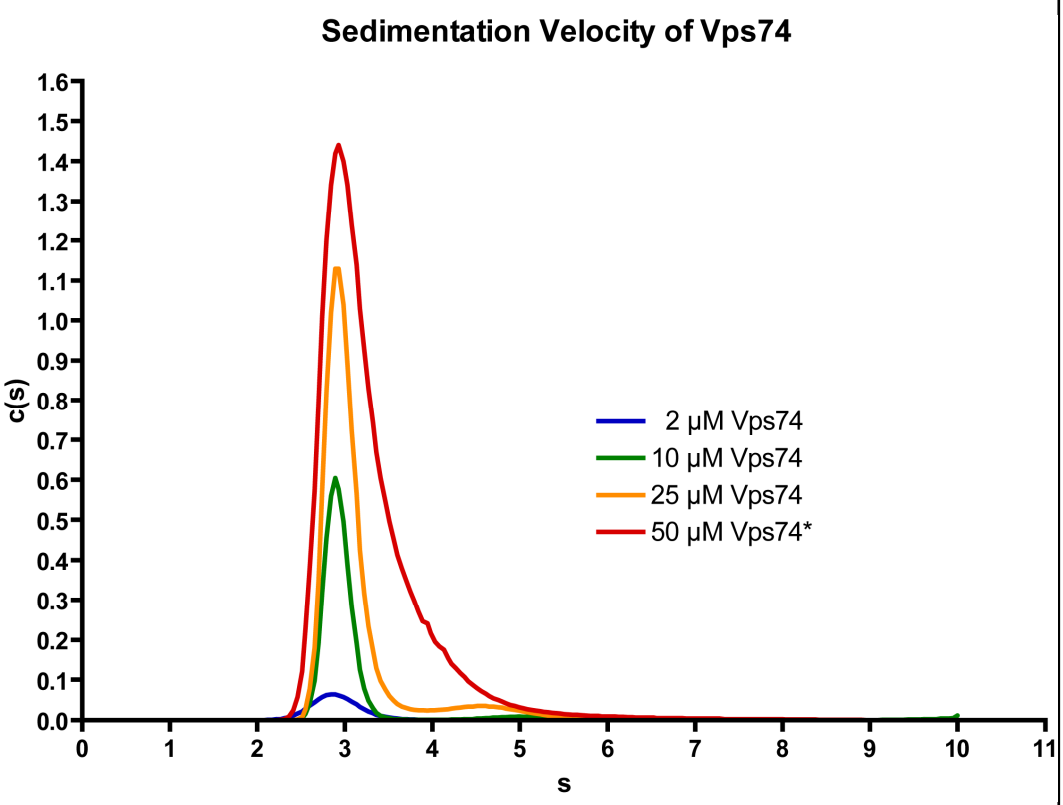

Figure 12. Vps74 is monomeric under simple buffer conditions.

Samples of Vps74 at the indicated concentrations were analyzed by sedimentation velocity analytical ultracentrifugation, and are presented as c(s) distributions. The peak at approximately 2.8 S corresponds well with monomeric Vps74. Weak monomer-dimer equilibrium is evidenced by a small 4.6 S peak in 25 µM sample, and a broad shoulder to the monomer peak at 50 µM. The sedimentation of 50 µM Vps74 was monitored at 250 nm, and the resulting c(S) data has been adjusted by a factor of 2.1 for comparison to data collected at 280 nm. Low S value artifacts have been excluded from the plot (see Methods 5.4.1).

The β-hairpin plays a key role in formation of the crystallographically-observed Vps74 tetramer. Its disruption would be predicted to prevent tetramerization, while leaving the central dimer intact. To test this, we generated a variant of Vps74 lacking the β3/β4 hairpin, Vps74Δ(197-208). In the absence of Ca/EGTA this variant sediments as a monomer (solid red curve in Fig. 13). In the Ca/EGTA buffer that promotes tetramer assembly in full-length Vps74, Vps74Δ(197-208) was found to sediment only as a distribution of monomer and dimer, with no evidence of tetramer (dashed red curve in Fig. 13). We also constructed variants, in the context of Vps74Δ59, that target amino acids in the peripheral interface. Vps74Δ59(R120A/F121A) alters side chains that form a salt bridge and hydrophobic contacts, while Vps74Δ59(Y274A) alters a tyrosine that makes a hydrogen bond as well as van der Waals contacts. Both peripheral interface variants sediment as monomer in the absence of Ca/EGTA, but as monomer and dimer in the presence of Ca/EGTA (Fig. 14). These data confirm that the same interactions that stabilize tetramer in the crystal structure are required for tetramer formation in solution.

Oligomerization in solution is not exclusively dependent Ca/EGTA. Several divalent cations (Zn^{2+}, Mn^{2+}, Co^{2+}, but not Mg^{2+}) drive tetramer or higher order oligomer formation in combination with EGTA or EDTA (D. Dersh; not shown). This argues against a specific interaction between Vps74 and metal or chelator, and further suggests that tetramer is stabilized by an interaction with the metal ion/chelator complex.

We asked whether pH influences Vps74 oligomerization. Vps74 crystals formed at pH 6.0, whereas the above SV AUC studies were conducted a pH 7.5. SV AUC experiments in the absence of Ca/EGTA yield monomeric distribution of Vps74 from pH 6.0 to 7.5 (Fig. 15). By contrast, in the presence of Ca/EGTA we observe decreased oligomer formation at pH 6.0 compared to pH 7.5. Histidine side chains do not make stabilizing contacts in any of the three tetramer interfaces, and crystals of tetramer formed at pH 6.0, so it is unlikely that this pH difference directly affects the stability of the tetramer interactions. We have already noted that free Ca^{2+} and EGTA alone do not stimulate tetramer formation. Thus, the observed pH sensitivity further suggests that tetramer stabilization occurs via the Ca/EGTA complex. The

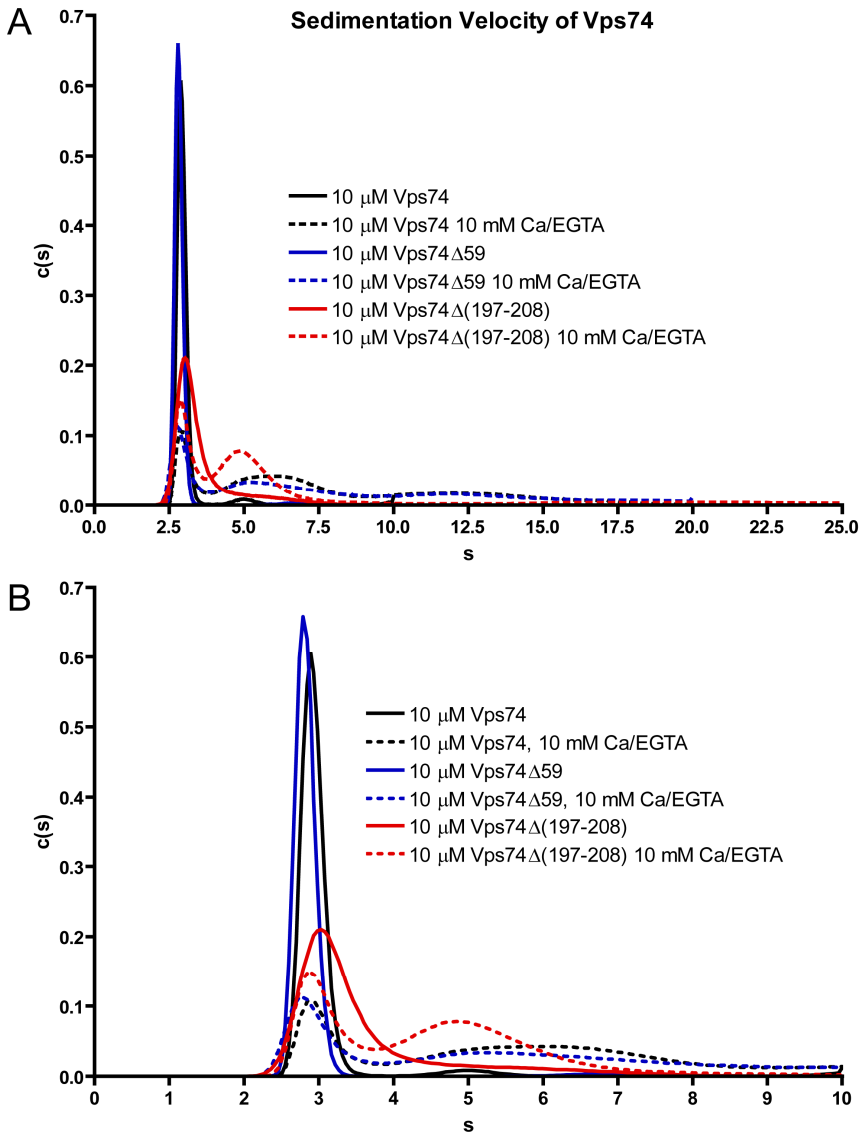

Figure 13. Vps74 forms oligomers in solution in the presence of Ca^{2+} and EGTA.

Vps74 and the indicated variants were analyzed by sedimentation velocity analytical ultracentrifugation, and are plotted as c(S) distributions to 25 S (A) or truncated to 10 S, to magnify the monomer-tetramer scale. Vps74 and Vps74Δ59 are monomeric in simple buffer conditions, but redistribute into a monomer/dimer equilibrium with some higher species in the presence of 10 mM $CaCl_2$, 10 mM EGTA. By contrast, Vps74Δ(197-208) populates a monomer/dimer equilibrium in the presence of Ca/EGTA.

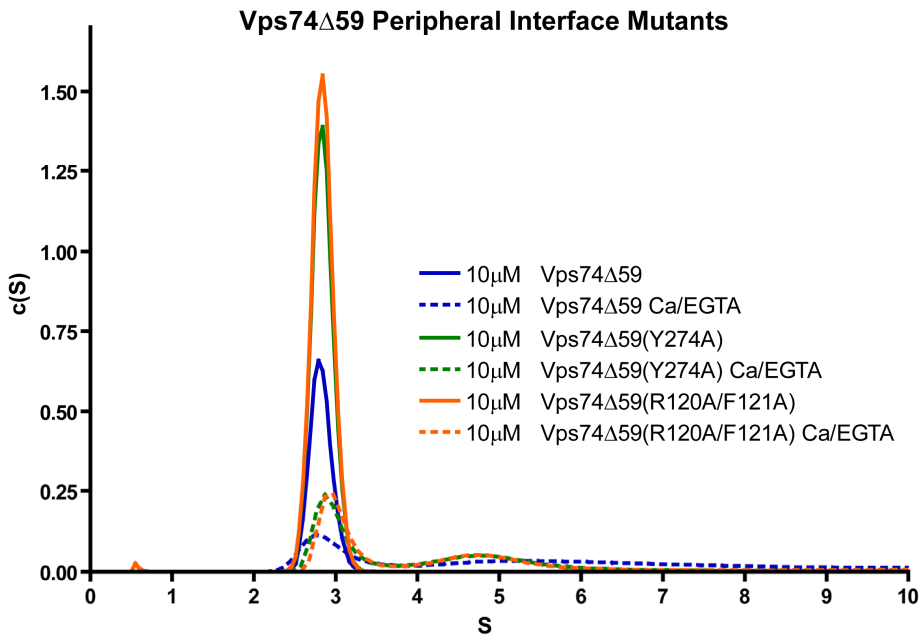

Figure 14. Alterations to amino acids that stabilize the peripheral interface affect tetramer formation.

Vps74Δ59, or variants containing mutations to residues in the peripheral interface, were analyzed by sedimentation velocity analytical ultracentrifugation, plotted as c(S) distributions. The sedimentation distribution of Vps74Δ59 shifts from monomeric to monomer/tetramer equilibrium in the presence of Ca/EGTA. Variants Vps74Δ59(Y274A) and Vps74Δ59(R120A/F121A) are monomeric in the absence of Ca/EGTA, but exist in monomer/dimer equilibrium in its presence.

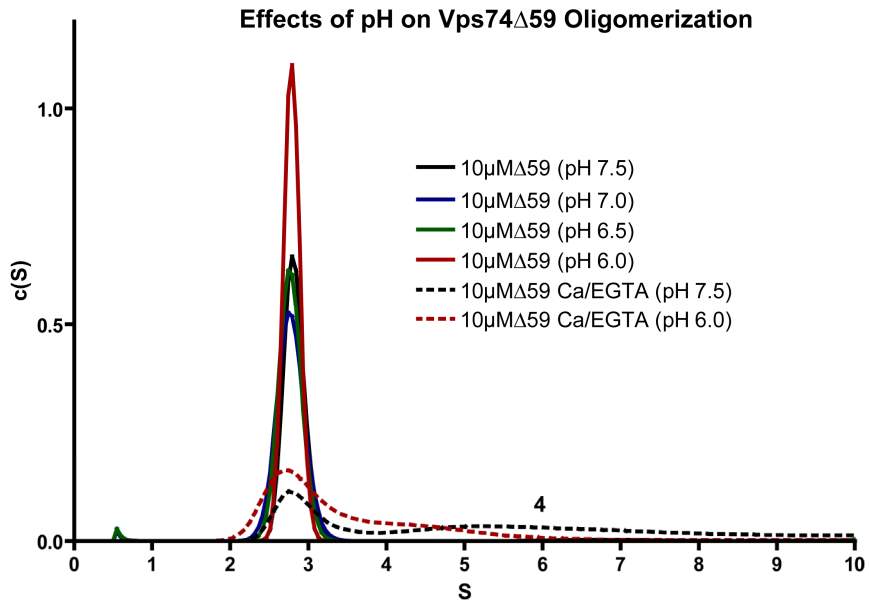

Figure 15. Vps74 oligomer formation occurs at pH 6.0, but to a lesser extent than at pH 7.5.

Samples of Vps74Δ59 between pH 7.5 and 6.0, with and without 10 mM $CaCl_2$ & 10 mM EGTA, were analyzed by sedimentation velocity analytical ultracentrifugation, and are plotted as c(S) distributions. In the absence of Ca/EGTA, Vps74Δ59 sediments as a single species. (The c(S) peak width and height varies, but the area under the curve is similar in all samples.) In the presence of Ca/EGTA at pH 7.5, Vps74Δ59 redistributes as monomer and tetramer. Oligomerization also occurs at pH 6.0, but to a lesser extent, populating dimer rather than monomer. As histidine residues do not stabilize any of the observed interfaces, this pH sensitivity is likely due to the pH dependence on the chelation of Ca^{2+} by EGTA.

amount of total Ca/EGTA in complex does not vary dramatically between pH 7.5 (99.7% of total, or 9.97 mM in this experiment) and pH 6.0 (91.4%, or 9.14 mM) [66], but the protonation state and conformation of the Ca/EGTA complex may, either of which may be critical for its interaction with Vps74.

We do not have a conclusive explanation for the effect of Ca/EGTA on Vps74 tetramerization. No strong electron density attributable to bound calcium or EGTA was evident in the structure, although the presence of such sites cannot be ruled out on the basis of these data alone. Whatever the mechanism of tetramer stabilization, this phenomenon allows us to test the effects of tetramer formation on mannosyltransferases binding. We find that addition of Ca/EGTA to the buffer in the GST-binding assay leads to a modest enhancement (approximately 2-fold) in binding of Vps74 to Kre2-GST (Fig. 16).

1.2.8 The β-hairpin is required for Vps74 function, suggesting a link between function and tetramer formation

Structural evidence suggests that the tetramer is physiologically relevant. To determine if oligomerization is required for Vps74 function, we asked whether Vps74 variants lacking the entire β-hairpin, Vps74Δ(197-208), or the hydrophobic FFLF motif (Fig. 8C), Vps74Δ(201-204), can functionally substitute for wild type Vps74 in yeast. Neither variant rescues the glycosylation defect (Fig. 17A) or the Kre2-GFP localization defect (Fig. 17C). These altered forms also fail to bind Kre2-GST (Fig. 17B). Thus, the hydrophobic β-hairpin is required for Vps74 function, and function is correlated with tetramer formation: solution conditions that promote oligomerization enhance the interaction with mannosyltransferases peptides, and Vps74 variants incapable of tetramer formation are non-functional in cells.

Finally, we observe that GFP fusions of Vps74 variants incorporating alterations to the β-hairpin, Vps74Δ(197-208) and Vps74Δ(201-204), fail to localize to the Golgi and instead adopt a cytosolic distribution (Fig. 17D). This suggests either that the recognition of resident Golgi proteins is requisite for Vps74 Golgi localization, or that the β-hairpin is directly involved in recognition of Golgi membranes (in fact, both are likely: see Section 1.3)

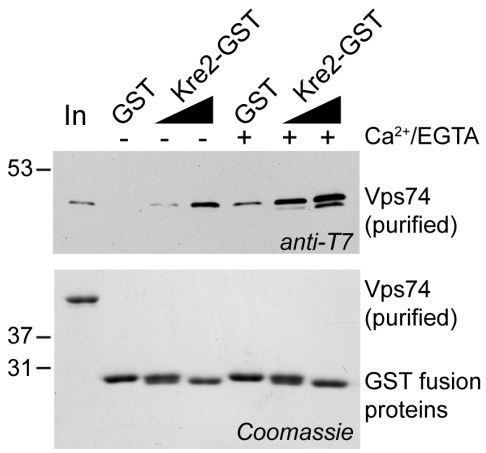

Figure 16. Stabilization of tetramer formation by addition of Ca^{2+}/EGTA enhances binding to Kre2-GST.

As in Fig. 3A, Kre2-GST constructs were immobilized to GSH agarose beads at varying ratios of Kre2-GST:GST alone. Beads were incubated with T7-tagged vps74 in binding buffer with or without 10 mM $CaCl_2$, 10 mM EGTA. A ≈ 2-fold enhancement in pulled down Vps74 is observed in the presence of Ca/EGTA.

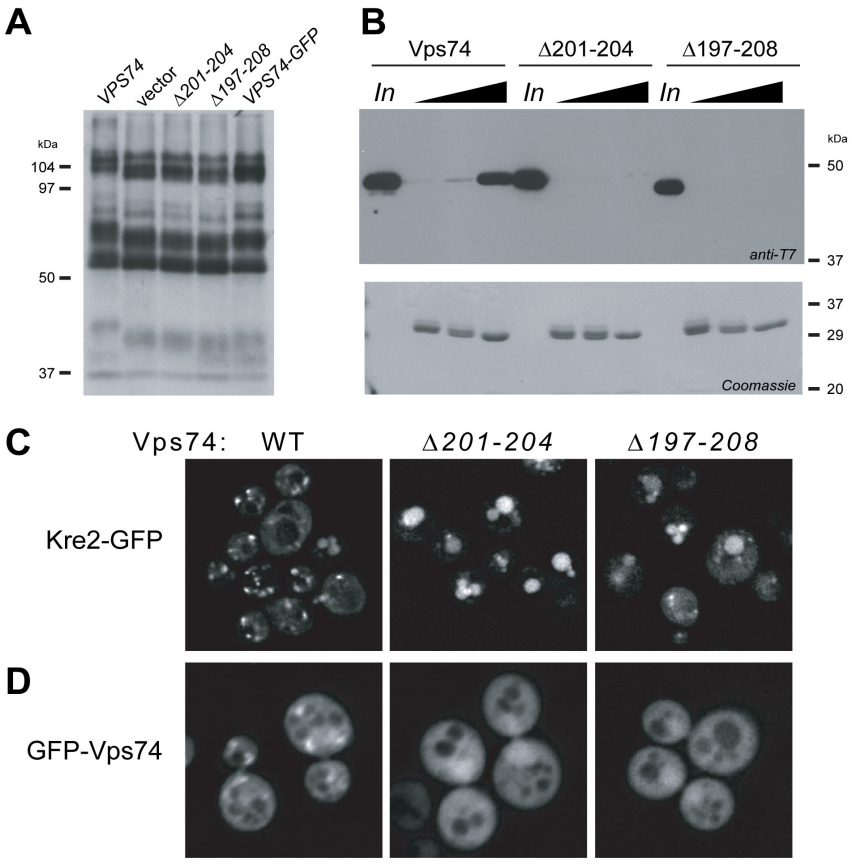

Figure 17. Alterations that disrupt tetramer formation abrogate Vps74 function.

(A) ^{35}S Cys/Met pulse-chase of *vps74Δ* cells transformed with the indicated constructs, as in Fig. 1A, Constructs incorporating alterations to the β-hairpin, or a C-terminal GFP fusion, do not rescue the glycosylation defect. (B) As in Fig. 3B, immobilized Kre2-GST fusions were used to pull down T7-tagged Vps74 or altered forms Vps74Δ(201-204) or Vps74Δ(197-208) from *E. coli* lysates. Lanes marked "In" contain 1% lysates input. (C) The localization of GFP tagged Kre2 was visualized by fluorescence microscopy in *vps74Δ* yeast expressing the indicated forms of Vps74. Punctate Golgi localization of Kre2 is only maintained by wild type Vps74. (D). The localization of the indicated GFP-Vps74 constructs expressed in *vps74Δ* yeast was visualized by fluorescence microscopy. GFP-Vps74 variants with alterations to the β-hairpin are cytosolic, rather than Golgi localized. (*Experiments done by J. Liu, C.S. Wood, C.G. Burd; adapted from [19].*)

1.3 Vps74 and GOLPH3 bind phosphoinositides, with specificity for PtdIns4*P*.

1.3.1 The crystal structure of GOLPH3 exhibits high structural homology with that of Vps74

To better understand the function of the animal homolog of Vps74, GOLPH3, we determined its X-ray crystal structure to 2.9 Å resolution (Fig. 18, Table 3). Crystals of a truncated version of GOLPH3 (GOLPH3Δ51) lacking the first 51 amino acids (approximately equivalent to the first 59 amino acids of Vps74, which are disordered in the Vps74 crystal structure and generally poorly conserved throughout fungi and animals; Fig. 9) were grown from a solution containing 0.9 M ammonium sulfate and, as discussed in detail below, contain a bound sulfate ion that suggests the location of a PtdIns4*P* binding site (see Section 1.3.2).

The GOLPH3 structure exhibits strong structural homology to that of Vps74 (Figs. 19 & 20) as expected based on sequence similarity, with an overall backbone atom root mean square deviation (RMSD) between GOLPH3 and Vps74 of ≈ 1.0 Å. Four parallel α-helices (α1, α2, α8, α12) are surrounded by 6 roughly orthogonal amphipathic helices, and, as observed in Vps74, a β-hairpin motif (β3/β4) projects away from this helical core. The major differences between GOLPH3 and Vps74 lie on the face opposite this β-hairpin. In Vps74 the α7/α8 loop and the region between α10 and α12 are longer than in GOLPH3 (Figs. 9 & 20). These longer loops, along with helix α11, pack tightly against the core of Vps74 on one face of the protein, creating a surface that is much less positively charged than the analogous surface of GOLPH3 (Figs. 20 & 22). GOLPH3Δ51 is a monomer in the crystallographic asymmetric unit, but the β-hairpin of one molecule makes an interaction with a symmetry related molecule that is equivalent to the "β-hairpin interface" observed in the Vps74 tetramer (Fig. 8, Table 4).

An intermolecular disulfide bond (C84-C108; Fig. 21) links symmetry related GOLPH3Δ51 molecules. Purified GOLPH3Δ51 was found to precipitate in the absence of reducing agent, and this precipitate re-dissolved upon the addition of reducing agent. Similarly, crystal formation was oxidation dependent, indicating that the observed disulfide bonds are

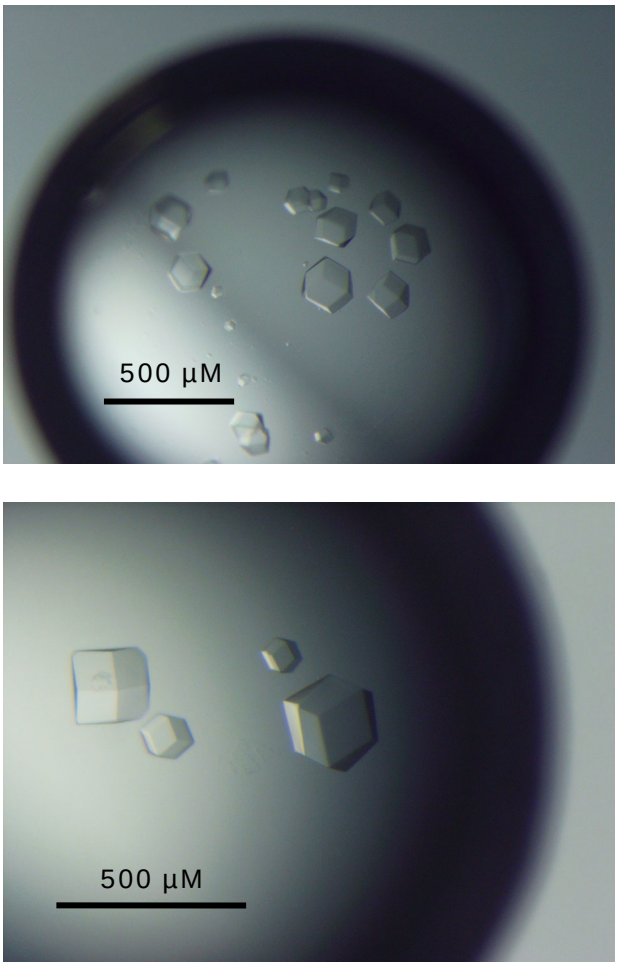

Figure 18. Crystals of GOLPH3Δ51.

Representative crystals of GOLPH3Δ51. 200 × 200 × 200 µm crystals grew in 0.9 M $(NH_4)_2SO_4$, 0.1 M MES, pH 6.0.

Table 3. GOLPH3 Crystallographic Statistics

	GOLPH3Δ51 (Se-met)[a]	GOLPH3Δ51 (native)
Data Collection Statistics[b]		
Space group	I432	I432
Unique cell dimensions	a, b , c = 174.4 Å	a, b , c = 174.6 Å
X-ray source	APS 23-ID-B	APS 23-ID-B
Wavelength	0.97951 Å	1.07221
Resolution limit	2.9 Å	2.3 Å
Observed/Unique	36,448/10,180	175,831/19,854
Redundancy	3.6 fold	8.9 fold
Completeness (%)	98.3 (99.6)	97.1 (78.2)
R_{sym}[c]	0.133 (0.586)	0.060 (0.619)
$\langle I/\sigma \rangle$	11.9 (2.4)	36.4 (3.2)
Refinement Statistics		
Resolution limits	50-2.9 Å	50-2.3 Å
No. reflections/No. test set	10,152/492	18,800/1013
R factor (R_{free})[d]	0.213 (0.266)	0.211 (0.266)
Model		
Protein	aa 59-297	aa 57-297
	6 sulfate ions	6 sulfate ions
	9 water molecules	109 water molecules
Total number of atoms	1,939	2050
RMSD bond lengths (Å)	0.017	0.016
RMSD bond angles (°)	1.779	1.655

a. Native data was collected on a selenomethionine derivitized protein crystal.

b. Numbers in parenthesis refer to last resolution shell.

c. $R_{sym} = \sum|I_h - \langle I_h \rangle| / \sum I_h$, where $\langle I_h \rangle$ = average intensity over symmetry equivalent measurements.

d. R factor = $\sum|F_o - F_c| / \sum F_o$, where summation is over data used in the refinement; R_{free} includes only 5% of the data excluded from the refinement.

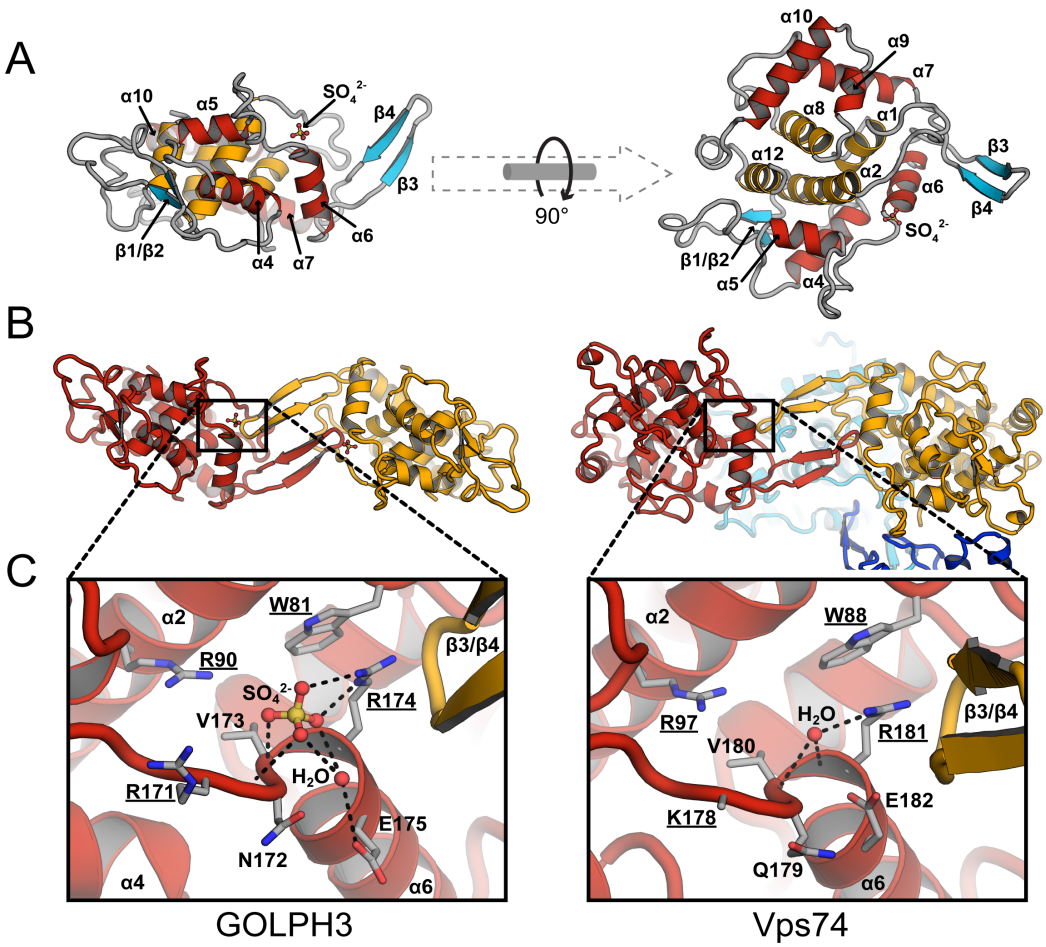

Figure 19. The structure of GOLPH3 indicates a putative PtdIns4*P* binding site.

(A) Orthogonal views of GOLPH3. Six amphipathic α-helices (red) surround four core α-helices (gold). A β-hairpin (β3/β4, cyan) projects away from the helical domain, as in Vps74 [19]. A well ordered sulfate is shown as ball-and-stick. (B) A β-hairpin interaction between crystallographic dimers of GOLPH3 (left) is nearly identical to an interaction stabilizing the Vps74 tetramer (right). The two additional subunits of the tetramer are shown (gold, cyan). (C) Magnified view of the sulfate binding pocket (boxed in B) of GOLPH3, and the homologous region of Vps74, with sulfate and surrounding amino acids shown as sticks. A water molecule occupies a similar position to the sulfate ion in GOLPH3. K178 is disordered in Vps74. Underlined amino acids were altered to alanine (see Section 1.3.5).

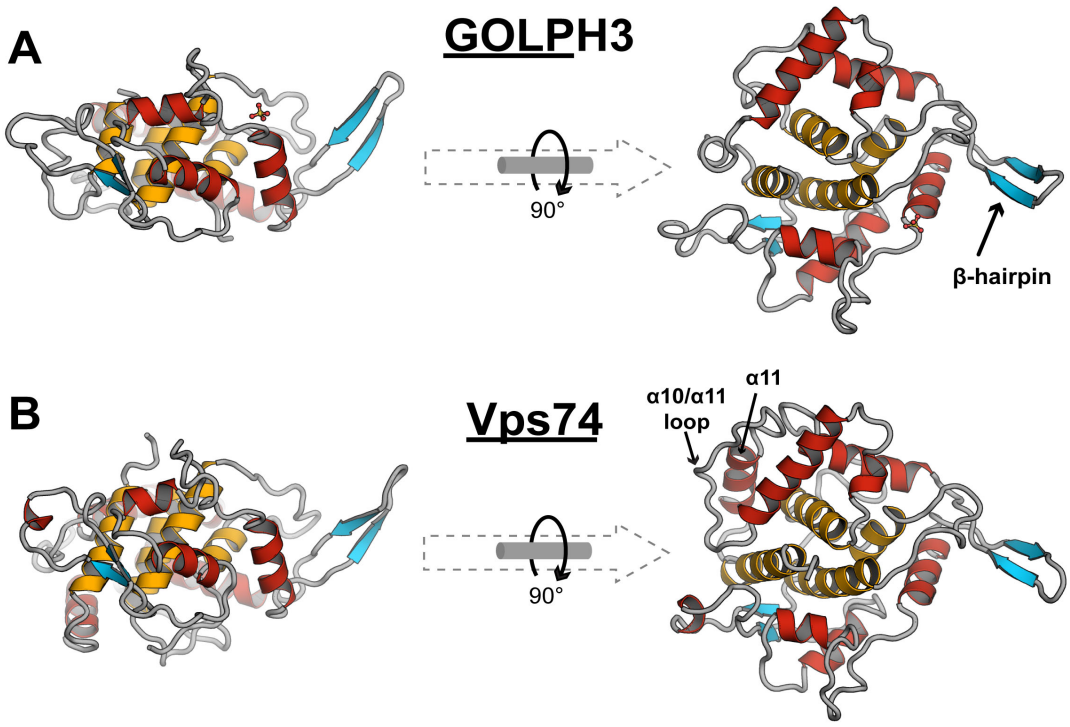

Figure 20. Comparison of GOLPH3 and Vps74 structures.

Orthogonal cartoon views of GOLPH3 (A) and a single subunit of Vps74 (B) (PDB ID 2ZIH, chain A). GOLPH3 and Vps74 share homologous α-helical folds. The four core hydrophobic helices (gold) are conserved in GOLPH3, as are most of the peripheral amphipathic helices (red). Helix α11 and the α10/α11 loop are absent in GOLPH3.

responsible for stabilizing the crystal form. The existence of these bonds prevents formation of interactions in the GOLPH3 crystal analogous to those observed in the Vps74 tetramer. GOLPH3 C84 is homologous to Vps74 S91, and in the central Vps74 subunit S91 lies near the β-hairpin. The presence of a disulfide bound molecule at this position sterically prevents both the "inner" and "peripheral" interactions. However, the absence of these interactions in the GOLPH3Δ51 crystal does not preclude their formation by GOLPH3 under physiological conditions.

1.3.2 The GOLPH3 structure reveals a conserved positively charged pocket, occupied by a bound sulfate ion

Our studies of Vps74 reveal a direct but relatively weak interaction between Vps74 and the cytosolic regions of mannosyltransferase enzymes, which may not be sufficient to account for the robust localization of Vps74 and GOLPH3 to the Golgi apparatus. Thus we considered the possibility that these proteins recognize other components enriched in the Golgi, including the phospholipid PtdIns4*P*. The structure of Vps74 provides no specific indication of a phospholipid binding site. By contrast, several ordered sulfate ions from the crystallization solution are modeled in the final refined GOLPH3Δ51 structure. Crystallographically bound anions have served as indicators for the presence of lipid headgroup binding sites in several instances, including the p47^{phox} PX domain [81] and the Hrs FYVE domain [82]. We asked whether this may also be true in GOLPH3. Of the six observed anions, the most well ordered and the one making the most extensive contacts with GOLPH3 (151 Å^2 buried surface area, compared to 95 – 125 Å^2 buried by the other sulfates) lies in a positively charged pocket formed by the α1-α2 loop, α2, and α5 (Fig. 19C & Fig. 22). This pocket is also observed in the *S. cerevisiae* Vps74 structure (Figs. 19C & 24) and the amino acids lining this pocket are well conserved throughout fungal and animal lineages (Fig. 9 & Fig. 23). The sulfate in the GOLPH3 pocket is stabilized by electrostatic interactions with R174, hydrogen bonds to backbone atoms of residues N172 and V173, and indirect hydrogen bonds to α5 via an ordered water molecule (Fig. 19).

Table 4. Details of the GOLPH3 β-hairpin interaction

	β-hairpin interaction
Number of Hydrogen bonds (≤ 3.5 Å)	6
Salt bridge interactions	2
Area occluded from solvent[a] (Å^2)	1095 Å^2
Polar area (% of total)	312 Å^2 (28 %)
Apolar area (% of total)	783 Å^2 (72 %)
Shape complementarity[b]	0.68

a. Average area occluded on each subunit.

b. Shape complementarity calculated with the SC module in CCP4 [8, 83].

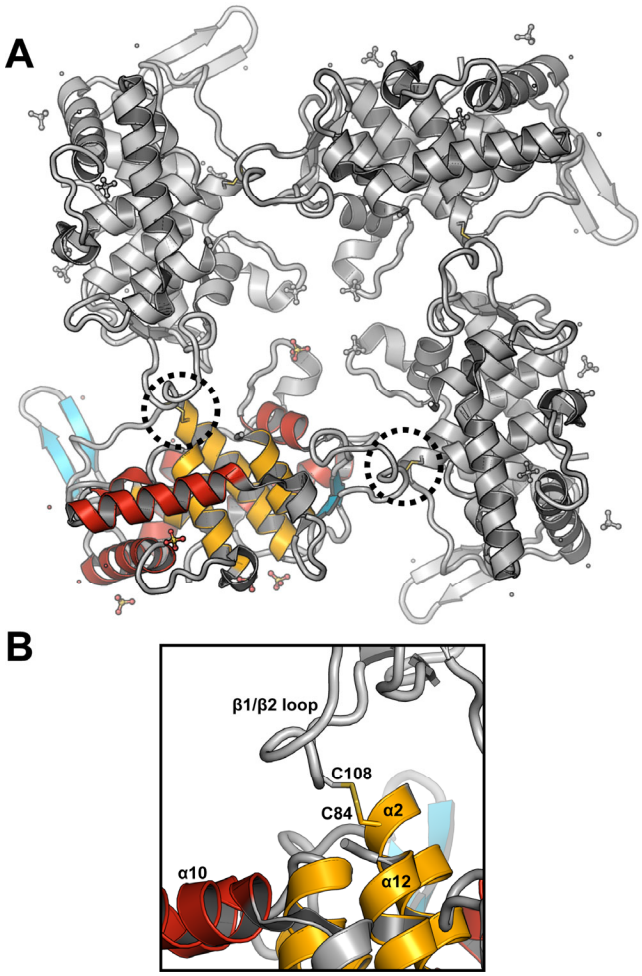

Figure 21. Disulfide bonds between symmetry related GOLPH3 molecules stabilize the crystal structure.

(A) Cartoons show GOLPH3 molecules related by crystallographic four-fold symmetry. Disulfide bonds (sticks, and indicated by dashed circles) link each GOLPH3 molecule (for example, the colored copy) to two symmetry related copies (gray). Four GOLPH3 molecules form a single covalently linked unit. Additional molecules are related by a two-fold symmetric β-hairpin interaction (Fig. 19B) and minor crystallographic contacts. (B) A detailed view of the disulfide bond between C84 of one GOLPH3 molecule (colored) to a C108 on a symmetry related copy. GOLPH3 is a cytosolic molecule, so these disulfide bonds are not expected to occur physiologically.

The features of this basic pocket are reminiscent of the binding sites of some structurally characterized phosphoinositide binding modules (Fig. 24). PLCδ_1 specifically binds PtdIns(4,5)P_2 in a basic binding pocket between loops β1-β2 and β3-β4 of its PH domain [84]. Binding is stabilized by electrostatic interactions between the inositol phosphates and five basic side chains, and by polar contacts with serine and tryptophan residues. Positively charged pockets in the PX domains of SNX3 and p40^{phox} specifically bind PtdIns3*P* [85, 86], stabilized by electrostatic interactions (one or two arginine side chains coordinate the 3-phosphate; a lysine interacts with 1-phosphate), by hydrogen bonds, and by van der Waals contacts between a conserved tyrosine and the inositol ring. Specific binding of the EEA1 FYVE domain to PtdIns3*P* is stabilized by the interaction of one arginine side chain with each phosphate, and by polar interactions with the inositol ring [87]. The GOLPH3 and Vps74 pocket includes three conserved basic residues (R90, R171, and R174 in GOLPH3) that may interact with the 1- and 4-phosphates of PtdIns4*P*, a conserved tryptophan (W81) that may pack against the inositol ring, and additional polar side chains that may stabilize the lipid head group (Figs. 19C & 24). In addition there is evidence that membrane insertion of conserved hydrophobic side chains contributes to the stability of membrane interactions [85-87]. The GOLPH3/Vps74 pocket is situated near the base of the hydrophobic β-hairpin, and near the hydrophobic pocket that accepts the tip of the partner β-hairpin to stabilize the β-hairpin interaction. Hydrophobic side chains from the β-hairpin may stabilize lipid binding by projecting into the lipid bilayer.

We have attempted to determine the crystal structure of phosphoinositide bound to GOLPH3Δ51 by soaking and by co-crystallizing soluble short chain phosphoinositides or headgroups alone into GOLPH3Δ51 crystals. Data were collected on GOLPH3Δ51 crystals soaked for over two months in a cryo-stabilizing solution containing 0.5 mM di-C_4-PtdIns4*P*. These crystals diffracted to a resolution of 2.3 Å (Table 3), but the resulting electron density is consistent with a bound sulfate in the proposed PtdIns4*P* binding pocket, with no evidence of lipid headgroup (Fig. 25). The high concentration of sulfate present in the crystallization buffer likely prevents binding of PtdIns4*P*.

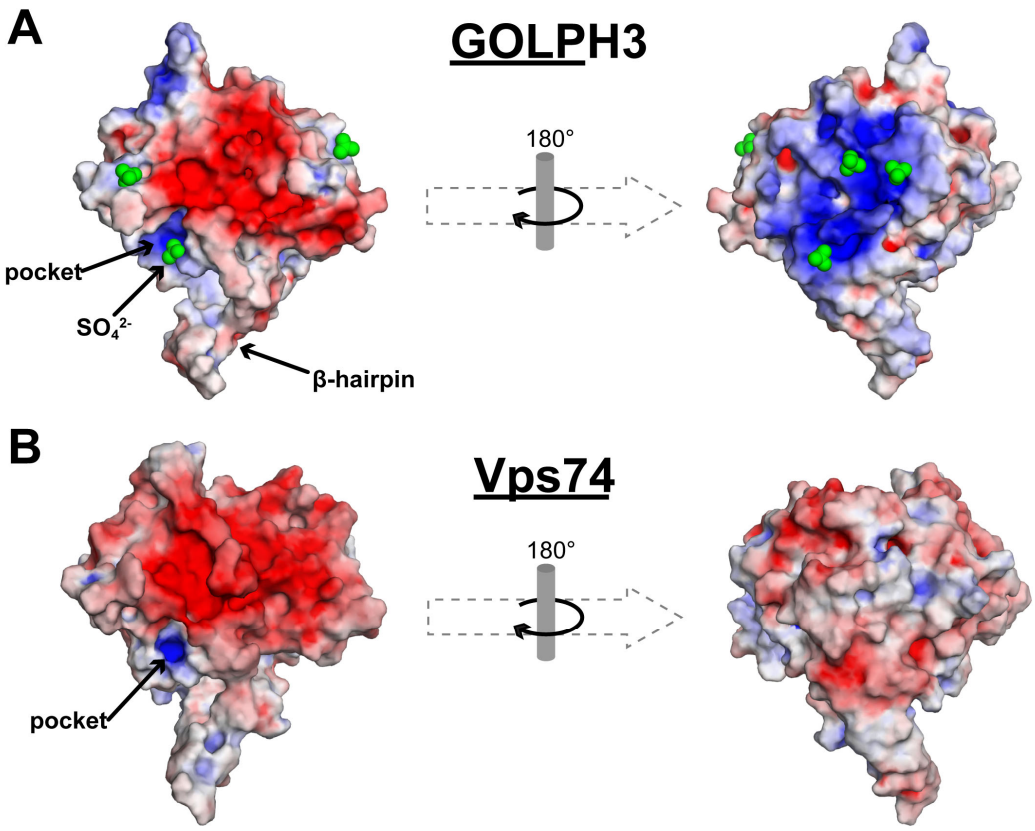

Figure 22. Comparison of the surface electrostatics of GOLPH3 and Vps74 monomers.

The surface electrostatic potential of GOLPH3 and Vps74 monomers is shown, on a scale of -8 kT (red) to +8 kT (blue). Crystallographically bound sulfate ions in GOLPH3 are shown as spheres. A well ordered sulfate ion lies in a positively charged pocket in GOLPH3, near the hydrophobic β-hairpin motif. This pocket is conserved in Vps74. The face opposite the pocket (rightmost orientation) is less charged in Vps74 than in GOLPH3, due largely to the amino acid composition of α11 and the α10/α11 loop, which are absent in GOLPH3. Electrostatic surfaces were generated by the Adaptive Poisson-Boltzmann Solver (APBS) [13] in Pymol [12].

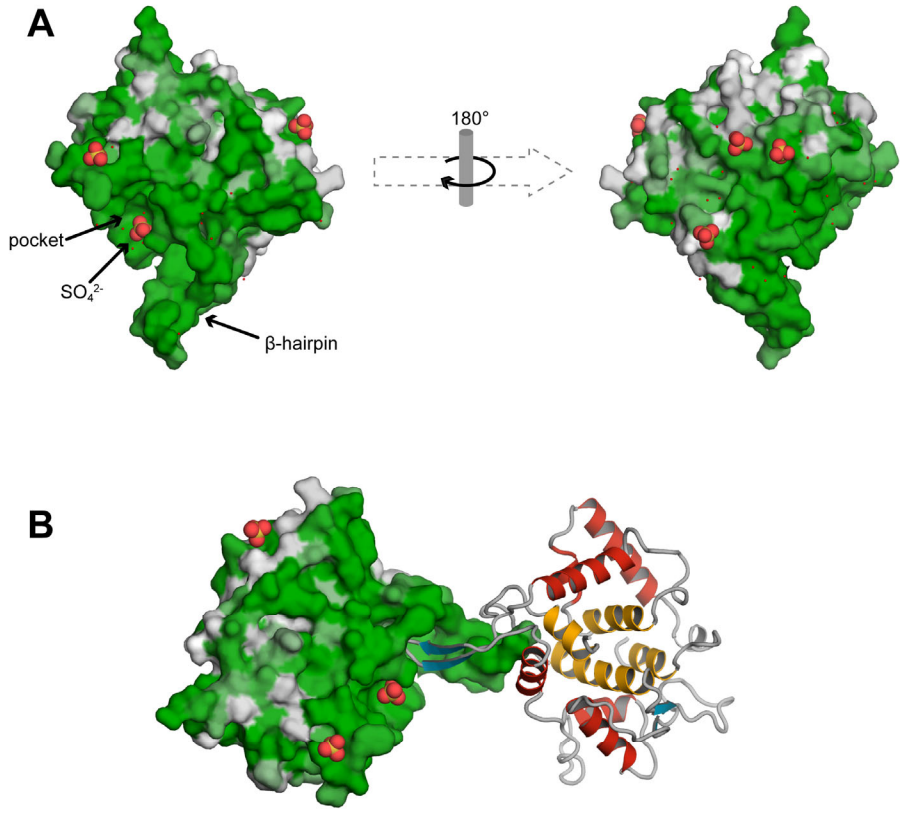

Figure 23. Conservation of GOLPH3 surface resides.

Surface representations of the GOLPH3 crystallographic monomer (A) and β-hairpin dimer (B) are shown. Sequence conservation, based on a multiple sequence alignment between animal homologs of GOLPH3, is color coded by residue on a scale of green (most conserved) to white (least conserved). Residues involved in the β-hairpin interaction and the nearby sulfate binding pocket exhibit a high degree of conservation.

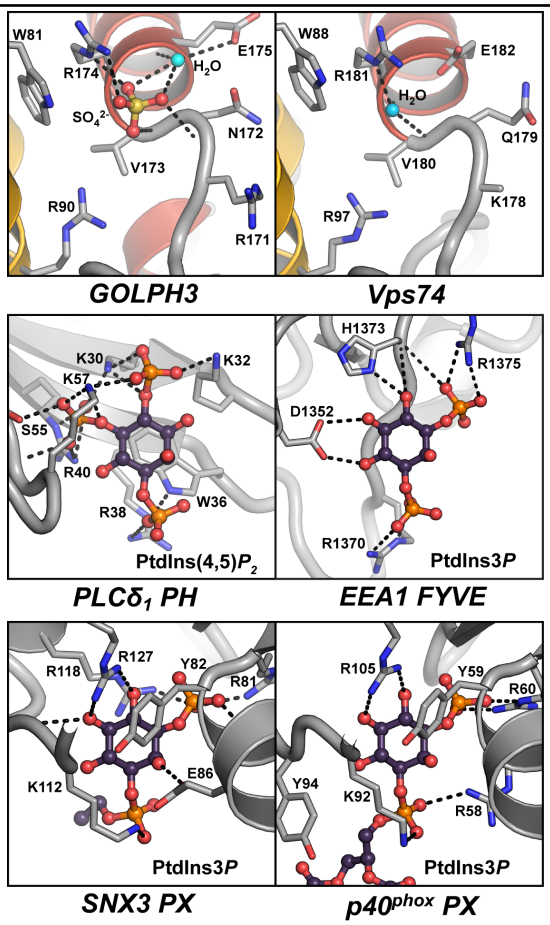

Figure 24. The GOLPH3 and Vps74 putative PtdIns4*P* binding pocket shares features with structurally characterized PI binding motifs.

Detailed views of the putative PtdIns4*P* binding pockets of GOLPH3 and of Vps74 are shown, as well as of the phosphoinositide binding pockets of PLCδ_1 PH domain (PDB ID: 1MAI), SNX3 PX domain (1OCU), p40^{phox} PX domain (1H6H), and EEA1 FYVE domain (1JOC), bound to their respective ligands. Ligands are represented as ball and sticks, with carbon atoms colored purple. A bound sulfate occupies the center of the pocket in GOLPH3, and an ordered water molecule occupies approximately the same position in Vps74.

1.3.3 Disruption of a PtdIns kinase and PtdIns phosphatase affect Vps74/GOLPH3 localization

The existence of a well conserved positively charged surface cavity occupied by sulfate in the GOLPH3 crystal structure suggests that GOLPH3 may bind phosphoinositides (PIs). Further evidence for this interaction comes from a yeast screen to identify gene deletions or temperature-conditional alleles that cause defects in the localization of GFP-Vps74.

In wild type yeast, Vps74 is localized primarily to the Golgi apparatus, with some occurrence in the cytosol. GOLPH3, which can functionally substitute for Vps74 in yeast, localizes even more robustly to Golgi. In the *pik1-83* and *frq1-1* temperature-conditional yeast strains, after incubation at the restrictive temperature both GFP-Vps74 and GFP-GOLPH3 lose their Golgi localization and adopt a cytosolic and nuclear distribution (Fig. 26A). *PIK1* encodes the only Golgi associated PtdIns 4-kinase in yeast, and *FRQ1* encodes a Golgi-localized calcium-binding protein that binds and activates Pik1 [88]. Disruptions in Stt4 and Lsb6, the other two PtdIns 4-kinases in yeast, do not alter Vps74 localization (data not shown). Thus, GOLPH3 and Vps74 Golgi localization in yeast requires a functioning Golgi PtdIns 4-kinase. This dependence may indicate a direct interaction between Vps74/GOLPH3 and Pik1/Frq1, or an interaction between Vps74/GOLPH3 and the product of this enzyme, PtdIns4*P*.

To distinguish between these possibilities, we assessed the distribution of Vps74 and GOLPH3 in *sac1Δ* yeast, which lack the Golgi and ER localized Sac1 phosphatidylinositol phosphatase. Sac1 dephosphorylates several phosphoinositide substrates, but Golgi localized Sac1 is thought to function largely as a PtdIns4*P* phosphatase [89, 90]. Deletion or mutation of *sac1* results in a several fold increase in PtdIns4*P* levels and an enrichment of PtdIns4*P* in the ER [89]. In *sac1Δ* cells, Vps74 and GOLPH3 localize to the ER as well as to the Golgi (Fig. 26B). Thus, Vps74 and GOLPH3 adopt a cytosolic distribution in yeast deficient in the synthesis of Golgi PtdIns4*P*, and localize to the ER when PtdIns4*P* is enriched in this organelle. Together these data strongly suggest that Vps74 and GOLPH3 specifically recognize PtdIns4*P*.

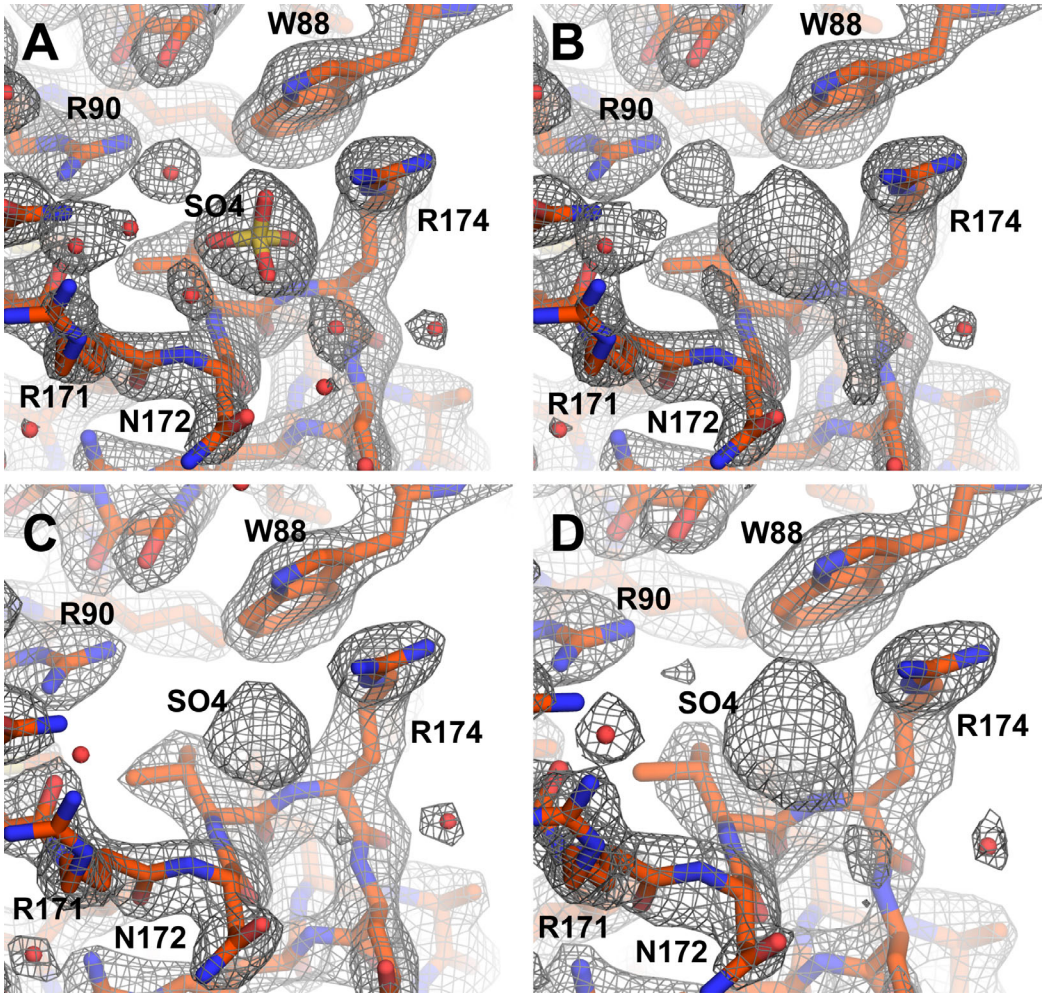

Figure 25. Electron density surrounding the ordered sulfate in GOLPH3.

Crystals of GOLPH3Δ51 were soaked with 0.5 mM diC_4-PtdIns4*P*. Stick representations are shown of the PI binding site in the resulting model, along with 1.5 σ $2f_O$-f_C maps calculated with (A) and without (B) a bound sulfate ion and neighboring water molecules; a map after density modification by DM [8], and a simulated annealing omit map by CNS calculated without the sulfate ion and neighboring waters [27]. The central region of electron density corresponds well with a sulfate ion or an inositol phosphate. However, no electron density is observed that correspond well with the inositol ring of PtdIns4*P*. View is approximately equivalent to that in Fig. 19C.

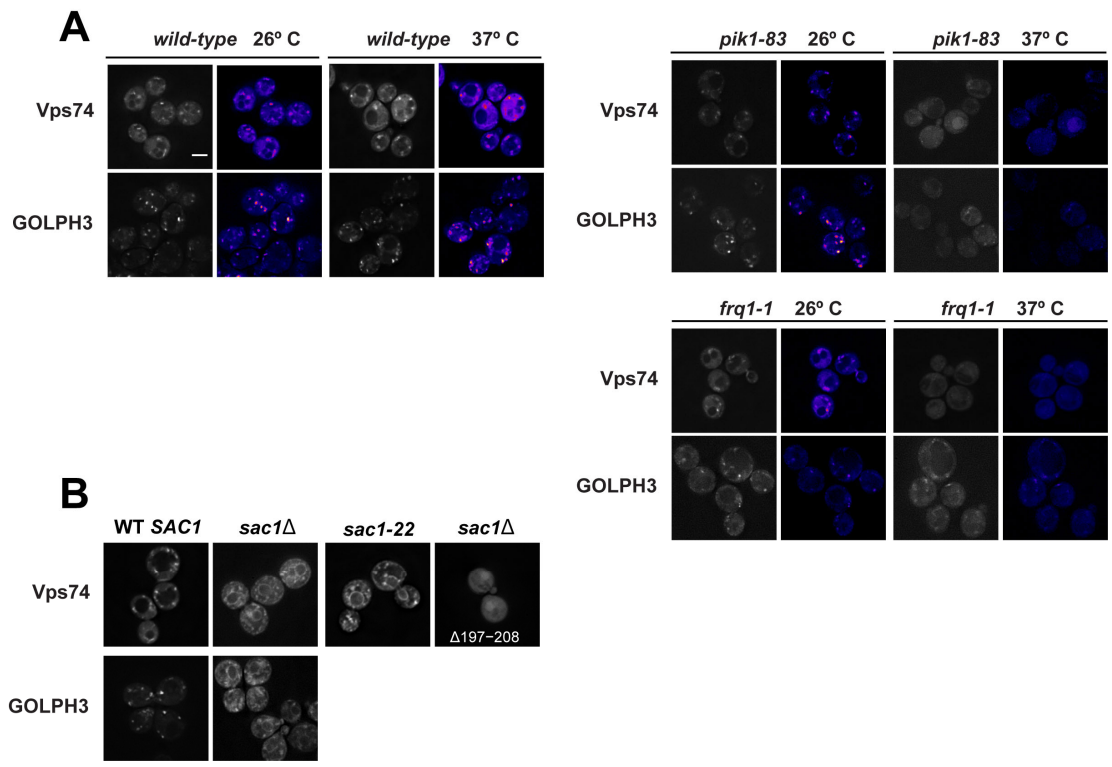

Figure 26. Golgi localization of Vps74 and GOLPH3 is altered upon disruption of the Pik1/Frq1 PtdIns 4-kinase or the Sac1 lipid phosphatase.

(A) Fluorescence microscopy showing the localization of GFP-tagged Vps74 and GOLPH3 in wild type yeast, as well as *pik1* and *frq1* temperature-conditional strains, at permissive (26°C) and restrictive (37°C) temperatures. Grayscale and false color images (yellow indicating highest intensity) are shown. At the restrictive temperature in *pik1* and *frq1* mutants, Vps74 and GOLPH3 distribution is cytosolic rather than Golgi associated. (B) Fluorescence microscopy showing the localization of GFP tagged wild type GOLPH3, Vps74, and Vps74Δ(197-208) is shown in wild type yeast, and strains carrying mutations or deletions of *sac1*. Disruption of *sac1* drives GOLPH3 and Vps74 to the ER, in addition to the Golgi apparatus.

(From work by C.S. Wood and C.G. Burd. Adapted from [28].)

1.3.4 Vps74 and GOLPH3 bind PtdIns4*P* in lipid binding assays

Yeast localization data, in conjunction with structural evidence, argue that Vps74 and GOLPH3 directly bind PtdIns4*P*. Surface plasmon resonance (SPR; Biacore) was used to assay the interaction between purified Vps74 or GOLPH3 with several PIs. Vesicles composed of dioleoylphosphatidylcholine (DOPC) with or without 3 % (mole/mole) PtdIns3*P*, PtdIns4*P*, PtdIns5*P,* or PtdIns(4,5)P_2 were applied to an L1 Biacore sensorchip [91]. Equilibrium binding of Vps74 and GOLPH3 to these lipid surfaces was assayed (Fig. 27) and dissociation constants (K_D values) were determined (Table 5).

Vps74 binds to PtdIns4*P* with a K_D of 8.9 ± 0.3 μM, an affinity at least fivefold tighter than to the other lipid surfaces tested. (The exact selectivity is uncertain due to the upper limit in K_D determination.) GOLPH3 binds PtdIns4*P* with a K_D of 2.6 ± 0.2 μM, approximately fourfold tighter than Vps74. This is consistent with the more robust targeting of GOLPH3 to Golgi in yeast (see Section 1.3.3). GOLPH3 likely exhibits poorer selectivity for PtdIns4*P*, having a substantial K_D of 10.2 ± 0.5 μM for the interaction with PtdIns3*P* surfaces, and affinities of approximately 25 μM for PtdIns5*P* and PtdIns(4,5)P_2 surfaces. Vps74 and GOLPH3 affinities for PtdIns4*P* lipid surfaces are similar to those of proteins targeted to the Golgi via PH domains (OSBP, FAPP1, Osh2), exhibiting K_D values from 1 to 6 μM [92, 93].

1.3.5 Mutations to the Vps74 or GOLPH3 residues lining the charged pocket abolish lipid binding, localization, and function

To test whether the amino acids lining the GOLPH3 and Vps74 sulfate binding pocket are important for PtdIns4*P* binding, *in vivo* targeting, and the function of Vps74 and GOLPH3, we created variants with alanine substituted at two of the amino acids underlined in Fig. 19C: Vps74(W88A/R87A) and Vps74(K178A/R181A) in Vps74; GOLPH3(W81A/R90A) and GOLPH3(R171A/R174A) in GOLPH3. In yeast and in *E. coli* these variants are expressed in equivalent amounts compared to wild-type GFP-tagged Vps74 and GOLPH3, indicating that these substitutions do not grossly affect the stability of the proteins (not shown). Circular

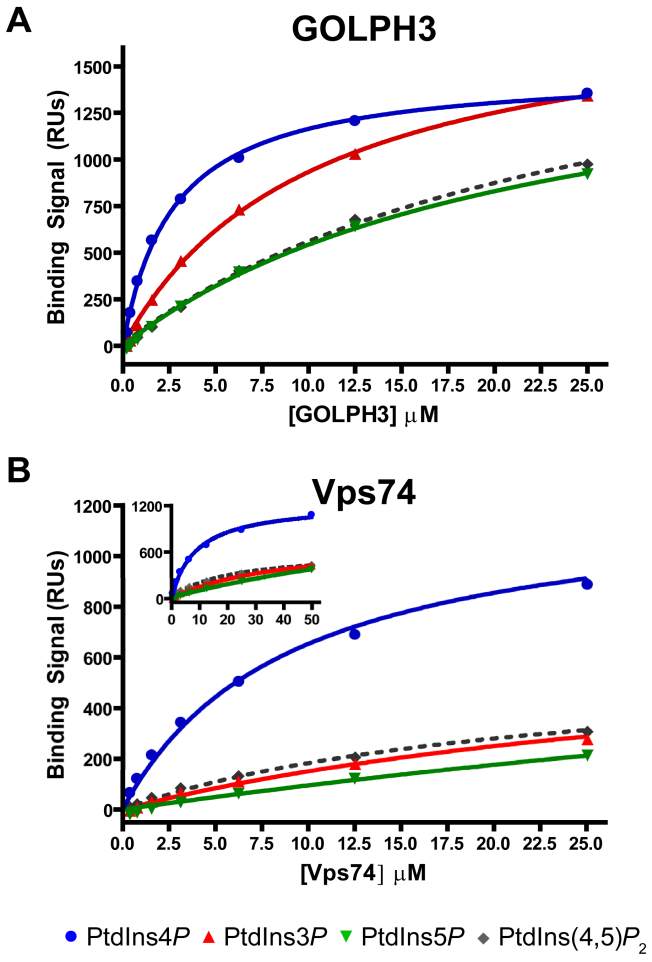

Figure 27. GOLPH3 and Vps74 specifically bind PtdIns4*P*.

Biacore SPR of GOLPH3 (A) and Vps74 (B) binding to phosphoinositides. Samples of the indicated concentration of GOLPH3 or Vps74 were passed over a L1 chip (Biacore) prepared with 3% (mole/mole) of the indicated phosphoinositide in a DOPC background. The steady-state SPR response for a representative dataset, corrected for background binding to a DOPC surface, is plotted against protein concentration. The inset (B) shows data for 50 μM Vps74. Curves were generated by a fit to a simple one-site binding equation. K_D values from at least three independent experiments are listed in Table 5.

Table 5. K_D values for the binding of Vps74 and GOLPH3 to phosphoinositides, determined from SPR.

	PtdIns4*P*	PtdIns3*P*	PtdIns5*P*	PtdIns(4,5)P_2
Vps74	8.9 ± 0.3 μM	> 50 μM	> 50 μM	> 50 μM
GOLPH3	2.6 ± 0.2 μM	10.2 ± 0.5 μM	≈ 25 μM	≈ 25 μM

dichroism (CD) spectroscopy of Vps74 variants further confirms that the alterations do not grossly destabilize the protein (Fig. 28). In SPR-based assays (Fig. 29), none of the Vps74 or GOLPH3 mutants detectably binds PtdIns4*P*, or any other phosphoinositide. Moreover, when expressed in *vps74*Δ and *sac1*Δ cells, the mutant forms of Vps74 and GOLPH3 do not localize to the Golgi or to the ER and are observed only in the cytosol and nucleus (Fig. 30). Importantly, the function of Vps74 is also abolished by these mutations: in cells expressing each mutant form as the sole source of Vps74, a GFP-tagged form of a medial Golgi mannosyltransferase, Kre2-GFP, is not efficiently retained in the Golgi apparatus, just as in *vps74*Δ cells (Fig. 30). These results firmly establish that the sulfate binding pocket originally identified in the X-ray crystal structure of human GOLPH3 is critical for PtdIns4*P* binding and Golgi targeting of GOLPH3 and Vps74, and they demonstrate that PtdIns4*P* binding is essential for the function of Vps74 in Golgi homeostasis.

Additionally, we note that deletion of the β-hairpin, which is required for Vps74 oligomerization and function [19], abolishes binding to PtdInsP*4* (Fig. 29). This raises the possibility that hydrophobic side chains in the β-hairpin may stabilize membrane binding by insertion into the hydrophobic milieu, as has been suggested for binding of PX and FYVE domain to PtdIns3*P* containing membranes [85, 86].

1.4 Vps74/GOLPH3 homologs among eukaryotes and bacteria

1.4.1 Homologs occur in unikont eukaryotes

We show that Vps74 and GOLPH3 function to maintain the spatial organization of Golgi proteins, and that this occurs in part through recognition of PtdIns4*P* at Golgi membranes. Sequence searches by BLAST [94] reveal one or more GOLPH3/Vps74 homologs in most unikonts [95], including animals, choanozoa (protists closely related to animals; a homolog occurs in *Monosiga brevicollis*), fungi, and amoebozoa (ameboid protists, including slime molds; a homolog occurs in the slime mold *Dictyostelium discoideum*). These proteins share a high

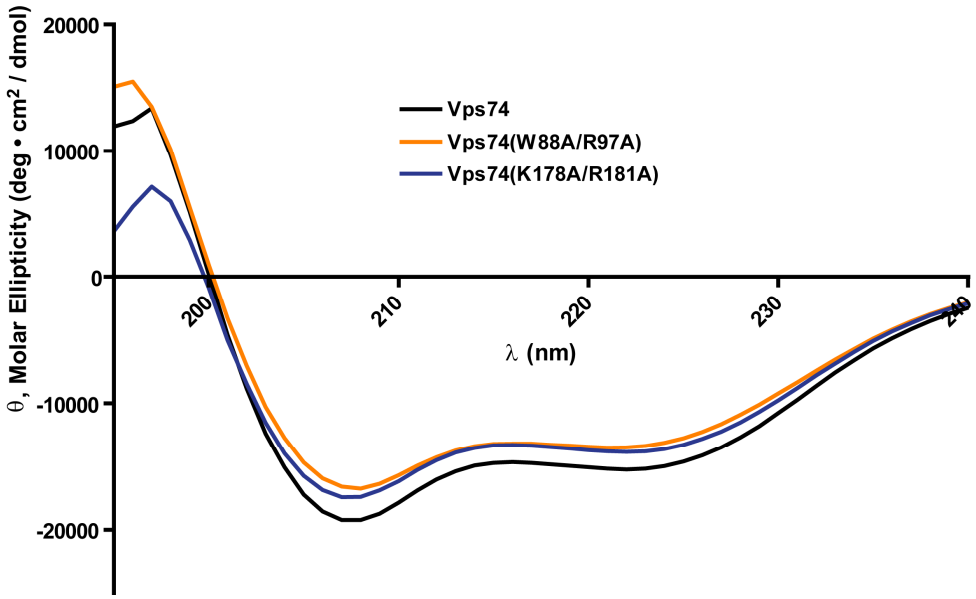

Figure 28. Circular dichroism spectra of Vps74 PtdIns4*P* binding pocket mutants.

Circular dichroism spectra at 37°C of wild type and the indicated Vps74 variants is shown. Samples contained 1.5 μM protein in 5 mM sodium borate, 150 mM sodium fluoride, pH 7.5. Wild type and mutant Vps74 are predominantly α-helical, and the point mutations have introduced no gross changes in secondary structure. Wild type Vps74 is untagged (H6-Vps74 after TEV digestion), while Vps74(W88A/R87A) and Vps74(K178A/R181A) contain N-terminal T7 and hexahistidine tags. This additional unstructured region in the mutants accounts for the slight overall decrease in secondary structure. GOLPH3 was not analyzed by CD due to poor solubility in UV transparent buffers at 37°C.

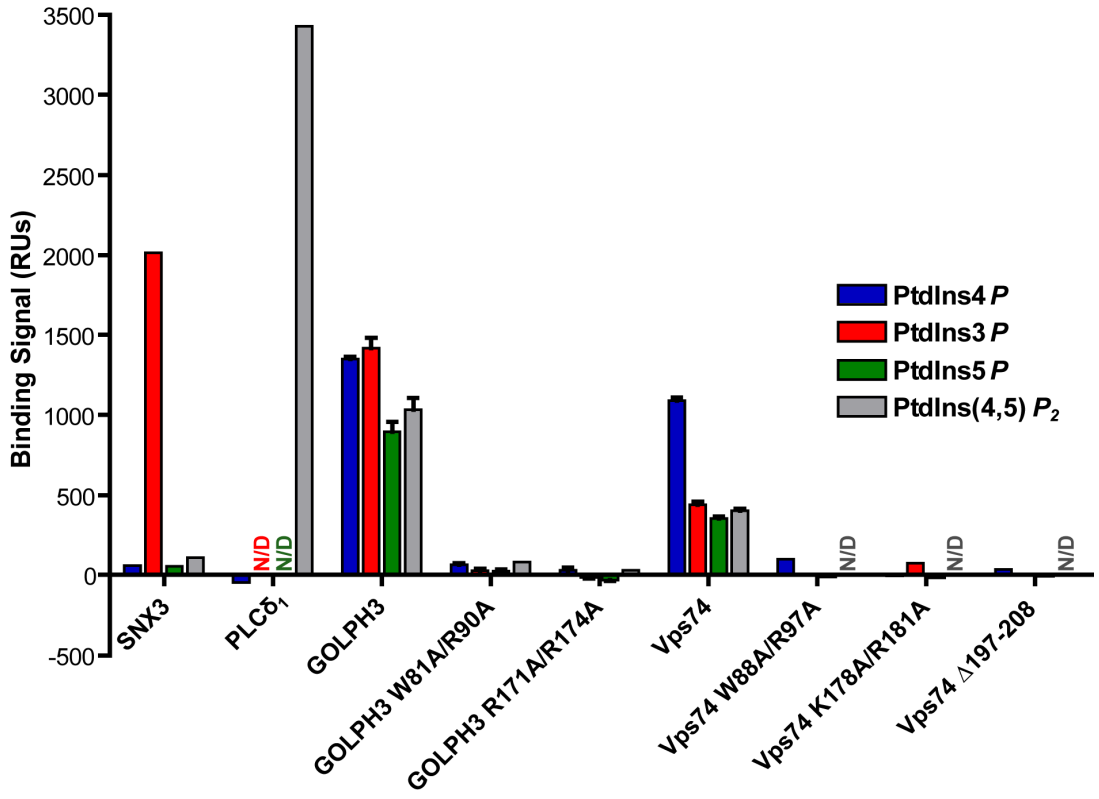

Figure 29. Equilibrium binding of GOLPH3 and Vps74 variants to phosphoinositide surfaces.

Equilibrium SPR response values are plotted (corrected for background binding to DOPC surfaces) from the interaction of a single concentration of the indicated protein with lipid surfaces incorporating 3% (mole/mole in a DOPC background) of the indicated phosphoinositides. SNX3 and PLCδ_1 controls have specificity for PtdIns3*P* and PtdIns(4,5)P_2 surfaces, respectively. Alteration of amino acids lining the sulfate binding pocket abolishes the interaction with lipid surfaces, as does disruption of the conserved β-hairpin. Control proteins were assayed at 25 μM; 50 μM samples were used for all other proteins.

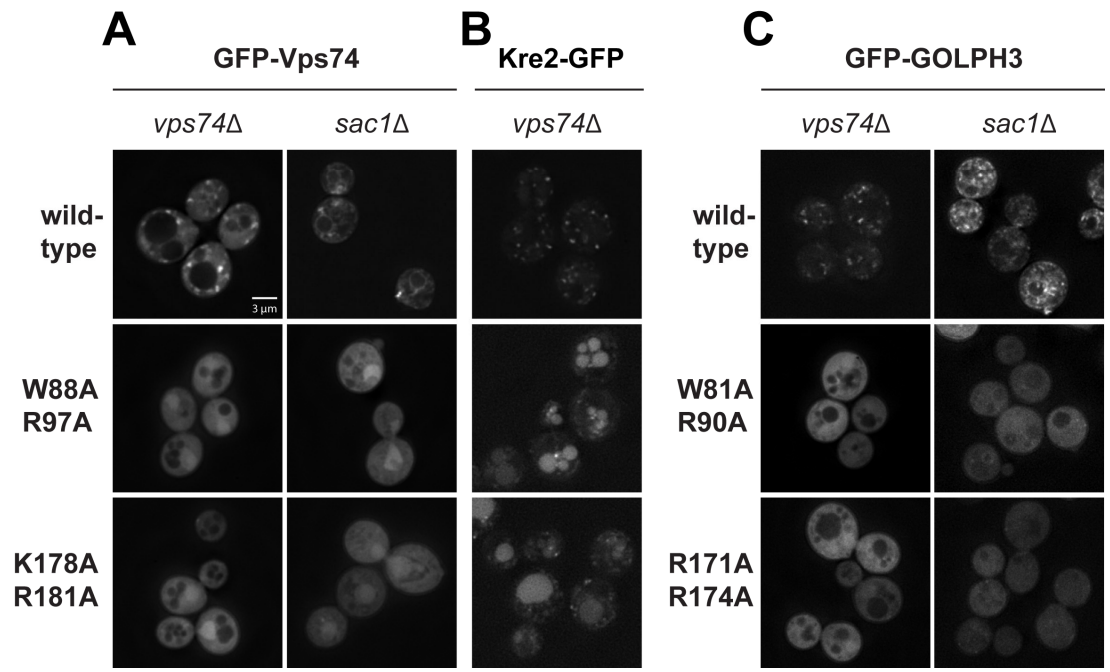

Figure 30. Disruption residues surrounding the PtdIns4*P* binding pocket abolishes GOLPH3 and Vps74 function.

(A) The localization of the indicated GFP-tagged variants of Vps74 (left legend) is shown by confocal microscopy in *vps74Δ* or *sac1Δ* deletion strains. (B) The localization of Kre2-GFP is shown by confocal microscopy in *vps74Δ* cells transfected with the indicated forms of Vps74 (left legend). (C) The localization of the indicated GFP-tagged GOLPH3 variants (left legend) is shown by fluorescent microscopy in the *vps74Δ* or *sac1Δ* deletion strains. GFP-GOLPH3 and GFP-Vps74 incorporating alterations to pocket residues lose Golgi localization and are found in the cytosol in *vps74Δ* and *sac1Δ* cells. Golgi localization of Kre2-GFP is also lost in cells expressing altered forms of Vps74, and Kre2-GFP accumulates in the vacuole.

(From work by C.S. Wood and C.G. Burd. Adapted from [28].)

degree of sequence identity (38% between *S. cerivisciae* Vps74 and *H. sapiens* GOLPH3; Fig. 31), which argues that they serve an important cellular function. However, homologs do not occur in all eukaryotes; notably, none has been identified in plants. (At the time of writing, we note that a hypothetical *Zea maize* gene product, GenBank ID: ACF86141.1, has significant sequence similarity to Vps74. However, the sequence of this protein is identical to that of Vps74 from the fungal plant pathogen, *Gibberella zeae*. This *Z. maize* entry is likely a sequencing artifact.) Eukaryotes lacking Vps74/GOLPH3 homologs possess functioning Golgi organelles, of similar topology and composition to those of unikonts.

Curiously, we have also identified Vps74/GOLPH3 homologs among some prokaryotes. We term these proteins Bhovs (bacterial homologs of Vps74). A BLAST search against prokaryotic sequences using *S. cerevisaie* Vps74 as a search query yields multiple hits with expectation scores (e-scores) below 1×10^{-4}. Bhovs occur in at least seven prokaryotic genera, encompassing Gram-positive and -negative species across five bacterial phyla, in species that occupy diverse environments.

Bhovs are not present in all bacteria (for instance, no homolog is detected in *E. coli*), and none has been identified in Archaea, which are thought to be more closely related to eukaryotes. This discontinuous distribution of homologs across distantly related organisms suggests that Vps74/GOLPH3 and Bhov proteins are related by horizontal gene transfer (i.e., xenologs). Bhov proteins are more similar to one another than to their eukaryotic homologs and occur in free living bacteria that have no special association with eukaryotes. This suggests that eukaryotic Vps74/GOLPH3 arose from Bhov, rather than vice versa.

1.4.1 Structure and function of the bacterial homologs of Vps74/GOLPH3

To our knowledge, no Bhov has been functionally characterized. However, we are able to make some functional predictions based on sequence comparison to eukaryotic homologs (Fig. 32). Key Vps74/GOLPH3 structural features are present in Bhov proteins. Hydrophobic

A

Sequence identity between indicated Vps74/GOLPH3 homologs	GOLPH3 *H. sapiens*	GOLPH3 *C. elegans*	GOLPH3 *M. brevicollis*	Vps74 *S. pombe*	Vps74 *S. cerevisiae*	GOLPH3 *D. discoideum*	Bhov *I. loihienses*
GOLPH3 *H. sapiens*	1.00						
GOLPH3 *C. elegans*	0.63	1.00					
GOLPH3 *M. brevicollis*	0.65	0.59	1.00				
Vps74 *S. pombe*	0.40	0.35	0.43	1.00			
Vps74 *S. cerevisiae*	0.38	0.33	0.37	0.53	1.00		
GOLPH3 *D. discoideum*	0.15	0.14	0.17	0.14	0.16	1.00	
Bhov *I. loihienses*	0.19	0.20	0.20	0.18	0.18	0.19	1.00

B

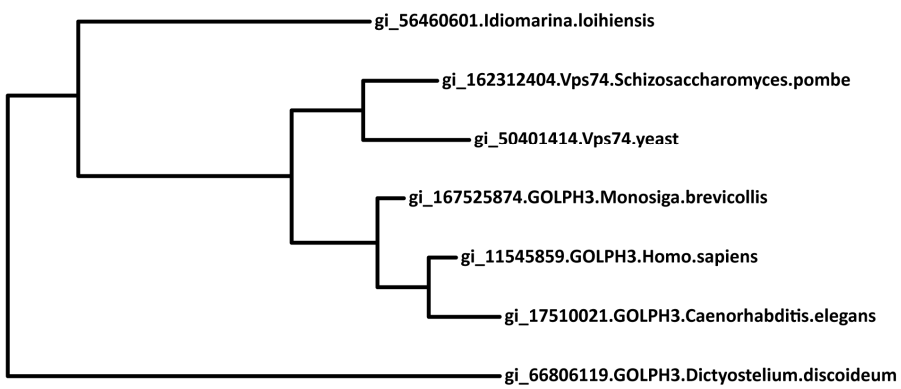

Figure 31. Sequence similarity among Vps74/GOLPH3 homologs.

(A) Matrix of sequence identity values between Vps74/GOLPH3 homologs. All organisms are eukaryotes, with the exception of *Idiomarina loihiensis,* a proteobacterium. Sequence identity has been defined as: (no. identities)/(mean seq. length) [11]. (B) Phylogram illustrating the sequence divergence and apparent evolutionary relationship between Vps74/GOLPH3 homologs. Tree arm length is proportional to sequence divergence. *I. loihiensis* Bhov is basal to both animal and fungal homologs. GOLPH3 from the slime mold *D. discoideum* is relatively divergent, and may serve a different or more specialized function in this organism. The phylogram was generated from a multiple sequence alignment using the web-based program, Phylogeny.fr [30].

core helices α1, α2, and α8 are well conserved. The α1/α2 loop and the hydrophobic β-hairpin, β3/β4, which stabilize oligomer formation, are also conserved. This suggests that Bhov proteins adopt a structurally homologous fold and are able to form β3/β4 hairpin mediated dimers.

By contrast, Bhovs lack the unstructured ≈ 60 amino acid N-terminal region of Vps74 and GOLPH3, suggesting that this region developed only in ancestral eukaryotes. Bhovs also likely lack a PI binding pocket. Two residues lining the PtdIns4*P* binding pocket in Vps74, W88 and R97, are clearly not conserved in Bhovs. The degree of conservation of the remaining residues lining this pocket (Vps74 amino acids 178 – 182) is less apparent, but we observe that these residues lie in a less well conserved region. This is not surprising, given the general absence of PI phospholipids in bacteria [96].

We have cloned and purified Bhov from the marine gammaproteobacterium, *Idiomarina loihiensis* (a kind gift from Dr. Maqsudul Alam, University of Hawaii at Manoa). Size exclusion chromatography reveals that *I. loihiensis* Bhov forms large but soluble aggregates in solution at pH 7.5 and 150 mM NaCl. Biacore studies indicate that this protein does not bind to PtdIns3*P*, PtdIns4*P*, PtdIns5*P*, or PtdIns4,5P_2 (data not shown).

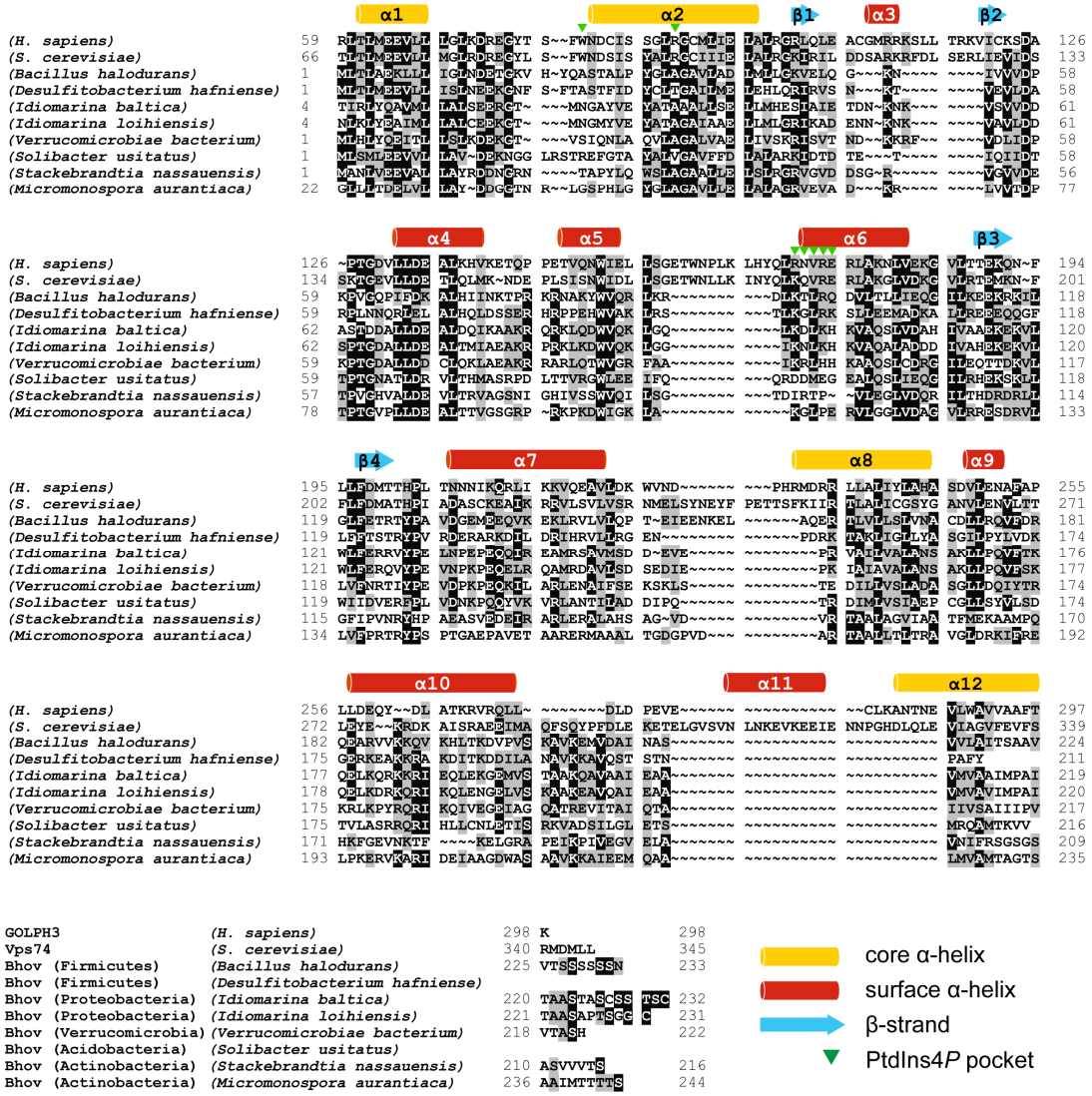

Figure 32. Multiple sequence alignment of GOLPH3, Vps74, and selected Bhov homologs.

A multiple sequence alignment shows sequence homology between animal GOLPH3, fungal Vps74, and several bacterial Bhov homologs. Bacterial phyla are shown (lower left). Vps74 secondary structure elements are indicated, as are the conserved residues lining the GOLPH3/Vps74 PtdIns4*P* binding pocket. Secondary structure is generally conserved across homologs, while the residues around the PtdIns4*P* are not conserved in bacteria. The alignment was generated by ClustalW [1, 4] and manually optimized.

2 Discussion of Vps74 and GOLPH3 data

2.1 Fungal Vps74

2.1.1 The function of Vps74 in yeast

Vps74 is required for the proper retrograde retrieval and Golgi localization of yeast mannosyltransferases. This requirement is consistent with a role as an adaptor protein, linking Golgi enzymes to retrograde trafficking machinery. To fulfill this role, Vps74 must (1) interact with cargo, (2) bind to retrograde trafficking components, (3) recognize Golgi membranes, and (4) associate with one or more of these elements in a regulated/reversible fashion. There is evidence that Vps74 possessed each of these characteristics.

Vps74 co-localizes with Golgi mannosyltransferases *in vivo*, and directly binds the cytosolic N-terminal regions of these enzymes in solution [19, 45]. Alanine scanning on the N-terminal cytosolic portion of yeast mannosyltransferases has revealed a loosely conservedconsensus sequence that is required for interaction with Vps74, and for the proper Golgi localization of these enzymes [45]. This sequence is composed of hydrophobic and basic residues, in the form of (F/L)-(L/I/V)-X-X-(R/K) [45]. Disruptions to this recognition sequence or alterations to key regions of Vps74 abolish the *in vitro* interaction between these proteins, and lead to mislocalization of these enzymes [19, 45]. Recognition of this cytosolic signal sequence by Vps74 fulfills the first of the above criteria.

We have identified surface features of the Vps74 tetramer that suggest an interaction with a protein binding partner. The tetramer forms a 70 Å wide, 30 Å deep cavity lined with well conserved amino acids (Fig. 11), which is distinct from the PtdIns4*P* binding pocket. This cavity seems larger than required to bind a ≈ 10 amino acid retention motif, and may instead interact with conserved components of retrograde trafficking complexes. We observe other regions of amino acid conservation or uniform charge, as well, all of which are candidate interaction surfaces (Fig. 11). Additionally, several reports have identified interactions between Vps74 and retrieval proteins *in vitro*. Tu et al. have shown in pulldown experiments that Vps74 interacts with COPI coatomer subunits Ret1, Ret2, Sec21, Sec26, Sec27, and Sec28 [45]. Scott et al.

report a two-hybrid assay interaction between GOLPH3 and human Vps35, a retromer complex component [62]. These interactions satisfy the second criterion above.

We and others have noted that Vps74 and GOLPH3 specifically bind PtdIns4*P* [28, 61], the major PI component of Golgi membranes [48]. We hypothesize that this interaction permits Vps74 to recognize Golgi membranes as the subcellular destination of cargo proteins (fulfilling the third of the above criteria). However, the relatively weak affinities of Vps74 and GOLPH3 for PtdIns4*P* (≈ 8.9 and 2.6 μM, respectively) are insufficient to account for the robust Golgi targeting observed in cells. Additionally, PtdIns4*P* recognition alone cannot explain the specificity of this localization, as this PI also occurs in the plasma membrane [48], to which these proteins do not localize. (The problem is exacerbated for GOLPH3, which exhibits some affinity to endosomally enriched PtdIns3*P*. [48]) Thus, Golgi localization by Vps74/GOLPH3 likely involves simultaneous detection of several Golgi markers, including PtdIns4*P*, abundant cargo proteins, and other Golgi residents. Coincidence detection has been proposed to account for the specificity of other proteins with PI recognition motifs [47], including $p47^{phox}$ [81] and oxysterol binding protein (OSBP) [93].

As an adaptor protein, Vps74 must maintain at least a small cytosolic pool in order to recognize mislocalized cargo. We hypothesize that this is accomplished by regulated, dynamic interactions between Vps74 and cellular components (which would address the final of the above criteria). This is supported by evidence that GOLPH3 dynamically associates with the Golgi *in vivo* in a nucleotide dependent fashion [69]. Structural observations also indicate opportunities for regulation. A potential phosphorylation site exists in the α1/α2 loop [19], which stabilizes tetramer formation. Also, the unstructured N-terminal region is not required for Golgi localization or tetramer formation [19], and may well serve a regulatory role. We observe that Vps74 is monomeric under most solution conditions, suggesting that tetramer formation *in vivo* must be stabilized by a protein partner. Furthermore, it seems unlikely that tetrameric Vps74 binds PtdIns4*P*, as the binding pocket is occluded in the Vps74 tetramer, and the hydrophobic β-hairpin may be required to stabilize lipid binding via membrane insertion.

Thus, there is strong evidence that Vps74 functions as an adaptor molecule in fungi, linking a subset of resident Golgi proteins to COPI and retromer retrograde trafficking machinery. It is curious that pulldown interactions have been observed with COPI [45], as the coatomer is typically characterized as mediating Golgi-to-ER retrieval [40]. Vps74 may serve to extend the functionality of COPI vesicles, directing trafficking of proteins to the Golgi via a retention motif distinct from typical COPI sorting signals [97]. This hypothesis is supported by the observation that Sac1 disruption results in accumulation of PtdIns4*P* in the ER, along with mislocalization of Vps74 and mannosyltransferases to this organelle (Fig. 26).

Our work with Vps74 has led to several key questions. Does GOLPH3 recognize specific retrograde cargo in animals? (The family of fungal α1,2-mannosyltransferases associated with Vps74 is absent in animals.) Under what conditions does Vps74/GOLPH3 interact with retrograde trafficking components, and what are the structural details of these interactions? Finally, and reflecting the greatest gap in our current understanding: how are these proteins regulated in cells? Further investigation is clearly required to address these questions.

2.1.2 Tetramer formation by Vps74

We show that Vps74 forms tetramer, and possibly higher order oligomers, in the crystal structure and in solution. Oligomer formation in solution is driven by millimolar mixtures of various divalent metal ions and either EGTA or EDTA. The molecular basis for this phenomenon is not clear. There is no crystallographic evidence that metal ions or chelator are bound in the Vps74 complex, and oligomer formation in solution occurs with multiple divalent metal ions and with the structurally similar chelators, EGTA and EDTA. Thus, it is unlikely that tetramer stabilization is due to the binding of a specific metal ion, or to the chelation of an ion purified along with the protein. Rather, we hypothesize that the entire chelated metal complex weakly interacts with Vps74 such that it stabilizes tetramer and alters monomer/tetramer equilibrium. Given an estimated K_D of oligomerization of ≈ 1 mM in the absence of Ca/EGTA and ≈ 10 μM in its presence (see Section 1.2.7), Ca/EGTA stabilizes oligomer formation by only ≈ 2.7 kcal/mol,

or the equivalent of about two hydrogen bonds [98, 99]. This is consistent with a relatively weak association of Ca/EGTA complex and Vps74 tetramer. Even if the effect on tetramer formation is a serendipitous artifact rather than a specific interaction with a physiologically relevant cofactor, it has allowed us to establish that tetramer formation is correlated with function.

Our work suggests a relationship between formation of the Vps74/GOLPH3 tetramer and binding to membranes, cargo, and other cellular components, although the nature of this relationship remains unclear. We show that the presence of Ca/EGTA enhances the pulldown of purified Vps74 by the cytosolic region of Kre2 (Fig. 16). This may indicate that the tetrameric species is specifically interacts with Kre2; alternatively, this effect may be a consequence of avidity. The presence of a conserved cavity in the tetramer structure suggests that tetramer formation is required to interact with protein partners. However, occlusion of PtdIns4*P* binding sites in the tetramer suggests that this species cannot bind membranes.

Additionally, we note that a conformational change likely accompanies the transition between monomeric to oligomeric states. The subunit conformation in the Vps74 and GOLPH3 crystal structures involves the projection of the hydrophobic β-hairpin away from the bulk of the helical domain. To exist as a soluble monomer, we predict that the β-hairpin packs against the protein surface, minimizing exposed hydrophobic surface area. The α1/α2 loop at the base of the β-hairpin may also change conformation to accommodate the hydrophobic side chains.

Although our current understanding of Vps74 function is limited, the data presented here allows us to propose an initial model for the function of this protein in cells. We suggest that, in the absence of interaction partners, Vps74 exists as a monomer in the cytosol and adopts a conformation (not yet observed crystallographically) in which the β3/β4 hairpin packs against the helical core. Interactions with retrograde trafficking elements in downstream trafficking compartments promotes tetramer (or higher order oligomer) formation, which in turn strengthens interactions with the cytosolic components of mislocalized cargo. Vps74 serves as an adaptor protein during transit. Vps74 recognizes the destination compartment by virtue of PtdIns4*P* levels. Tetramer subsequently dissociates in favor of monomeric association with PtdIns4*P*

containing membranes. We anticipate that post-translational modifications will be identified that regulate various aspects of Vps74 activity. Future work will be required to test these hypotheses and better understand the molecular details of Vps74 function in fungi.

2.2 Animal GOLPH3

2.2.1 The function of GOLPH3 in animals

Several lines of evidence indicate that GOLPH3 has similar functionality to Vps74. GOLPH3 and Vps74 share high sequence and structural similarity, including regions required for Vps74 function (the β-hairpin and the PtdIns4*P* binding pocket, Fig. 9). Like Vps74, GOLPH3 is Golgi localized and specifically interacts with PtdIns4*P*. GOLPH3 interacts with the cytosolic regions of mannosyltransferases in solution and can functionally replace Vps74 in yeast [19, 45]. Knockdown of GOLPH3 in animal cells results in mislocalization of a viral transmembrane protein [61]. However, homologs to the fungal α-1,2 mannosyltransferase enzymes trafficked by Vps74 in yeast do not occur in animals. While Golgi proteins with similar topology exist in animals, no endogenous cargo proteins have yet been reported for GOLPH3.

It has also been suggested that GOLPH3 and Vps74 have distinct functions [61]. The ability of GOLPH3 to substitute for Vps74 may occur as a consequence of structural conservation, rather than functional homology. Additionally, GOLPH3 plays an important role in linking Golgi membranes to the actin cytoskeleton via a direct interaction with Myo18A [61], such that knockdown of GOLPH3 alters Golgi topology and affects vesicle budding. Vps74 is not known to perform a similar function in fungi, which lack a Myo18A homolog.

2.2.2 The role of GOLPH3 in human disease

Defects in trafficking and glycosylation are associated with several human diseases [44]. However, no diseases are known to be caused by deletion or mutation of GOLPH3. It is possible that, as is the case with Vps74 in yeast, GOLPH3 function is dispensable for normal growth. Alternatively, severe disruptions in GOLPH3 may be lethal during embryonic

development. Global or conditional knockout studies may address these possibilities and shed more light on GOLPH3 function in animals.

Interestingly, *GOLPH3* has been shown to be a *bona fide* oncogene [62]. Focal amplification of the *GOLPH3* genomic locus occurs in 56% of lung, and > 30% of ovarian, breast, prostate, and melanoma tumors. GOLPH3 overexpression results in increased cell size and activates mTor signaling; conversely, knockdown of endogenous GOLPH3 in tumor cell lines decreases cell growth [62]. The specific mechanism linking GOLPH3 activity to mTor is unknown, though several potential pathways linking GOLPH3 to cancer have been noted [100]. GOLPH3 interacts with components of the retromer complex [62], disruption of which is known to alter Tor signaling in yeast [101]. Additionally, alteration of retromer activity may affect receptor internalization [62]. The oncogenic properties of GOLPH3 may also arise from its significant (though weaker) affinity for PtdIns3*P* (Fig. 27). Overexpressed GOLPH3 may partially localize to membranes enriched in PtdIns3*P*, incidentally mislocalizing cargo and altering cell signaling.

It may be insightful to investigate the effects on cell growth of GOLPH3 variants lacking the ability to bind lipid or to form tetramer. Along similar lines, it may be possible to increase GOLPH3 specificity for PtdIns4*P* by altering amino acids near the binding pocket, or substituting for homologous amino acids from Vps74. If the oncogenic effects of GOLPH3 require binding to PtdIns3*P*, overexpression of a higher selectivity GOLPH3 variant should not activate mTor signaling.

2.3 The function of Bhov in prokaryotes

No Bhov has been studied in detail, so the role of these proteins is still speculative. Structural features involved in tetramer formation, including the β-hairpin, are conserved in bacteria. Thus, Bhovs are likely capable of oligomer formation and possibly membrane association, given the conservation of the hydrophobic β-hairpin. It seems probable that these bacterial proteins serve a completely separate function among prokaryotes than in eukaryotes,

and that the PtdIns4*P* binding and protein sorting functions of Vps74 and GOLPH3 arose only eukaryotic organisms.

3 Inhibition of the extracellular domain of Epidermal Growth Factor Receptor by antibody-based inhibitors

3.1 Introduction

3.1.1 ErbB receptor tyrosine kinases

The epidermal growth factor receptor (EGFR, human EGF receptor 1 [HER1], ErbB receptor 1 [ErbB1]) is a plasma membrane receptor tyrosine kinase (RTK) that regulates cell growth and proliferation [102]. EGFR is a well characterized oncoprotein in epithelial derived cancers, and is an important target of cancer therapeutics [103]. Along with ErbB2/HER2/Neu, ErbB3/HER3, and ErbB4/HER4, EGFR is one of four members in humans of the ErbB RTK family (named for homology to the avian erythroblastosis virus transforming protein, v-*erb*B [104]). ErbB receptors are type I transmembrane proteins, composed of an amine-terminal extracellular ligand binding region, a single transmembrane helix, and an intracellular region that includes a juxtamembrane segment, a tyrosine kinase domain, and a carboxy-terminal regulatory segment [105].

Activation of ErbB receptors results in auto-phosphorylation of tyrosine side chains in the C-terminal region by the cytosolic TK domain. The phosphorylated regulatory region serves as a scaffold for the recruitment of a variety of adaptors and signaling enzymes (including Shc, PLCγ, PI3K) that contain Src homology 2 (SH2) or phosphotyrosine binding (Ptb) domains, resulting in the activation of downstream pathways (MAPK, PKC, Akt, among others) [106, 107]. ErbB signaling ultimately promotes cell growth, migration, and differentiation, and plays an important role in the development of the skin, lungs, pancreas, kidneys, nervous system, and other tissue types [102, 108-110].

Given their proliferative effects on cell growth, it is perhaps not surprising that aberrant activation of ErbB receptors is a frequent component of human cancers. EGFR involvement (through overexpression or mutation) occurs in a significant fraction of cancers of the breast, bladder, prostate, kidney, head and neck, as well as gliomas and non small cell lung carcinoma (NSCLC) [106, 111, 112]. ErbB2 overexpression occurs in lung, pancreas, and colon cancer,

and is particularly prevalent in breast cancer (≥ 30%) [106, 113]. ErbB receptor involvement is correlated with tumor malignancy and poor prognosis [114].

3.1.2 Activation of ErbB receptors

ErbB ligands include epidermal growth factor (EGF), transforming growth factor α (TGFα), and amphiregulin, which are recognized by EGFR; betacellulin, heparin-binding EGF (HB-EGF), and epiregulin, which bind to EGFR and ErbB4; and the neuregulins (NRG1 – 4), which bind ErbB3 and/or ErbB4 [106, 110, 115]. No soluble ligand is known for ErbB2. Activated ErbB receptors form a range of homo- and heterodimers, which stimulate distinct repertoires of downstream signaling pathways [106].

ErbB receptors follow the general paradigm of activation observed in most characterized RTK members, with some key distinctions. Ligand binding leads to receptor dimerization, thereby driving association and trans-activation of the intracellular tyrosine kinase domains. For several RTKs (vascular endothelial growth factor receptor [VEGFR] and fibroblast growth factor receptor [FGFR]) and non-RTK receptors (growth hormone receptor [GHR] and erythropoietin receptor [EPOR]) dimerization is stabilized by bivalent binding to ligand (or to two ligands, in the case of FGFR binding to FGF and heparin sulfate) [116-119]. By contrast, crystallographic and biochemical studies have shown that the ligand binding and dimerization surfaces of ErbB receptors are distinct [1, 2].

The extracellular region of human ErbB receptors comprises four domains (Fig. 33). Domains I (amino acids 1 – 165) and III (310 – 481) share ≈ 37% sequence similarity [17] and adopt a rigid β-helix fold, similar to that of leucine rich repeat domains [120]. Cysteine rich domains II (166 – 309) and IV (482 – 618), which share ≈ 17% sequence identity, adopt an elongated and somewhat flexible structure of successive disulfide-bonded modules [17]. In the absence of ligand, EGFR has been shown by X-ray crystallography and small angle X-ray scattering (SAXS) studies to adopt a “tethered” or “autoinhibited” conformation [14, 17] (Fig. 33). In the tethered state, domains II and IV interact (burying 922 $Å^2$ total surface area) largely through contacts between domain IV and a loop that projects from domain II (residues 240 –

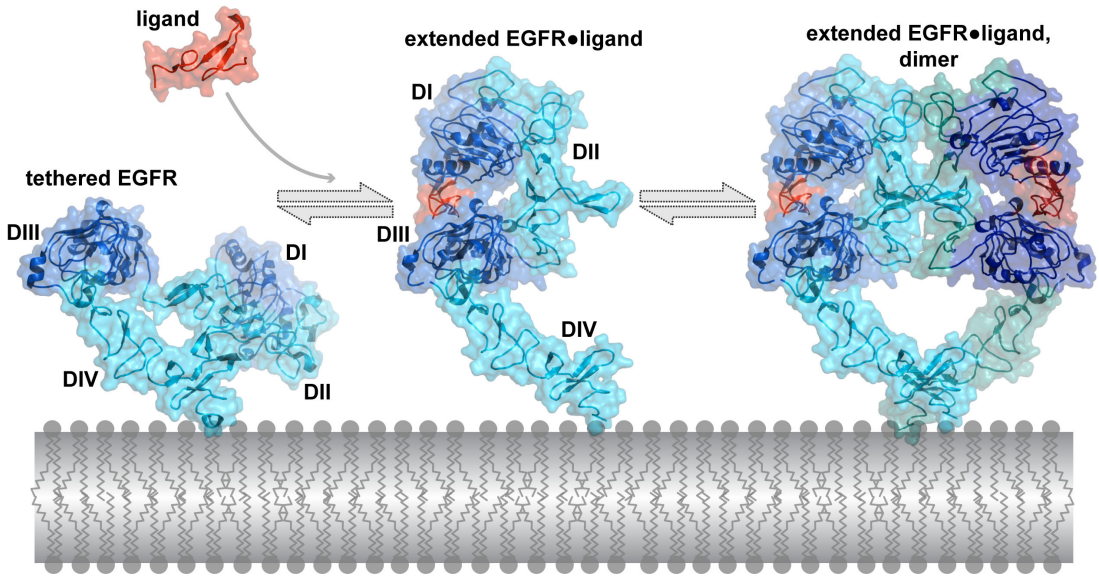

Figure 33. EGFR structure and activation.

Structural models of the EGFR extracellular region are shown above a representation of a membrane bilayer. Structural studies of the EGFR extracellular region reveal two distinct states, which are thought to exist in dynamic equilibrium. A tethered conformation predominates in the absence of ligand. Contacts between domains II and IV partially stabilize this conformation, burying dimerization surfaces on domain II, and separating domains I and III. Bound ligand stabilizes an extended conformation, in which tether interactions are absent, dimerization surfaces on domain II are exposed, and ligand interacts with surfaces on domains I and III. The extended conformation of EGFR forms dimers, stabilized largely by contacts on domain II. Ligand binding and dimerization are cooperative; intermediate states are predicted but have not been experimentally observed. The tethered conformation was generated from PDB 1NQL. The extended conformation and EGFR homodimer were generated from 1IVO, with domain IV modeled from the analogous region of 1YY9 chain A. EGFR transmembrane and intracellular regions are not shown in this representation.

260) [17]. In this inactive state, the ligand binding surfaces on domains I and III are spatially separated by ≈ 50 Å.

Binding of ligand is associated with a large conformational change in the extracellular region (Figs. 33 & 34). Domains I and II rotate with respect to domains III and IV, bringing about an "extended" conformation in which domains I and III bind ligand bivalently (occluding ≈ 1500 $Å^2$ total surface area) [1, 2]. In this conformation, the autoinhibitory interaction from the tethered state is abolished, instead exposing surface area on domains II and IV. Surfaces on domain II, and to a lesser extent on domain IV, symmetrically interact (≈ 1200 $Å^2$ total) to stabilize dimerization of "extended" EGFR [1, 2]. Interestingly, the differences between the unliganded/tethered and liganded/extended conformations of EGFR involve more than rigid body motions. Contacts between domains II and III in the extended state induce conformational changes in domain II, which repositions domain II residues to facilitate dimer-stabilizing interactions [21, 105].

Structural data on ErbB3 suggests that the extracellular regions of this receptor undergoes a similar activation scheme to EGFR, and can adopt tethered and extended conformations [14]. Similar behavior is expected for ErbB4. However, ErbB2 (for which no soluble ligand is known) exists only in the extended conformation. Thus activation of ErbB2 is thought to occur exclusively through heterodimerization with the ligand activated forms of other ErbB receptors.

3.1.3 Evidence of additional complexity of EGFR activation

The crystal structures of the EGFR extracellular region reveal two distinct states: extended [1, 2] and tethered [17] (Fig. 33). These structures suggest a simple model of activation in which unliganded receptor exists in the tethered state, bivalent binding of ligand stabilizes the extended state, and symmetric dimerization of extended conformation receptors promotes association of the TK domains (Figs. 33 & 34A) [121]. However, there is evidence that this model cannot account for the full complexity of receptor activation in cells.

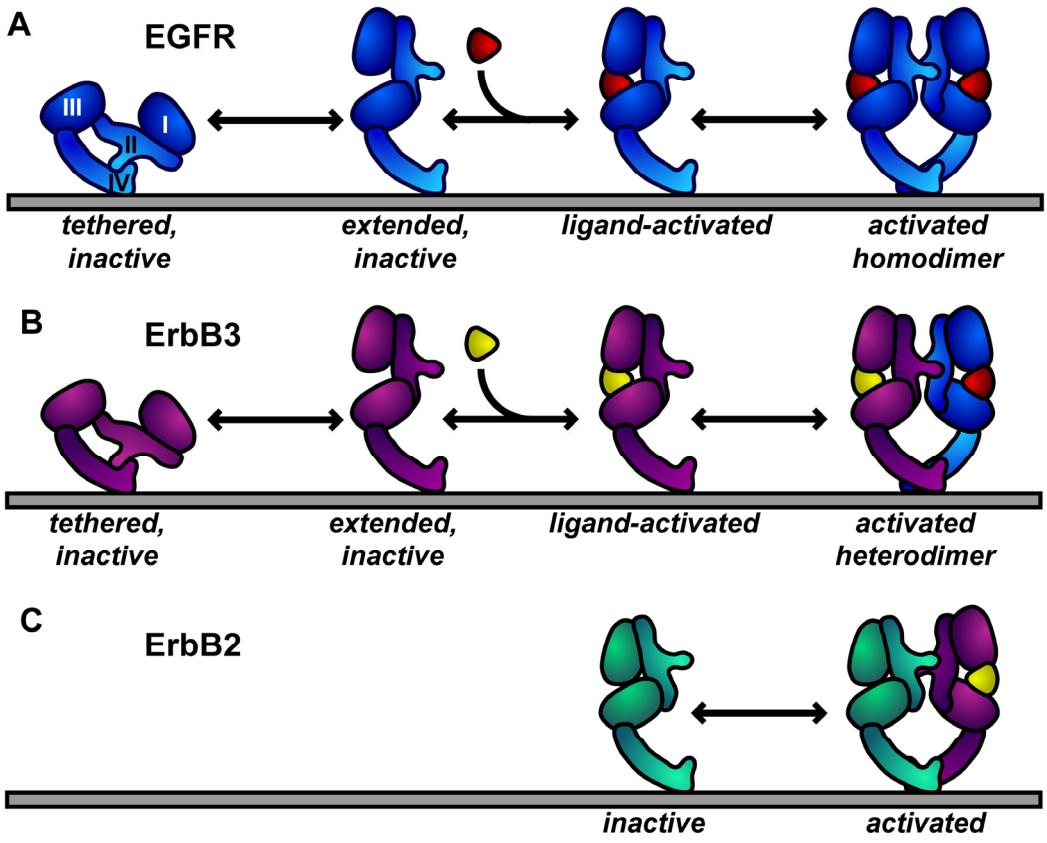

Figure 34. Ligand induced dimerization of ErbB receptor extracellular regions.

(A) Ligand binding stabilizes a conformational change in the EGFR extracellular region, favoring homo- and ErbB heterodimer formation. Tethered monomer [17] and homodimers of ligand-bound EGFR in an extended state [1, 2] have been observed crystallographically; intermediate states are inferred. (B) Tethered structures of ErbB3 and ErbB4 have been observed [22, 23], and are similar to that of EGFR. ErbB4 is thought to form homodimers (as in A), while ErbB3 does not. Both ErbB3 and ErbB4 form heterodimers with other ErbB receptors. (C) ErbB2, for which no ligand is known, adopts an extended conformation [24] and forms heterodimers with ligand-activated ErbB receptors.

It has been known since the 1980s that ligand binding to cell surface EGFR expressed at the cell surface results in Scatchard plots that are curved and concave-up [102] (Fig. 35A). This phenomenon has been interpreted as representing two receptor populations (although other interpretations are possible): a small population of "high affinity" sites for ligand (≈ 5%, 10-100 pM K_D [25, 122]) and a larger population of "low affinity" sites (≈ 95%, 1-10 nM K_D [25, 122]). This behavior at the cell surface requires the presence of the intracellular region [122, 123], and does not occur in purified preparations of the receptor extracellular region [124]. Multiple hypotheses have been proposed to explain these observations, including differential binding to extended and inactivated receptor populations, heterogeneous receptor density at the cell surface [125], the involvement of an unknown interaction partner [126], and negative cooperativity [127]. No definitive molecular evidence has been reported to explain the phenomenon, but recent modeling and ligand competition studies support negative cooperativity (Fig. 35) [18, 123]. Whatever the mechanism underlying this feature of ligand binding, it may have implications on receptor activation, signaling, and inhibition.

3.1.4 Inhibition of ErbB receptor activation

Given the correlation between aberrant activation of ErbB receptors and the progression of human cancers, ErbB receptors are an obvious target for anti-cancer therapeutics [115]. Inhibitors interfere with receptor activation by two distinct approaches. One group, the ErbB tyrosine kinase inhibitors (TKIs), targets the enzymatic activity of the intracellular kinase domain. These TKIs typically are ATP-mimetic small molecules that diffuse across the plasma membrane and compete for (or covalently block) ATP binding to the ErbB TK. FDA approved ErbB TKIs include Gefitinib (AstraZeneca), Erlotinib (Genentech), and Lapatinib (GlaxoSmithKline) [103, 110].

The second group of ErbB inhibitors abrogates receptor activation via the extracellular region. These function by blocking ligand binding, sterically preventing conformational changes associated with receptor activation, preventing receptor dimerization, or a combination of these

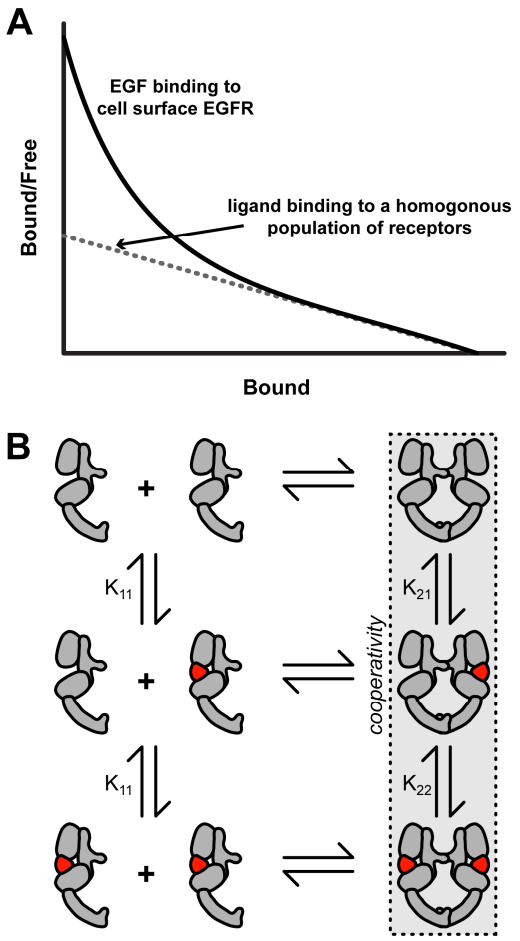

Figure 35. Receptor species predicted by negative cooperativity.

(A) Ligand binding to a single class of receptor binding sites results in a linear Scatchard plot (dashed gray line). However, EGF binding to EGFR at the cell surface produces a curved, concave-up Scatchard plot. (B) Receptor states in a system that exhibits cooperativity (adapted from [18]; compare to Fig. 34A). Receptor extracellular regions are shown in gray; ligand in red. Cooperativity influences ligand binding in the boxed region. Negative cooperativity occurs when binding of a first ligand to an unliganded dimer (K_{21}) occurs with higher affinity than binding of a second ligand (K_{22}). [25].

mechanisms [128] (Fig. 36). Small molecules that disrupt the interactions involved in ligand binding and dimerization have not been identified, possibly due to the large surface areas involved. Monoclonal antibodies have been isolated that inhibit activation by binding to the extracellular region of the receptors. Antibodies recognizing EGFR and ErbB2 have been studied since the 1980s. A number of anti-ErbB antibodies are now known, which exhibit distinct modes of binding and inhibition (Table 6). These antibodies have proven useful as biochemical tools in the study of receptor function [129, 130], as histological agents, and as cancer therapeutics (Table 6)

At the molecular level, both mAbs and Fab fragments are typically capable of inhibiting receptor activation. However, intact antibodies may additionally elicit a variety of immune responses are mediated by the Fc constant region. Fc receptors (FcRs) occur on the surface of natural killer cells, phagocytes, and other immune cells that may destroy antibody targeted tumor cells by phagocytosis or antibody-dependent cellular cytotoxicity (ADCC) [131]. Fc regions are also recognized by subunits of the complement system, which lyses antibody bound cells by complement dependent cytotoxicity (CDC) [132]. These immune responses are an important component of the overall anti-tumor activity of antibody therapeutics, although both types of response are recruited more robustly by antibodies of subtype IgG1 than IgG2 (Table 6) [131, 132]. The relative importance of ErbB receptor inhibition and of immune response recruitment, with respect to anti-tumor efficacy, has not been clearly evaluated.

The following sections briefly describe notable and clinically relevant antibodies directed against EGFR and ErbB2, beginning with those for which the interaction with antigen has been structurally characterized.

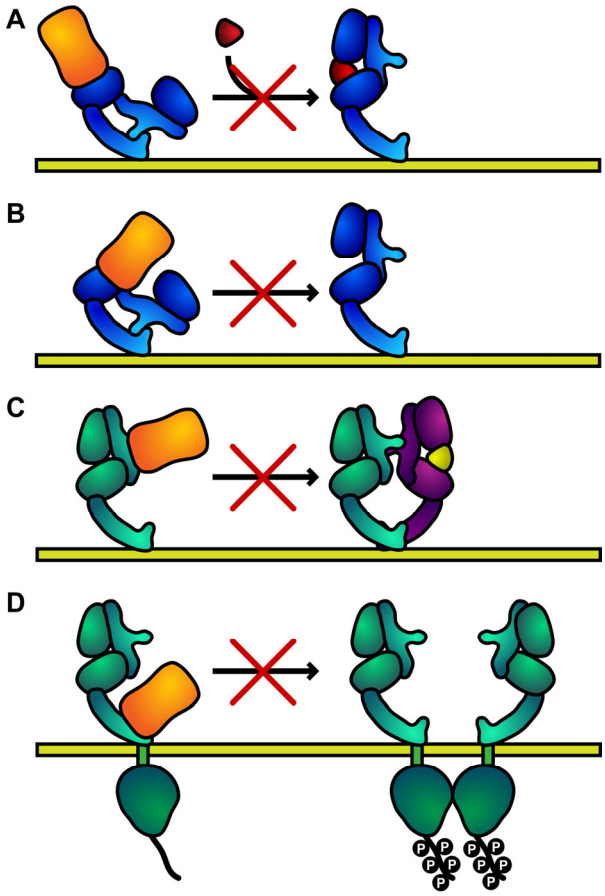

Figure 36. Inhibition of ErbB activation by antibody binding.

Characterization of inhibitory ErbB antibodies has revealed several distinct inhibitory routes. (A) Cetuximab, IMC-11F8, panitumumab, and zalutumumab directly occlude the ligand binding site on EGFR domain III. (B) Matuzumab sterically obstructs adoption of the extended conformation of EGFR (and thus high affinity ligand binding, as well as dimerization) without directly competing with ligand binding. (C) Pertuzumab disrupts receptor dimerization of ErbB2. (D) Trastuzumab recognizes an extracellular juxtamembrane epitope on ErbB2 and may inhibit activation by preventing association of ErbB2 TK domains. Additional mechanisms have also been proposed for the antiproliferative effects of this antibody.

Table 6. Properties of selected antibodies to the extracellular region of ErbB receptors

Underlined antibodies are characterized in Section **Error! Reference source not found.**.

mAb	IgG type	Epitope	Drug Name	Comments	Other Refs
Antibodies directed against EGFR					
Antibodies derived from mice immunized with A431 cell (or similar) EGFR (intact cells or membrane preparations)					
225/ C225	IgG1	Domain III [5]	Cetuximab/Erbitux	Approval for use against colorectal and head and neck cancers; ongoing trials for other indications	[133-135]
425/ EMD72000	IgG2a	Domain III [6]	Matuzumab	Phase II trials; development halted	[136, 137]
R1		Domains I/II [129]	-	Does not block ligand stimulated activation	[138]
2E9	IgG1	Domain I [129]	-	Blocks low affinity cell surface EGF binding sites	[139]
108	IgG2a	Domain III [129]	-	Blocks high affinity cell surface EGF binding sites	[25]
13A9	IgG1	Domain III[29]	-	Inhibits TGFα binding and activation of EGFR; does not alter EGF binding or signaling	[140]
29.1		Carbohydrate	-	Used in purification of active EGFR	[141]
Antibody derived from mice immunizedwith purified placental EGFR					
R3/h-R3	IgG2a	Domain III [142, 143]	Nimotuzumab	Phase III trials	[144]
Fully human monoclonal antibodies derived from transgenic mice immunized with A431 preparations					
E7.6.3/ ABX-EGF	IgG2	Domain III [9]	Panitumumab/Vectibix	Approved for use against colorectal cancer; in trials for additional cancer types	[145]
2F8/ HuMax-EGFr	IgG1	Domain III [10]	Zalutumumab	FDA fast track status for head and neck cancer; Phase III trials	[146]
Fully human monoclonal antibody derived from library screening					
IMC-11F8	IgG1	Domain III [147]	Necitumumab	Phase II trials	[148]
Antibody derived from mice immunized with fibroblasts expressing EGFRvIII					
806	IgG2b	Domain II [3]	ch806	Recognizes EGFRvIII and binds to a fraction of cells overexpressing wild-type EGFR. Completed Phase I trials.	[149, 150]
Antibodies directed against ErbB2/HER2					
Monoclonal antibody derived from mice immunized with NIH 3T3 cells that had been transformed with human ErbB2					
4D5	IgG1	Domain IV [24]	Trastuzumab/Herceptin	Approved for treatment of ErbB2-expressing breast cancer	[151]
2C4	IgG1	Domain II [31]	Pertuzumab/Omnitarg	Phase II trials	[152]

3.1.5 Inhibitory EGFR antibodies C225 and IMC-11F8

Of the antibody inhibitors to EGFR, the most extensively studied is the murine antibody 225 and its derivatives. mAb 225 was generated by immunizing mice with A431 cells (an epidermoid carcinoma cell line expressing high levels of EGFR [153]) and screening the resulting antibodies for inhibition of both EGF binding and receptor phosphorylation [133, 154]. A human/murine chimera of 225, IMC-C225 (Cetuximab/Erbitux), was developed as a therapeutic by Imclone, Inc. and has FDA approval for the treatment of colorectal cancer and cancers of the head and neck [134, 135].

mAb 225 competes with ligand for binding to EGFR at the cell surface [133] and in solution [5]. We have structurally characterized the interaction between FabC225 and the extracellular region of EGFR, revealing the basis of this competition [5]. FabC225 binds to an epitope on domain III that substantially overlaps with the ligand binding surface (Figs. 37A & 38C). Mutations to key residues on the surface of domain III affect binding of FabC225, binding of ligand, or both, confirming the existence of overlapping binding sites. Additionally, bound FabC225 sterically prevents adoption of the extended conformation, precluding the possibility of ligand independent dimerization and activation.

The fully human monoclonal antibody, IMC-11F8, inhibits similarly to mAb C225. IMC-11F8 was identified from a phage display library, based on its ability to bind A431 cells and to compete with mAb C225 for binding to EGFR [150, 155]. Not surprisingly, the mAb C225 and IMC-11F8 epitopes are nearly identical, and IMC-11F8 shows similar inhibitory characteristics [147]. This antibody is in Phase II trials for the treatment of colorectal cancer.

Together mAb C225 and IMC-11F8 represent a single mode of EGFR inhibition, involving direct competition of ligand binding along with steric blockade of the extended receptor conformation (Fig. 36A).

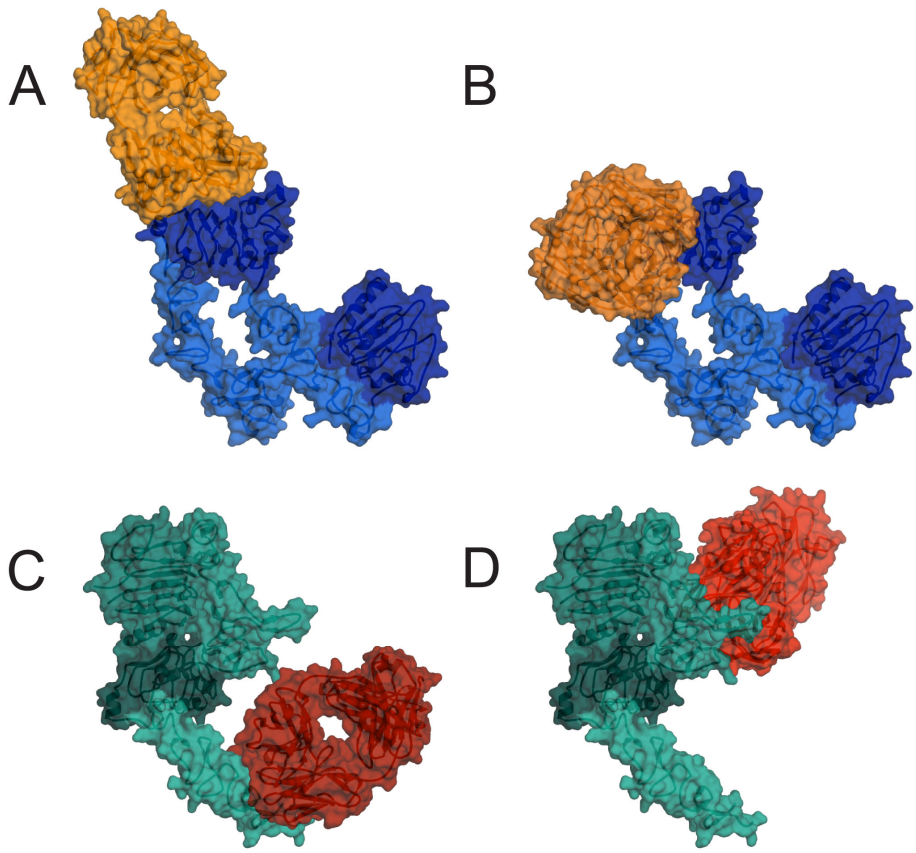

Figure 37. Crystals structures of Fab inhibitors in complex with EGFR and ErbB2. Molecular surface representations of (A) FabC225/cetuximab (light orange) bound to the EGFR extracellular region (sEGFR, blue), (B) FabEMD72000/matuzumab (orange) bound to a model of sEGFR (blue), (C) Fab4D5/trastuzumab (dark red) bound to ErbB2 (teal), and (D) Fab2C4/pertuzumab (red) bound to ErbB2 extracellular region (teal). The structurally characterized inhibitors of EGFR recognize domain III, while ErbB2 inhibitors bind domains II and IV. Figures were generated from PDB coordinates 1YY9 [1, 5] (A), 3C09 [6] (B), 1N8Z [24] (C), and 1S78 [31] (D). Domains I, II, and IV in (B) are modeled on 1YY9.

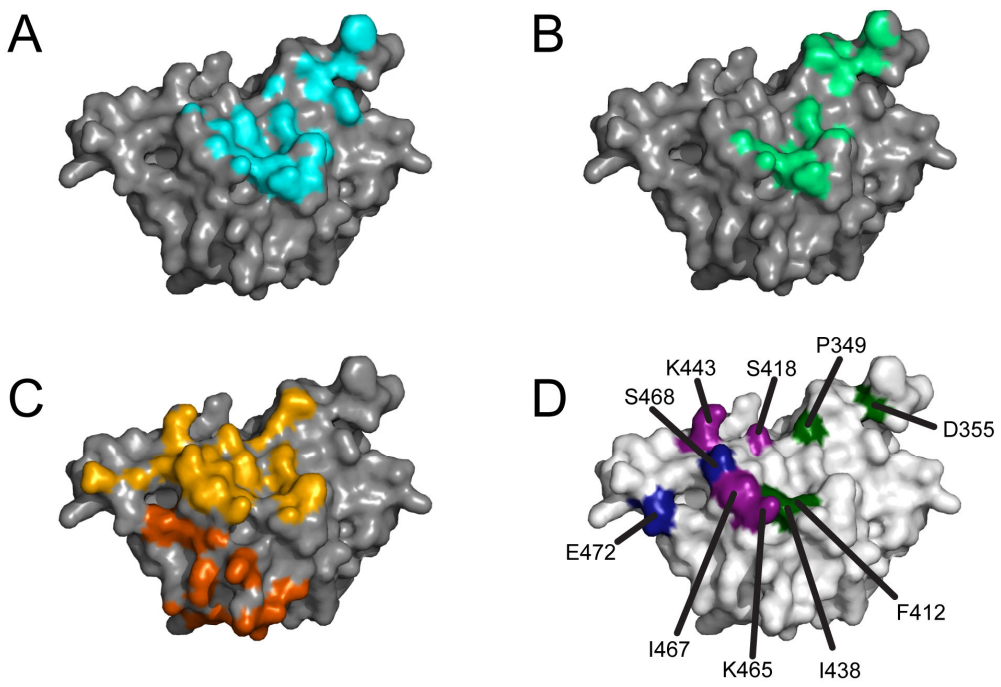

Figure 38. Comparison of epitopes and ligand binding sites on EGFR domain III.

Molecular surface representations of EGFR domain III. (A) EGF and (B) TGFα binding footprints are highlighted [1, 2]. (C) The FabC225 epitope (gold) significantly overlaps with the ligand binding sites, while the Fab72000 epitope (red) does not. (D) Mutation of the indicated amino acids affects binding of inhibitory antibodies panitumumab (green) [9], zalutumumab (magenta) [10, 14], and mAb 13A9 (blue) [29]. The domain III representation was generated from structure 1YY9 [5]. (A-C) Relevant complex structures (EGF:1IVO, TGFα:1MOX, FabC225:1YY9, Fab72000:3C09) were aligned to domain III. Atoms on domain III within 4 Å of bound ligand or Fab are highlighted.

3.1.6 Inhibitory EGFR antibody mAb 425/EMD72000

A separate mechanism of inhibition is achieved by the inhibitory murine antibody 425, which has been humanized for potential therapeutic applications as EMD72000/matuzumab (Merck KGaA). mAb 425 was generated in mice immunized with A431 cells, and was identified by screening the resulting antibodies for inhibition of EGF binding [136]. It inhibits receptor activation and tumor growth, although EMD72000 exhibited poor performance in Phase II clinical trials [137].

The structure of the Fab fragment of EMD72000 (Fab72000) has been determined in complex with the EGFR extracellular region (Fig. 37B) [6]. Fab72000, like FabC225, binds to domain III, but recognizes an epitope distinct from the ligand binding surface (Fig. 38C) [6]. Mutations to residues in domain III that disrupt EMD72000 binding to EGFR do not alter ligand binding. However, bound EMD72000 sterically blocks adoption of the active receptor conformation, by preventing domain I from adopting its extended position. Thus, EMD72000 prevents bivalent binding to ligand and dimerization induced receptor activation (Fig. 36B). The distinction between this mechanism and the competitive inhibition of mAb C225 may be therapeutically relevant. While the binding and inhibitory activity of mAb C225 (and similar antibodies) may be out-competed by high concentrations of ligand, EMD72000 inhibition is immune to such competition. This grants EMD72000, and similarly binding inhibitors, particular therapeutic utility for the treatment of tumors that involve overexpression of EGFR ligand [6].

3.1.7 Inhibitory antibodies of ErbB2

Antibodies that recognize and inhibit ErbB2 activation have been characterized, as well. As ErbB2 has no known soluble ligand (thus far) and adopts only an extended conformation [24, 31, 156], these inhibitors utilize mechanisms distinct from inhibitory EGFR antibodies.

Murine monoclonal antibody 2C4 was generated by immunizing mice with NIH 3T3 cells transfected with human ErbB2 [157]. mAb 2C4 blocks ligand dependent ErbB2/ErB3 heterodimer activation and cell growth in ErbB2 expressing tumor cell lines and xenograft

studies [158, 159]. mAb 2C4 has been humanized as a therapeutic (pertuzumab/Omnitarg) [160], and is in Phase II trials for treatment of breast, ovarian, prostate, and other cancers [152].

Structural characterization of the pertuzumab Fab/ErbB2 complex shows pertuzumab to bind domain II of ErbB2 (Figs. 37D, 39) [31]. ErbB2 heterodimerizes with other ErbB receptors [105, 115] and this interaction is thought to involve the arm region of domain II, as in ligand stabilized homodimerization of EGFR [1, 2]. The pertuzumab epitope lies largely on the dimerization arm (Fig. 39), which sterically prevents ErbB2 from engaging in a heterodimerization interaction (Fig. 36C) [31]. This structural observation explains the ability of pertuzumab to block heterodimer formation on the cell surface [159].

A second murine monoclonal antibody, 4D5, was generated by the same method as mAb 2C4 [157]. mAb 4D5 specifically recognizes ErbB2 [157, 161], decreases ErbB2 phosphorylation in breast cancer cell lines [162, 163], and decreases tumor growth of ErbB2 expressing tumors in xenograft mice [164]. mAb 4D5 was subsequently humanized (trastuzumab/Herceptin) for therapeutic applications [165]. Trastuzumab has FDA approval for the treatment of breast cancers [151].

The structure of the trastuzumab Fab bound to the ErbB2 extracellular domain has been determined [24]. Trastuzumab recognizes an epitope on the extracellular juxtamembrane region of the receptor, domain IV (Figs. 37C & 39). This epitope does not include regions thought to significantly stabilize receptor dimerization, and trastuzumab binding inhibits only some forms of ErbB2 heterodimerization (blocking ErbB2/EGFR, but not ErbB2/ErbB3 heterodimer formation) [166]. Since trastuzumab inhibits proliferation in cells in which ErbB2 is overexpressed and constitutively active, it may be that this antibody generally (independent of dimerization) inhibits close association and trans-activation of the TK domain (Fig. 36D). Additionally, it has been suggested that the antiproliferative effects of trastuzumab arise largely from antibody cytotoxicity, promotion of receptor internalization, and inhibition of ectodomain shedding [24, 167], rather than inhibition of dimerization.

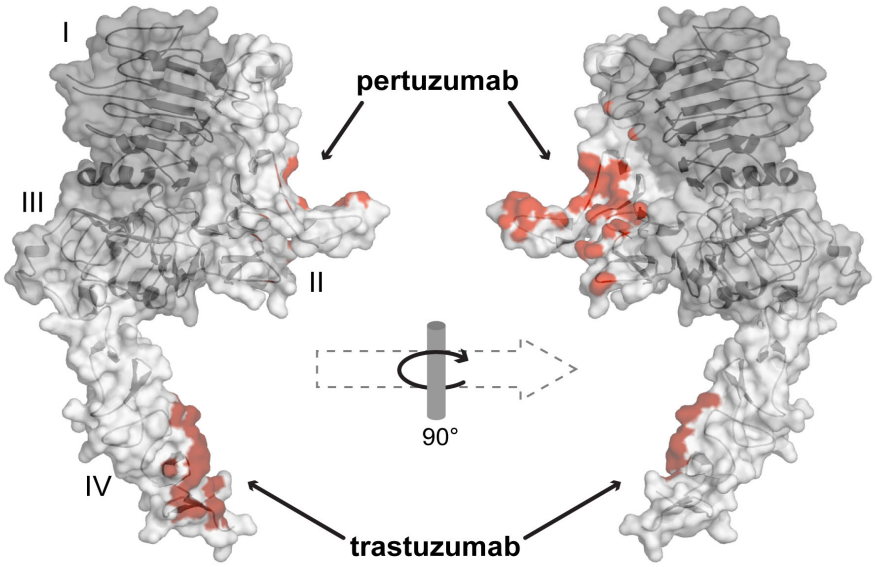

Figure 39. Epitopes of ErbB2 antibodies.

The epitopes of anti-ErbB2 antibodies pertuzumab (red) and trastuzumab (dark red) are shown on a molecular surface representation of the ErbB2 extracellular region. The pertuzumab epitope on domain II includes the dimerization arm and other amino acids that, in EGFR, mediate dimerization. The domain IV surface recognized by trastuzumab does not contribute strongly to dimer stabilization, by analogy with EGFR [21]. However, it is thought trastuzumab binding sterically affects intracellular self- or hetero-association [18, 24]. Atoms in ErbB2 within 4 Å of bound antibody are highlighted, based on PDB structures 1S78 (pertuzumab) and 1N8Z (trastuzumab).

3.1.8 Inhibitory EGFR antibody 806

Murine EGFR antibody 806 was derived from mice immunized with mouse fibroblast cells expressing EGFR variant III (also called de2-7, abbreviated here as EGFRvIII). EGFRvIII arises from an in frame deletion of exons 2-7 of the EGFR gene. In the resulting variant, amino acids 6 – 273 of wild type EGFR are replaced by a single glycine; domain I and much of domain II are absent [168]. This aberrant form of EGFR is constitutively active [169], although it is not known if this is caused by ligand independent activation of properly folded EGFRvIII, or by mislocalization and aggregation. EGFRvIII arises spontaneously in a portion of human tumors (≈ 24 - 76% of glioblastomas [170, 171]; ≈ 42% of head and neck tumors [172]), and tumor therapies targeting EGFRvIII may benefit from high selectivity for tumor cells.

Antibody 806 was selected for its ability to specifically recognize EGFRvIII over wild type EGFR, though it also binds to cells overexpressing wild type EGFR [173]. Epitope mapping [29, 174] and structural studies [3] have determined the epitope of mAb 806 (and the related antibody mAb 175 [3]) to be a disulfide bonded loop composed of residues 287 – 302 of the mature receptor. This region of domain II is mostly buried in both the extended and tethered conformations of wild-type receptor (Fig. 40). It is plausible that the mAb epitope is disordered in EGFRvIII as well as in some population of over-expressed wild-type receptor.

If mAb 806 proves to be an effective therapeutic, it may have advantages over other antibody EGFR inhibitors. Antibodies that recognize full length, folded EGFR bind to epithelial tissues [175, 176]. This increases the dosage required to adequately target EGFR expressing tumors, and may contribute to therapeutic side effects. By contrast, the specificity of mAb 108 for EGFRvIII may result in greater selectivity for tumor cells and decreased immunological side effects [3, 128]. mAb 806 (chimerized to ch806, now termed Necitumumab) has completed Phase I trials [149] and has been licensed to Abbott Laboratories for further development.

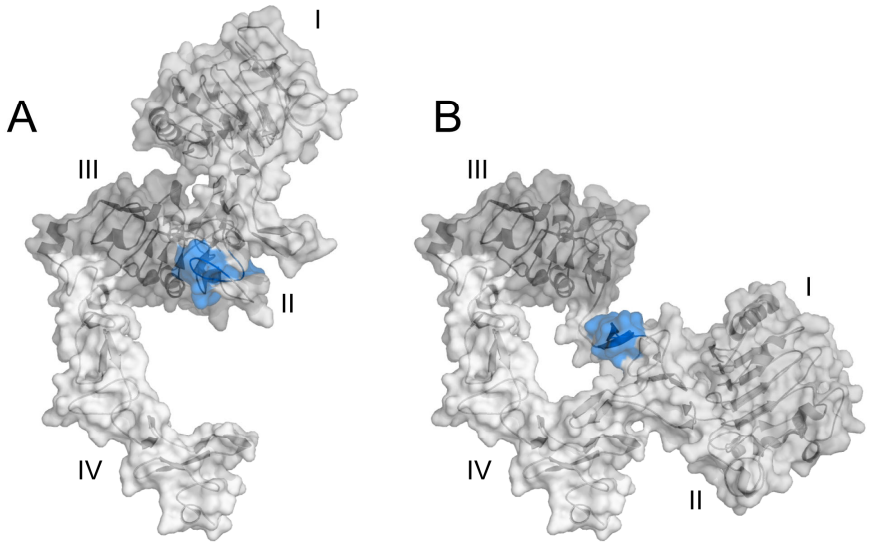

Figure 40. The epitope of mAb 806 on the EGFR extracellular region.

The epitope of antibody 806, amino acids 287 – 306 of EGFR, is highlighted (blue) in a molecular surface representation of EGFR extracellular region in (A) extended and (B) tethered conformations. In both conformations, the epitope is well packed within domain II and is inaccessible to mAb 806 [3]. mAb 806 is thought to recognize partially unfolded EGFR, as well as EGFR variant III, which lacks domain I and most of domain II [3]. The model of the extended conformation was generated from PDB ID 1IVO [2], and domain IV of this conformation was modeled from 1YY9 [5] by alignment with domain III. The tethered conformation was generated from chain A of 1YY9.

3.1.9 Other clinically important EGFR antibodies

Several other clinically important inhibitory EGFR antibodies have been described, but have not been as fully structurally characterized.

Inhibitory antibody E7.6.3 (panitumumab/Vectibix) was generated by immunizing transgenic XenoGraft mice, which produce fully human antibodies, with A431 cells [177, 178]. Panitumumab inhibits ligand binding and receptor phosphorylation, and epitope mapping indicates that its epitope overlaps with that of mAb C225 [9]. Panitumumab has been approved by the FDA for the treatment of colorectal cancer [145, 179], and is in trials for use against other cancers and in combination therapies.

Antibody 2F8 (HuMax-EGFr, zalutumumab) was created by immunizing HuMax transgenic mice, expressing fully human antibodies, with both A431 cells and purified EGFR [180]. Zalutumumab blocks ligand binding, receptor phosphorylation, and growth of EGFR expressing tumor cells [180]. Epitope mapping and electron microscopy protein tomography data are consistent with an epitope on domain III [10]. Zalutumumab is in Phase III trials for the treatment of head and neck cancer [146].

Murine antibody R3 was generated by immunizing mice with purified human EGFR, and has been humanized (h-R3/nimotuzumab/TheraCIM/Theraloc) for clinical investigation [144, 181]. Nimotuzumab bocks ligand binding and receptor phosphorylation in cells [144], and restricts tumor growth in xenograft mice [182]. The structure of the nimotuzumab Fab alone has been determined [143]. Modeling and mutational analysis suggest an epitope on domain III, partially overlapping that of cetuximab [143]. This antibody has been reported to produce few toxic side effects as a therapeutic [183], and has regulatory approval in a number of countries (including Brazil, India, and China [184]), though not yet in the US. Nimotuzumab is in Phase III trials for the treatment of glioma.

3.1.10 Non-clinical EGFR antibodies with unique properties

Several non-therapeutic antibodies directed against EGFR are notable for having unique binding or inhibitory characteristics. These antibodies have been characterized to some extent,

but they have not been the subject of structural study, and their epitopes have not been determined in detail.

Murine antibody 108 was generated by immunizing mice with EGFR expressing CHO cells, and was identified by its ability to bind A431 and HER14 cells (NIH-3T3 fibroblast cells stably expressing human EGFR) [25]. mAb 108 decreases ligand induced receptor phosphorylation [25] and inhibits tumor growth in mouse xenografts [185]. Though this antibody binds to domain III of the receptor [129], it does not robustly inhibit ligand binding [25]. Interestingly, mAb 108 specifically inhibits high affinity ligand binding, but not low affinity binding, at the cell surface.

Murine antibody 2E9 was generated by immunizing mice with membrane preparations from A431 cells, and was screened for the ability to pull down kinase functional EGFR [139]. mAb 2E9 binds to domain I of the receptor [129] and, in contrast to mAb 108, blocks only low affinity ligand binding at the cell surface [139].

Finally, murine antibody 13A9 was generated by immunizing mice with A431 cells followed by purified EGFR [140]. mAb 13A9 binds to domain III of the receptor [29] and inhibits the growth of A431 cells [157]. Curiously, this antibody competes with TGFα binding but not with EGF binding to EGFR [130, 140], even though the epitopes of these two ligands substantially overlap [1, 2] (Figs. 38A & B).

3.1.11 VHH inhibitors of EGFR

Conventional IgG antibodies, including the monoclonal antibodies detailed above, are composed of two heavy and two light chains (Fig. 41). The interaction with antigen is mediated by six complementarity determining regions (CDRs), three each on the heavy chain variable Ig domain (VH) and the light chain variable Ig domain (VL). The minimum antigen binding unit of these antibodies is a single Fab fragment (four Ig domains) or an Fv fragment (two variable Ig domains only). An unconventional class of IgG antibodies occurs naturally in camelids alongside conventional antibodies. Termed heavy chain antibodies (HCAb), these antibodies lack light

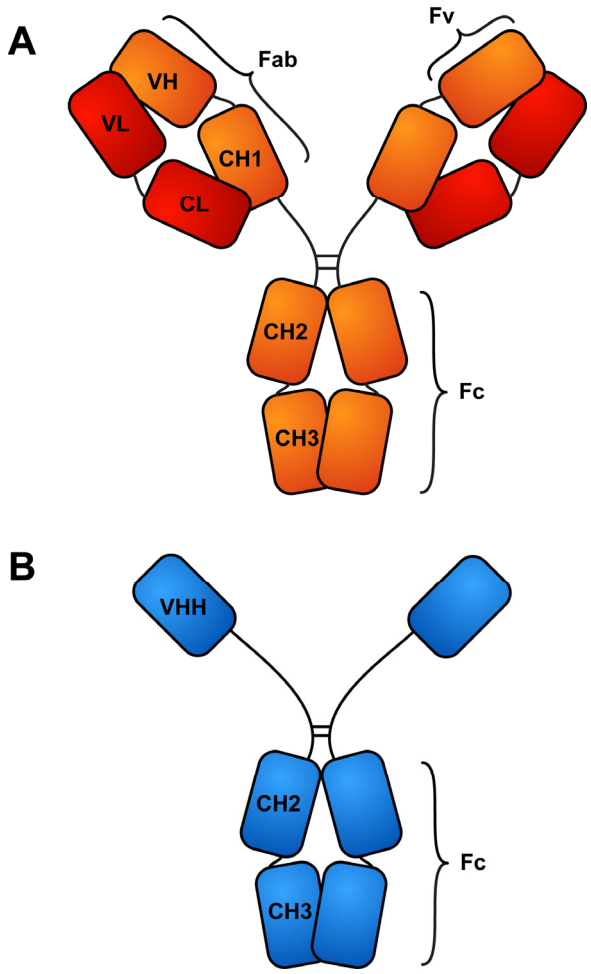

Figure 41. Conventional and heavy chain only antibody structure.

(A) A conventional IgG antibody is composed of two heavy and two light chains, with a total of twelve Ig domains. The antibody can be divided into three stably folded units: two Fab modules and one Fc module. Antigen binding is mediated by the heavy (VH) and light chain (VL) variable domains. An Fv fragment (one VH and one VL domain) is the smallest unit that contains an intact antigen binding surface. (B) Camelid HCAbs lack light chains as well as the heavy chain CH1 domain. The minimum antigen binding unit of a camelid HCAb is a single VHH domain.

chains and exist as homodimers of heavy chains only [186] (Fig. 41). Thus, the antigen binding surface of HCAbs is composed of only three CDRs on a modified heavy chain variable domain (VHH, also termed Nanobody [187]), and the minimum antigen binging unit is a single VHH (Fig. 41). (Single chain antibodies also occur in cartilaginous fish [termed IgNARs; immunoglobulin new antigen receptors], but arose independently and are only distantly homologous to HCAbs [188].) HCAbs lack the heavy chain constant domain (CH1) and often incorporate a relatively elongated hinge region (in IgG2a and IgG2b subtypes) [189]. The lengthened hinge increases the potential separation between antigen binding sites, partially compensating for the absence of CH1. Additionally, residues of the former VH/VL interface are altered in VHHs to compensate for the absence of the VL domain. By contrast, the HCAb Fc region is conventional and is a competent immune effector [190].

VHHs have received attention as an alternative to conventional mAbs as diagnostic and therapeutic agents, offering similar advantages to other truncated antibody constructs (diabodies [191], unibodies [192], single chain antibodies [193], etc.). As small single domain proteins, VHHs are well folded and are well produced in exogenous expression systems, including *E. coli*. Their small size may confer improved tissue penetration. Like conventional antibodies, VHHs are stable and resistant to thermal denaturation [194]; unlike conventional VH domains, VHH folding is reversible [194]. Also, notwithstanding the adaptations that have occurred to accommodate the absence of a VL, VHHs exhibit close homology to human VH domains. This may permit VHHs to be humanized for therapeutic use [195], although the antigenic properties of humanized VHHs are not well characterized.

Several reports describe anti-EGFR VHH fragments. VHH ORB1-83 was isolated from a phage display library of *Camelus bactrianus* VHH constructs on the basis of its ability to specifically recognize an N-terminal peptide from EGFRvIII [196, 197]. This VHH has been used as an immunohistochemical agent to study EGFRvIII expression in thyroid carcinoma [198]. VHHs have also been generated that recognize folded epitopes on the EGFR extracellular region [199, 200]. In one of these studies, *Llama glama* were immunized with intact A431 cells or A431

membrane preparations. Lymphocytes were isolated from immunized animals, and HCAb cDNA was used to construct a phage display library. Anti-EGFR VHHs from this library have been isolated, and have been found to block ligand binding, inhibit ligand dependent EGFR signaling, and restrict tumor growth [199]. These VHHs have also been used as delivery vehicles for radionuclides [201]. We have structurally characterized a subset of these VHH fragments bound to the EGFR extracellular region (see Section 3.3).

3.2 The application of structural knowledge to mAb inhibitors of EGFR

3.2.1 Domain level epitope mapping of antibody binding to sEGFR

To assess the domain specificity with which the antibodies in our panel (shown in bold and underlined in Table 6) interact with sEGFR, binding was measured between antibodies and wild type sEGFR, as well as three sEGFR truncation variants: (i) sEGFR501, which terminates after amino acid 501 and includes only the first disulfide bonded module of domain IV, (ii) sEGFRd3, which comprises the first four amino acids of mature EGFR followed by amino acids 311 to 514 of sEGFR, (iii) sEGFRd1, which terminates after amino acid 183, (iv) and sEGFR239, which terminates after amino acid 239, immediately preceding the dimerization arm. The mAb or corresponding Fab fragment was amine coupled to a Biacore surface, and samples of increasing concentration of sEGFR or the relevant altered form of sEGFR were passed over these surfaces. The equilibrium SPR response for each sample was plotted as a function of concentration and fit to a single site binding equation. Representative binding curves for each Fab surface are shown in Figure 42, and the mean equilibrium binding constants are listed in Table 7.

The binding of wild type sEGFR to Fab13A9 is the highest affinity for this group of antibodies (K_D = 1.1 ± 0.1 nM), similar to the affinity previously reported for binding to FabC225 (K_D = 2.3 ± 0.5 nM) [5]. Binding to Fab108 and Fab425 is somewhat weaker (K_D = 19.8 ± 2.2 nM

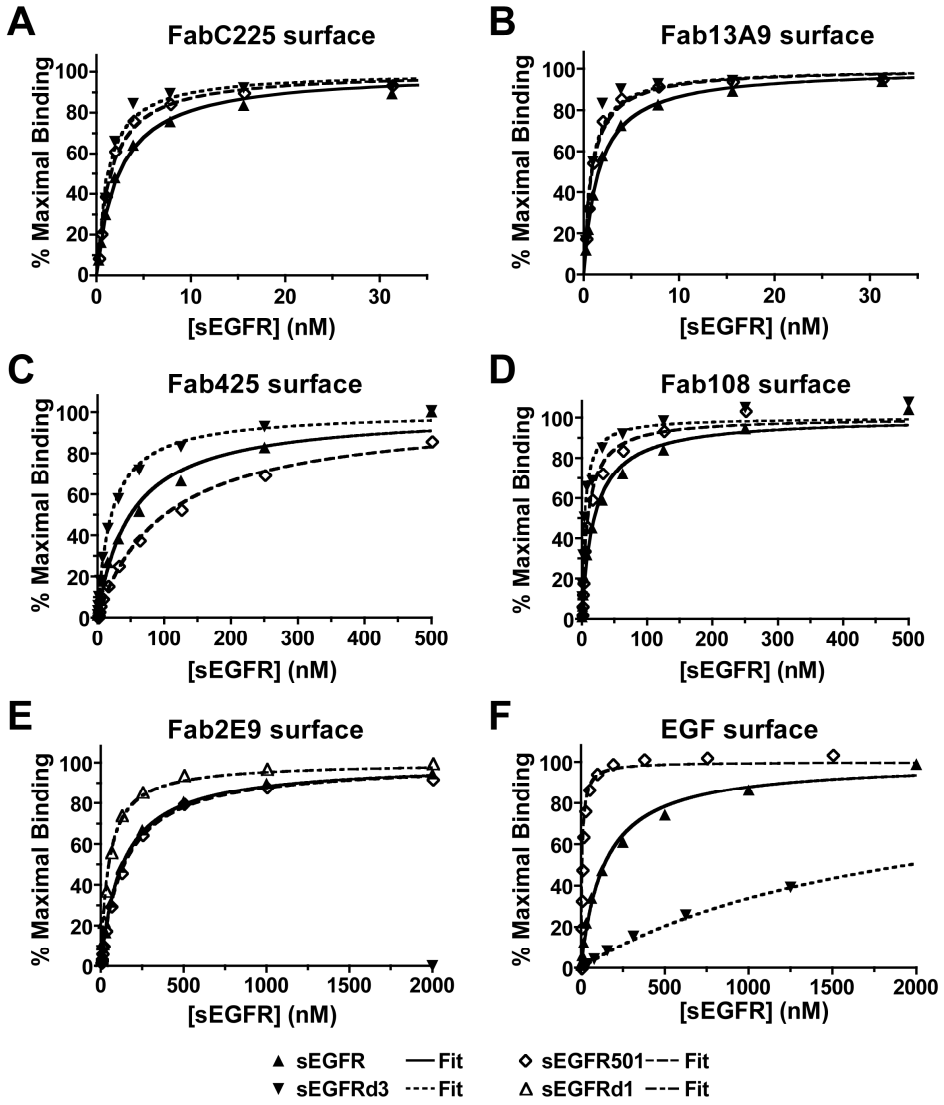

Figure 42. Equilibrium binding analysis for the interaction between sEGFR constructs and inhibitory Fab or EGF surfaces.

Increasing concentrations of the indicated sEGFR constructs were applied to Biacore surfaces of amine-coupled Fab or EGF ligand. The normalized equilibrium SPR response signal is plotted against analyte concentration. Representative binding curves are shown for each analyte/surface combination. Data were fit to a single site binding isotherm, and the resulting line of best fit is shown. K_D values are reported in Table 7.

and 46.4 ± 8.7 nM, respectively), and that to Fab2E9 is significantly less (K_D = 130 ± 18 nM). In each case, the equilibrium binding constants to immobilized mAb and Fab are similar, as is expected in this assay format since there is no avidity effect due to the bivalence of the antibody when it is immobilized on the surface. We were unable to obtain surface immobilization for mAb 2E9, thus no K_D is reported.

The longest of the sEGFR variants, sEGFR501, bound to all Fab surfaces. For Fabs C225, 108 and 13A9 the K_D values for the binding of sEGFR501 were the same or slightly higher affinity than for full-length sEGFR (Fig. 42 & Table 7), while against the Fab425 surface the affinity was reduced by approximately 2 fold. This trend is in stark contrast to the binding of sEGFR501 to immobilized EGF. As has previously been reported [21, 202], sEGFR501 binds to EGF with approximately 25 fold stronger affinity.

All of the antibodies, with the exception of mAb 2E9, bind to domain III of sEGFR. Indeed, sEGFRd3 binds with higher affinity to each immobilized mAb or Fab than is observed for full-length sEGFR, possibly due to the absence of steric effects imposed by the other domains of the receptor. Both sEGFRd1 and sEGFR239 bind to immobilized Fab2E9, with respective K_D values of 59.2 ± 14 nM and 65.7 ± 17 nM, confirming that the epitope for this antibody lies in domain I of EGFR, as was predicted on the basis of experiments using chimeric chicken/human EGFR molecules [129].

3.2.2 Domain III binding antibodies have distinct but overlapping epitopes.

To more precisely map the epitopes of those antibodies that interact with sEGFRd3, a panel of site specifically altered sEGFR variants was utilized. These altered sEGFR proteins were designed based on the X-ray crystal structures of sEGFR in complex with EGF/TGFα [2, 156] with FabC225 [5] and with the Fab fragment from matuzumab (FabEMD72000) [6]. For each sEGFR variant we determined the equilibrium binding constant for its interaction with each immobilized Fab. Data are summarized in Figs. 43 and 44, and K_D values are listed in Table 7.

Table 7. Equilibrium binding constants of sEGFR and sub-domain fragments to immobilized Fabs, mAbs and ligands.

K_D values are nM, unless specified otherwise.

	sEGFR	sEGFR501	sEGFRd3	sEGFRd1		sEGFR	sEGFRd3
FabC225	2.3 ± 0.3 nM	1.6 ± 0.2 nM	1.2 ± 0.3 nM	> 50 μM	**mAb C255**	2.37 ± 0.2 nM	0.26 ± 0.06 nM
Fab425	46.4 ± 8.7 nM	100.5 ± 0.6 nM	20.6 ± 1.4 nM	> 50 μM	**mAb 425**	53.9 ± 15 nM	10.6 ± 0.4 nM
Fab108	19.8 ± 2.2 nM	12.1 ± 1.0 nM	4.7 ± 0.5 nM	> 50 μM	**mAb 108**	22.4 ± 3.8 nM	2.35 ± 0.2 nM
Fab13A9	1.1 ± 0.1 nM	0.83 ± 0.02 nM	0.4 ± 0.02 nM	> 50 μM	**mAb 13A9**	1.47 ± 0.17 nM	0.24 ± 0.04 nM
Fab2e9	130 ± 18 nM	140.8 ± 1.2 nM	> 50 μM	59.2 ± 14 nM			
EGF	130 ± 4 nM	7.8 ± 3.5 nM	2.4 ± 0.4 μM	> 50 μM	**TGFα**	335 ± 31 nM	11.2 ± 1.2 μM

Table 8. Equilibrium binding constants for the binding of sEGFR with substitutions in domain III to immobilized Fabs and ligands.

Values are nM, unless specified otherwise.

	D355T/ F357A	Q408M/ H409E	Q408A/ H409A	Q384A	K443A	S468I	N473A	K443A/ N473A	K454A	K463A	S460A/ T459A
FabC225	4.2 ± 0.9	340 ± 50	25.1 ± 4.6	3.9 ± 1.0	2.9 ± 0.8	2.0 ± 0.8	1.9 ± 0.1	47.3 ± 0.5	2.0 ± 0.1	6.7 ± 0.3	1.8 ± 0.1
Fab425	97.5 ± 13	127 ± 27	65.8 ± 5.2	83.6 ± 16.1	40.8 ± 2.4	56.8 ± 0.9	66.6 ± 4.1	104 ± 4	1600 ± 200	> 10 µM	> 10 µM
Fab108	25.0 ± 2.8	33.2 ± 7.5	26.2 ± 6.8	18.1 ± 1.8	19.8 ± 1.3	24.0 ± 9.9	23.4 ± 8.1	33.4 ± 2.8	747 ± 7	6770 ± 760	5390 ± 186
Fab13A9	1.9 ± 0.4	3.3 ± 0.7	1.4 ± 0.4	1.2 ± 0.4	27.0 ± 0.3	49.0 ± 10	94.4 ± 9.3	> 10 µM	2.8 ± 1.3	9.9 ± 3.0	4.0 ± 1.7
EGF	> 10 µM	480 ± 70	51.1 ± 8.8	1000 ± 300	157 ± 27	140 ± 20	182 ± 25	nd	289 ± 13	229 ± 8	161 ± 2

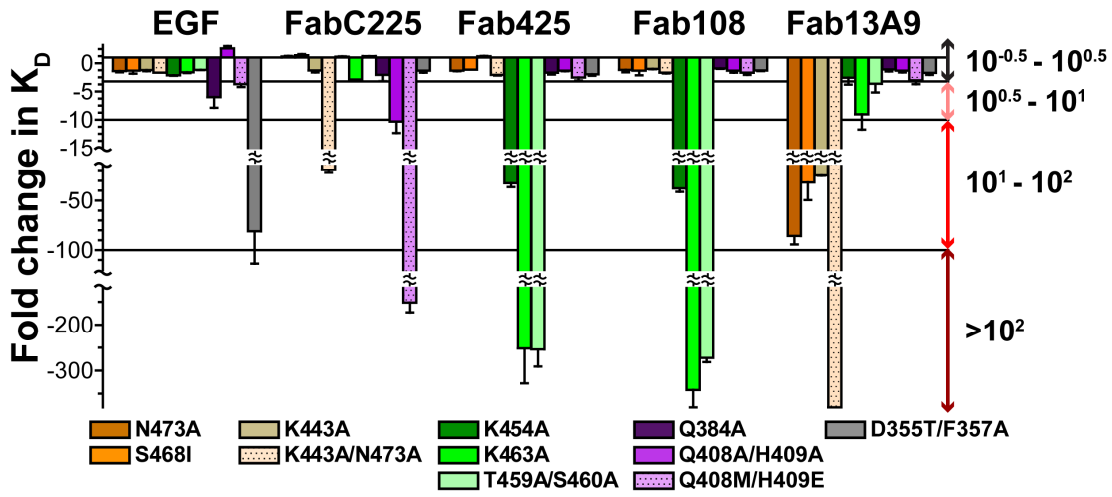

Figure 43. Amino acid substitutions on the surface of EGFR domain III specifically affect binding of ligand and inhibitory antibodies.

Binding of the indicated sEGFR variants to the indicated interaction partner was determined by Biacore (Table 7). Data are represented as fold change in K_D, versus biding to wild type sEGFR, on a logarithmic scale. (See Fig. 44)

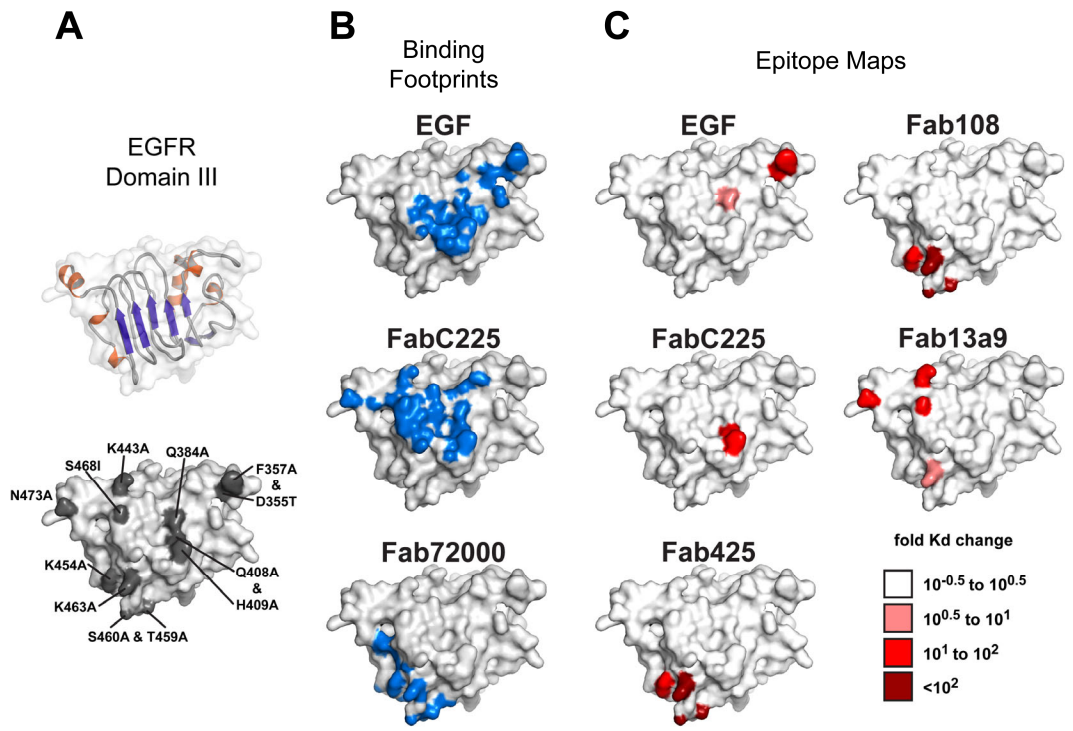

Figure 44. Epitope mapping on the surface of EGFR domain III.

(A) A cartoon of the secondary structure of EGFR domain III (a.a. 310 – 480) is shown with a transparent molecular surface (upper panel). Point mutations (gray) that disrupt binding of EGF, Fab425, Fab225, or both EGF and Fab225 are shown on the surface of domain III (lower panel). (B) The footprints of EGF, Fab225, and Fab72000 are shown on domain III. Blue surface indicates atoms on domain III within 4 Å of the respective binding partner. (C) Binding affinities of each indicated binding partner for each sEGFR variant indicated in (A) was determined by Biacore. The fold decreases in K_D was sorted on a logarithmic scale into four bins, and represented on the surface of domain III by color (white represents least change; dark red represents greatest decrease in binding). (See Fig. 43)

As we have previously reported, there is a “hot spot” for FabC225 binding at the center of the FabC225 epitope on domain III, close to a hydrophobic pocket that accommodates Y98 (Kabat numbering scheme) of FabC225 (Q408A/H409A) [5]. This same hydrophobic pocket accommodates the critical L47 of EGF, and substitutions in this region also affect EGF binding. By contrast, alterations in this region have no effect on the binding of Fab425. Fab425 binding is all but abolished against sEGFR variants with substitutions in the crystallographically defined epitope of FabEMD72000, as expected. Interestingly, these same alterations in domain III (K463A, K454A and T469A/S460A) have a similar dramatic effect on binding of sEGFR to immobilized Fab108, suggesting that the epitope of mAb 108 overlaps significantly with that of mAb 425.

There is a modest effect of the K463A variant upon the binding to Fab13A9. By contrast, three substitutions clustered near the C-terminal α-helix of domain III (K443A, S468I and N473A) result in a substantial decrease in sEGFR binding to Fab13A9, suggesting that the epitope for mAb 13A9 lies in this region of domain III. S468, as well as nearby residue E472, were previously identified as important for mAb 13A9 binding by yeast surface display [29]. This region of domain III lies within the C225 epitope but outside of the ligand binding surface. Individually, these single amino acids substitutions have minimal effect on the binding of sEGFR to immobilized FabC225, but a variant in which both K443 and N473 are substituted with alanine (K443A/N473A) does show significant reduced binding to FabC225, and binding to Fab13A9 is almost abolished (Fig. 43 & Table 7).

This limited fine epitope mapping of domain III suggests that some antibodies directly compete for binding to domain III, while others may bind simultaneously. To test this hypothesis and compliment the epitopes identified through amino acids alterations, we next investigated the effect of addition of an excess of each Fab upon the binding of sEGFR (100 nM) to each immobilized Fab (Fig. 45). Addition of a 10 fold molar excess of Fab2e9 (white bars) does not substantially reduce the equilibrium SPR value observed for the binding of sEGFR to any of the immobilized Fabs tested, as would be expected since the epitope for this antibody is quite

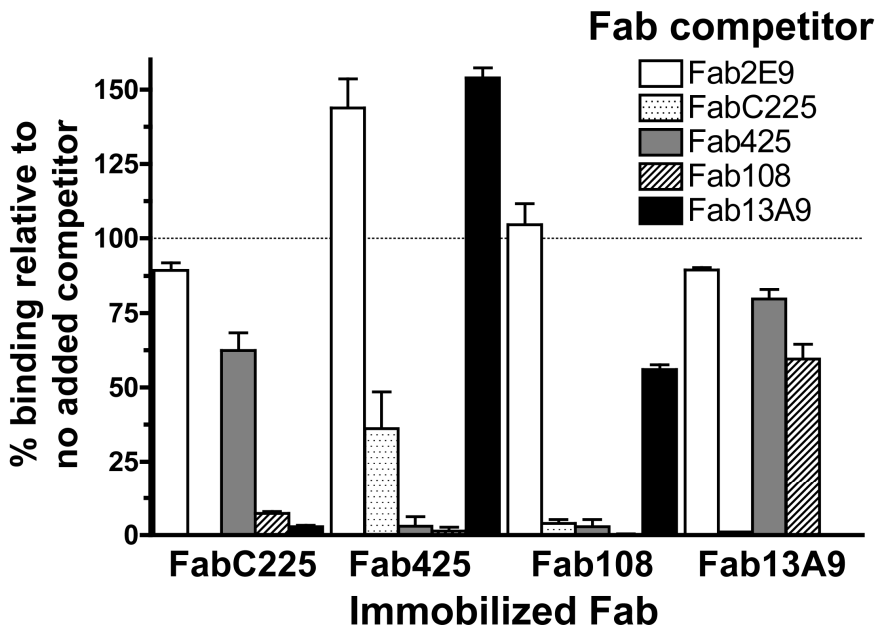

Figure 45. Binding competition of inhibitory EGFR Fabs.

The normalized SPR response is shown for equilibrium binding of 100 nM sEGFR to the indicated immobilized Fab in the presence of a 10 fold molar excess of the indicated soluble competitor Fab. The Fab 2E9, which binds domain I, does not compete with the domain 3 binding Fabs. Fab C225 competes strongly with Fabs 108 and 13A9, but only weakly with Fab 425. Fabs 425 and 108 compete strongly with one another. Fab 13A9 competes weakly with Fabs 425 and 108.

distinct from the others in this group and lies on domain I. Excess Fab225 (stippled bars) diminishes binding of sEGFR to all Fab surfaces, although only by about 50% to immobilized Fab425 (gray bars). A similar pattern of competition is seen when FabC225 is the immobilized species: addition of Fab13a9 (black bars) or Fab108 (striped bars) significantly reduces binding of sEGFR to immobilized FabC225, while addition of Fab425 (gray bars) reduces the SPR response for binding to FabC225 by only about 50%.

For both the Fab108 and Fab425 surfaces, addition of excess Fab425 or excess Fab108 blocks all binding of sEGFR, consistent with the suggestion that these two antibodies share a similar epitope. The epitopes for 425 and 108 are however not identical, as is evident in the ability of excess Fab13A9 to diminish the binding of sEGFR to Fab108 by about 60%, while the binding of sEGFR to Fab425 is not affected. Indeed there is an increase in the SPR response for sEGFR binding to Fab425 in the presence of excess Fab13A9, and vice versa, implying that Fab13A9 and Fab425 can bind to sEGFR simultaneously, and that their epitopes do not overlap.

3.2.3 Monoclonal antibodies 425, 108 and 13A9 do not abolish binding of sEGFR to EGF.

We next asked what effects the binding of these antibodies to their respective epitopes on domain III has on the ability of sEGFR to bind the ligand EGF. Samples of sEGFR at 600 nM containing increasing molar excesses of mAb C225, 425, 108, or 13A9 were passed over a Biacore surface to which EGF had been covalently coupled. As shown in Figure 46A, the binding of sEGFR to the EGF is abolished by addition of 0.5 fold molar excess of mAb C225 (an equimolar number of antigen binding sites). This is expected since the epitope of this mAb overlaps with ligand binding site on domain III of sEGFR, and we have previously reported that the FabC225 and FabIMC-11F8 compete efficiently for the binding of sEGFR to immobilized EGF [5, 147].

For the other antibodies, increasing molar excesses of mAb do not bring the relative equilibrium binding response to baseline, indicating that sEGFR is still able to interact with

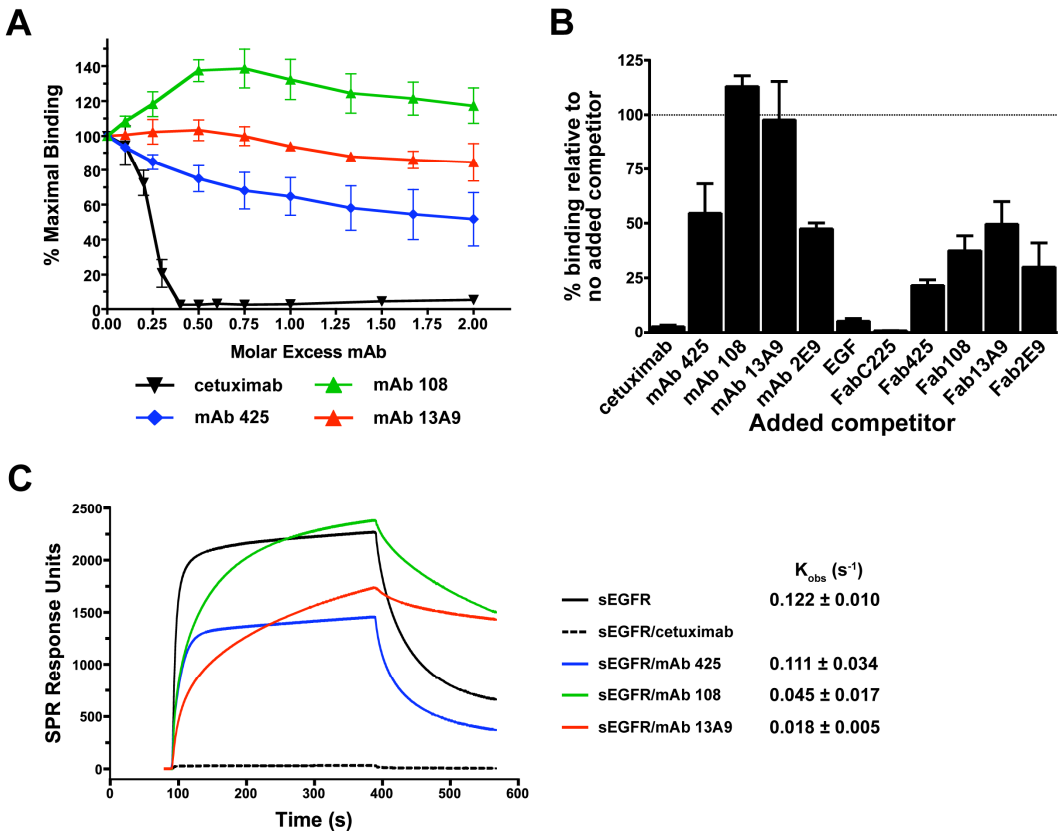

Figure 46. Ligand binding competition by inhibitory mAbs and Fabs.

(A) The equilibrium SPR binding of 600 nM sEGFR to immobilized EGF is shown in the presence of the indicated molar excess of mAb C225 (cetuximab), 425, 108, or 13A9. mAb C225 exhibits essentially complete inhibition of ligand binding at 0.5 molar excess. (B) The fold change in equilibrium binding of 600 nM sEGFR to immobilized EGF in the presence of soluble competitors is shown, relative to sEGFR binding alone. Binding was assayed with a 2 fold excess of mAbs C225, 425, 108, 13A9, and 2E9, a 5 fold excess of the respective Fabs, and a 5 fold excess of EGF. (C) SPR sensograms showing association and dissociation curves of sEGFR to immobilized EGF in the presence of a 2 fold molar excess of mAbs C225, 425, 108, or 13A9. Apparent association rates (K_{obs}) are shown.

immobilized EGF in the context of this assay. Rather, above a 0.5 molar excess of these mAbs (an equimolar concentration of EGFR binding sites), SPR response remains relatively constant. This could be explained either if (i) sEGFR but not the mAb:sEGFR complex binds to immobilized EGF, (ii) the mAb:sEGFR complex binds but with a substantially lower affinity than sEGFR alone, or (iii) a combination of these two effects occurs. Addition of mAb 108 to sEGFR in this assay leads to an initial increase in the SPR response that plateaus at a 1:1 ratio of sEGFR to antigen binding regions on the antibody (0.5 fold molar excess of antibody). This indicates that in this case there must be some association of the mAb:sEGFR complex. Addition to the sEGFR of a 2 fold molar excess of a control antibody (Ab-3, clone EGFR.1 NeoMarkers/Lab Vision) that does not block ligand binding to EGFR [138] resulted in an 2.5 fold larger SPR response compared to sEGFR alone, consistent in this case with binding of a 1:1 sEGFR:mAb complex to the immobilized EGF.

Examination of the SPR sensograms (Fig 46C) indicates that the apparent association rates for the binding of a 1:2 mixture of sEGFR and mAb 108 or 13A9 are substantially slower than for sEGFR alone (2.7 and 6.7 fold slower respectively). In contrast the association rate for the mixture with mAb 425 is not significantly altered (although in this case the equilibrium SPR signal is reduced by approximately 50 %). The dissociation phase for the mixtures with mAb 108 and 13A9 are also slower than for sEGFR alone, but the SPR response did not return to baseline without regeneration by low pH buffer. We did not attempt to fit the dissociation phase of the sensogram. The binding of mAb 108 and mAb 13A9 to sEGFR clearly influences the ability of sEGFR to bind to the immobilized EGF.

In Fig 46B the equilibrium binding response observed at 2 fold excess of mAb or 5 fold excess of Fab are shown for C225, 425, 108, and 13A9, as well values for 2 fold excess mAb 2E9 and 5 fold excess of EGF. Antibody 2e9 reduces the SPR response to a similar extent as mAb 425 and also has no effect on the apparent association rate compared to sEGFR alone (data not shown). In contrast to the effects observed upon addition of excess mAbs to sEGFR, addition of excess Fab to sEGFR reduces the SPR signaling for all Fabs tested, including

Fab425, 108 and 13A9 (Fig. 46B). The trends in the effect of addition of a 5-fold molar excess of Fab to sEGFR are however the same as for the mAbs. FabC225 abolishes all binding of sEGFR to immobilized EGF, while Fab425 and Fab2E9 reduce the binding to about 25% of that observed in the absence of added competitor. The effects for Fab108 and Fab13A9 are less, reducing the binding value by about 50%.

Addition of excess Fab reduces the SPR signal to a greater extent than observed for addition of each mAb, possibly due to the lower mass of Fab/sEGFR complexes compared to the mAb/sEGFR complexes that may associate with the EGF.

3.2.4 Monoclonal antibody 13A9 competes efficiently for binding of sEGFR to TGFα but not to EGF

We next tested the ability of mAb C225 and mAb 13A9 to compete with ligand binding to purified sEGFR. Samples containing 600 nM sEGFR and increasing molar excesses of mAb C225 or 13A9 were passed over a Biacore surface of immobilized EGF or TGFα. The equilibrium binding response was measured and is plotted against fold excess mAb in Figure 47A. A 0.5-fold molar excess of mAb C225 or mAb 13A9 (an equimolar concentration of receptor binding sites) completely eliminates equilibrium binding to immobilized EGF. At the same concentration, mAb C225 abolishes binding to immobilized TGFα. By contrast, and in accordance with prior reports [130, 140], mAb 13A9 reduces binding to immobilized TGFα by only $\approx$ 15%, even at a 2-fold molar excess.

We further compared the ability of mAb and Fab fragments of 13A9, C225, and 108 to inhibit binding of sEGFR to ligand. 600 nM of sEGFR combined with a 2-fold molar excess of mAb, a 5-fold excess of Fab, or 5-fold excess of EGF was passed over a Biacore surface of immobilized EGF or TGFα. The equilibrium binding response was measured, and is shown in Figure 47B as fold change relative to sEGFR binding in the absence of antibody. mAb C225 and its Fab fragment essentially abolish binding of sEGFR to both EGF and TGFα. The Fab fragment of 13A9 behaves similarly to full antibody, competing strongly with binding to TGFα, but not

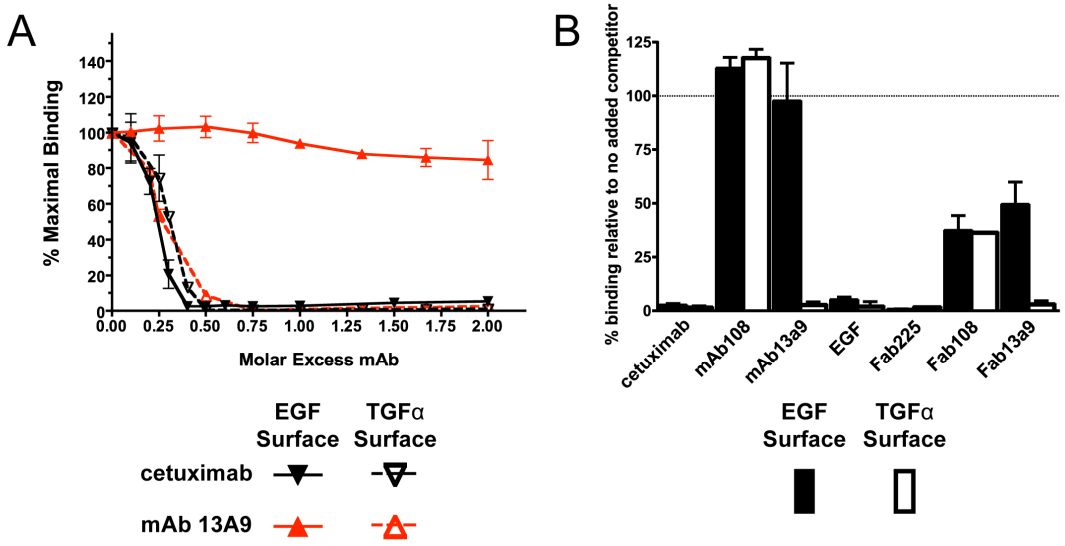

Figure 47. Antibody 13A9 competes for TGFα but not EGF binding to EGFR.

(A) The normalized SPR response is shown for equilibrium binding of 600 nM sEGFR to an EGF or TGFα surface in the presence of the indicated molar excess of mAb. An antibody:sEGFR ratio of 0.5:1 (equivalent to a 1:1 antigen binding site:sEGFR stiochiometry) effectively competes for sEGFR binding to immobilized TGFα but not EGF. Cetuximab competes for binding to both ligands. (B) Normalized SPR response is shown of 600 nM sEGFR binding to EGF or TGFα surfaces in the presence of an excess of the indicated mAb, Fab, or EGF, as in 46B. Cetuximab completely bocks binding to both ligands. Antibody 108 does not robustly block binding of either ligand. mAb and Fab 13A9 strongly competes for binding to TGFα, but only fractionally diminishes EGF binding.

EGF. By contrast, the presence of excess mAb 108 results in a slight increase in equilibrium response units for the interaction of sEGFR with both EGF and TGFα surfaces. The presence of Fab 108 diminishes equilibrium response values to approximately 40%.

3.3 Inhibition of EGFR activation by VHH antibody fragments

3.3.1 Inhibitory VHHs 7D12, IA1, and 9G8 bind to and inhibit EGFR

VHH antibody fragments have been evaluated as alternatives to conventional antibodies for a variety of applications [203, 204]. The anti-EGFR VHHs 7D12, IA1, and 9G8 were generated by immunizing *Llama glama* with EGFR expressing A431 cells and A431 membrane preparation [199, 201]. Lymphocytes were isolated and used to create an "immune" phage library of VHHs. VHH fragments were selected on the basis of their ability to specifically bind immobilized EGFR, and for competition with EGF. VHH 7D12 competes with mAb C225/cetuximab for binding to EGFR; VHHs IA1 and 9G8 compete with mAb 425 (the murine predecessor to humanized matuzumab/EMD72000) [205]. All three antibodies have been shown to inhibit receptor activation [199, 205].

VHH fragments have also been the subjects of structural studies, both alone and in complex with antigen [34, 35, 206-209]. The absence of a VL component and the presence of a lengthened CDR3 give VHHs a more wedge shaped paratope than Fabs, and VHH fragments often recognizes clefts or concave features on their antigen [204]. We sought to determine whether VHHs 7D12, IA1, and 9G8 recognize epitopes similar to those of conventional inhibitory antibodies, and how their smaller, more convex antigen binding surfaces interact with EGFR.

3.3.2 VHHs 7D12, IA1, and 9G8 recognize the EGFR extracellular region and bind with 1:1 stoichiometry

We first determined the dissociation constant for the interaction between soluble EGFR extracellular region (sEGFR) and immobilized VHH 7D12, IA1, or 9G8 by surface plasmon resonance (SPR; Biacore) (Fig. 48). To minimize bias due to orientation of the VHH, SPR assays were conducted with VHHs immobilized by two separate methods. Purified VHHs bearing a

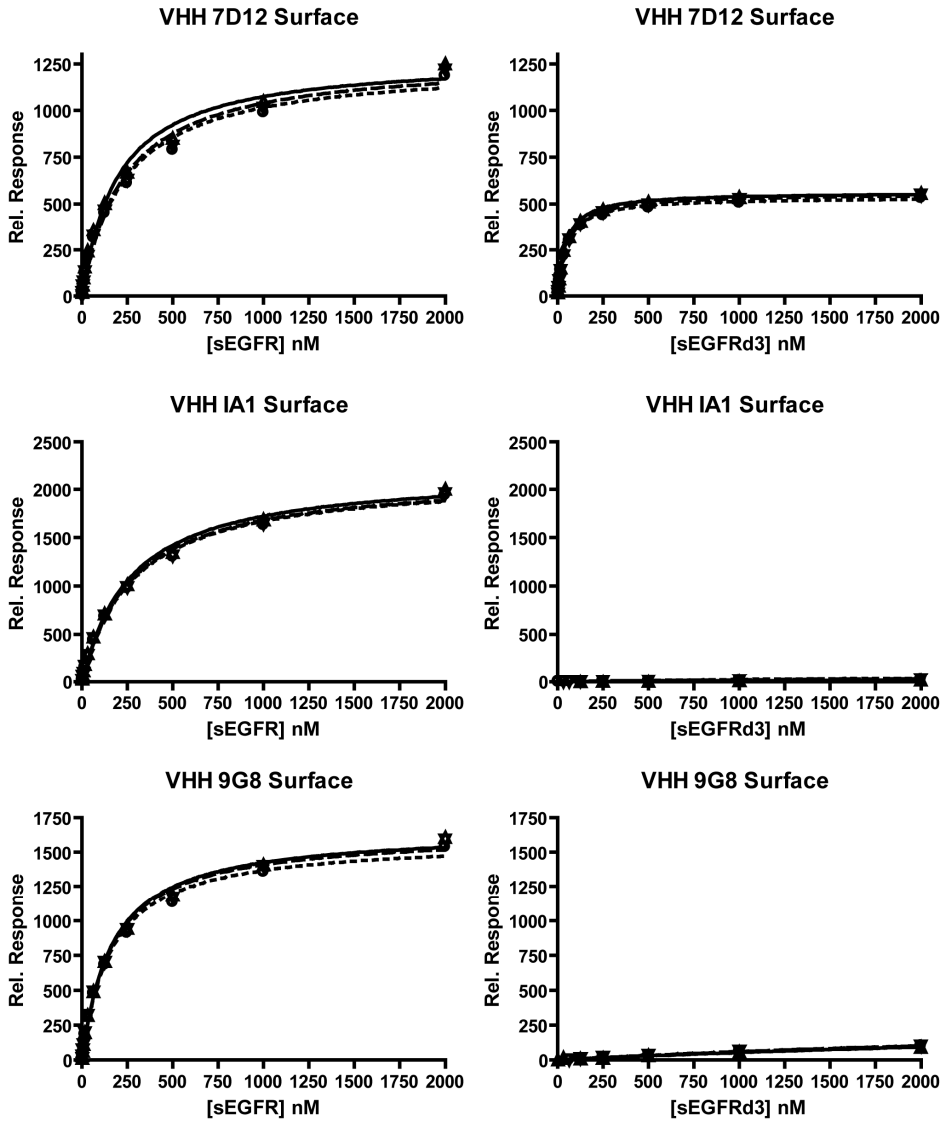

Figure 48. Equilibrium binding of sEGFR constructs to immobilized VHH.

Increasing concentrations of the indicated sEGFR constructs were applied to Biacore surfaces of amine-coupled VHH. The normalized equilibrium SPR response signal is plotted against analyte concentration for three replicate curves. Data were fit to a single site binding isotherm, and the resulting lines of best fit are shown. Replicate binding Dissociation constants are reported in Table 7.

Table 8. Equilibrium binding constants of sEGFR or sEGFRd3 to immobilized VHHs. Values are nM, unless otherwise specified.

	sEGFR	sEGFRd3
VHH 7D12	219 ± 20	47 ± 3.6
VHH 7D12 (SA)[1]	279 ± 19	
VHH IA1	276 ± 7.0	> 2.5 μM
VHH IA1 (SA)[1]	238 ± 42	
VHH 9G8	166 ± 1.2	> 5 μM
VHH 9G8 (SA)[1]	263 ± 76	

1. SA surfaces were created by immobilizing exogenously biotinylated VHH-AVIs to streptavidin coated SA sensor chips. All other surfaces were created by amine-coupling VHH-AVIs to CM5 sensor chips.

carboxy terminal *birA* biotinylation sequence followed by a poly-histidine tag were immobilized to biosensor chips by amine coupling, or were exogenously biotinylated and bound to streptavidin coated biosensor chips. The equilibrium SPR response was measured for a range of sEGFR concentrations. sEGFR binds to immobilized VHHs with dissociation constants (K_Ds) between 150 and 300 nM (Table 8). These affinities are 10 – 100 fold weaker than those observed for inhibitory antibodies and Fab fragments but are similar to those for ligand binding (Table 7).

The binding of a soluble construct comprising EGFR domain 3 (sEGFRd3) to VHHs 7D12, IA1, and 9G8 was assayed, as well. sEGFRd3 binds to VHH 7D12 with ≈ 4 fold tighter affinity than sEGFR, indicating that its epitope lies on this domain. Tighter binding to domain III than to full length sEGFR has been noted for conventional inhibitory mAbs that recognize domain III, presumably due to reduced steric effects (see Section 3.2.1). However, sEGFRd3 binds with > 10 fold weaker affinity to VHHs IA1 and 9G8, compared to sEGFR binding, suggesting that the epitopes of these VHH fragments lie at least partially outside domain III.

We confirmed the interactions observed in SPR experiments by examining the solution binding of these VHHs to sEGFR by sedimentation velocity analytical ultracentrifugation (SV AUC). 5 μM VHH and 5 μM sEGFR, or truncation variants, were analyzed individually or in combination. This concentration is ≥ 5-fold above the measured K_D for the VHH/sEGFR interactions, so stoichiometric mixtures of VHH and sEGFR are expected to sediment predominantly as a single species. We observe that VHH 7D12 binds stoichiometrically to both sEGFR and sEGFRd3 (Fig. 49), whereas VHH IA1 binds sEGFR but not the sEGFRd3 construct (Fig. 50).

To compare the epitope specificity of these VHH fragments we examined their ability to simultaneously bind sEGFR. SV AUC analysis of 5 μM sEGFR in the presence of 5 μM each of VHH 7D12 and VHH IA1 reveals the formation of a ternary complex (Fig. 51), as does analysis of sEGFR with both VHH IA1 and Fab C225 (Fig. 52). Thus, VHH IA1 is capable of simultaneous binding to sEGFR with either VHH 7D12 or Fab C225, indicating that VHH IA1 recognizes a

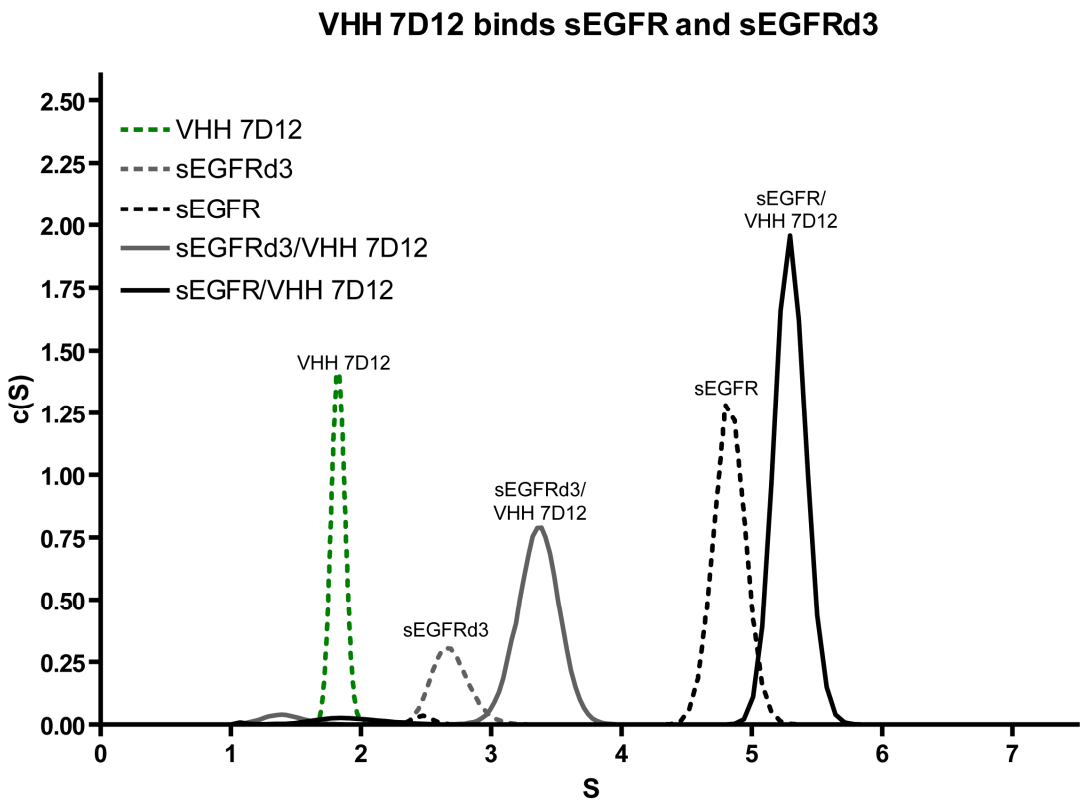

Figure 49. VHH 7D12 binds to sEGFR, and to sEGFRd3.

Sedimentation velocity c(S) species analysis of 5 μM VHH 7D12, 5 μM sEGFRd3, 5 μM sEGFR, or stoichiometric mixtures of these species. VHH 7D12 (dashed green), sEGFRd3 (dashed gray), and sEGFR (dashed black) sediment as discrete monodisperse species. VHH 7D12 sediments as a complex with either sEGFR (black) or sEGFRd3 (gray). Thus, the VHH 7D12 epitope lies on domain 3 of the EGFR extracellular region.

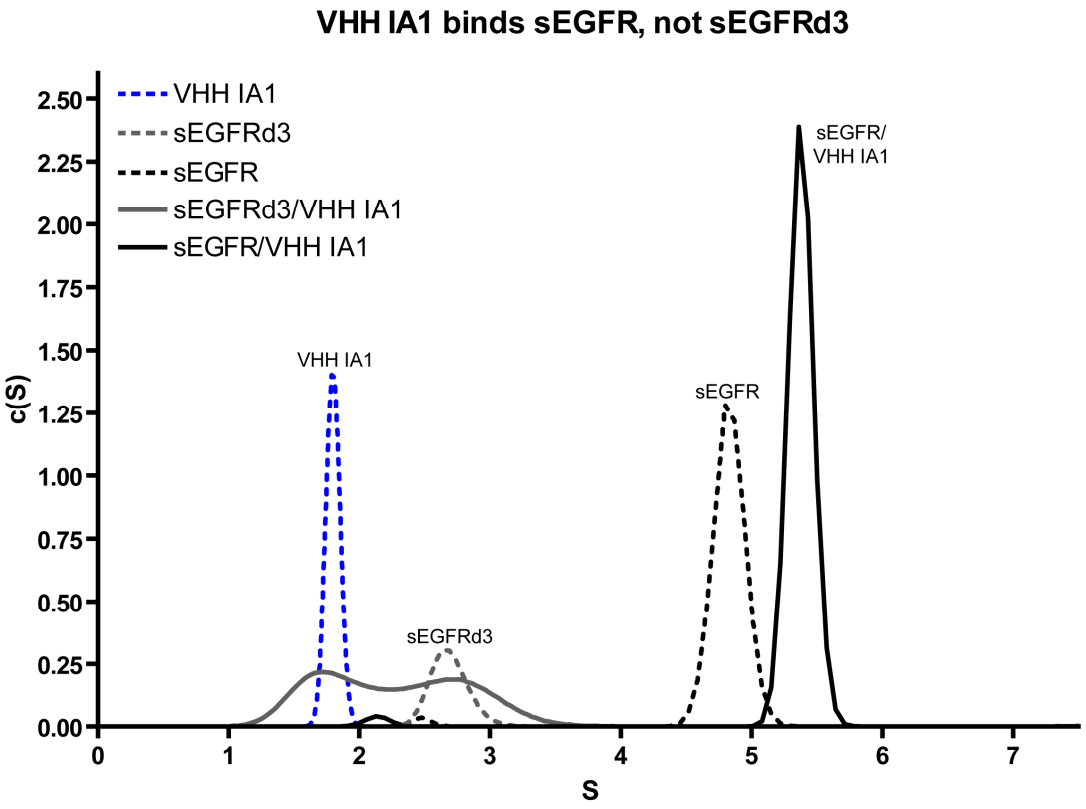

Figure 50. VHH IA1 binds to sEGFR, but not to sEGFRd3.

Sedimentation velocity c(S) species analysis of 5 μM VHH IA1, 5 μM sEGFRd3, 5 μM sEGFR, or stoichiometric mixtures of these species. VHH IA1 (dashed blue), sEGFRd3 (dashed gray), and sEGFR (dashed black) sediment as discrete monodisperse species. A mixture of sEGFR and VHH IA1 (black) sediments as a single complex, while a mixture of sEGFRd3 and VHH IA1 (gray) sediments as separate species and does not form complex. This suggests that the VHH IA1 epitope lies partially or entirely outside of the region contained in the sEGFRd3 construct (amino acids 1-4 fused to amino acids 310-514).

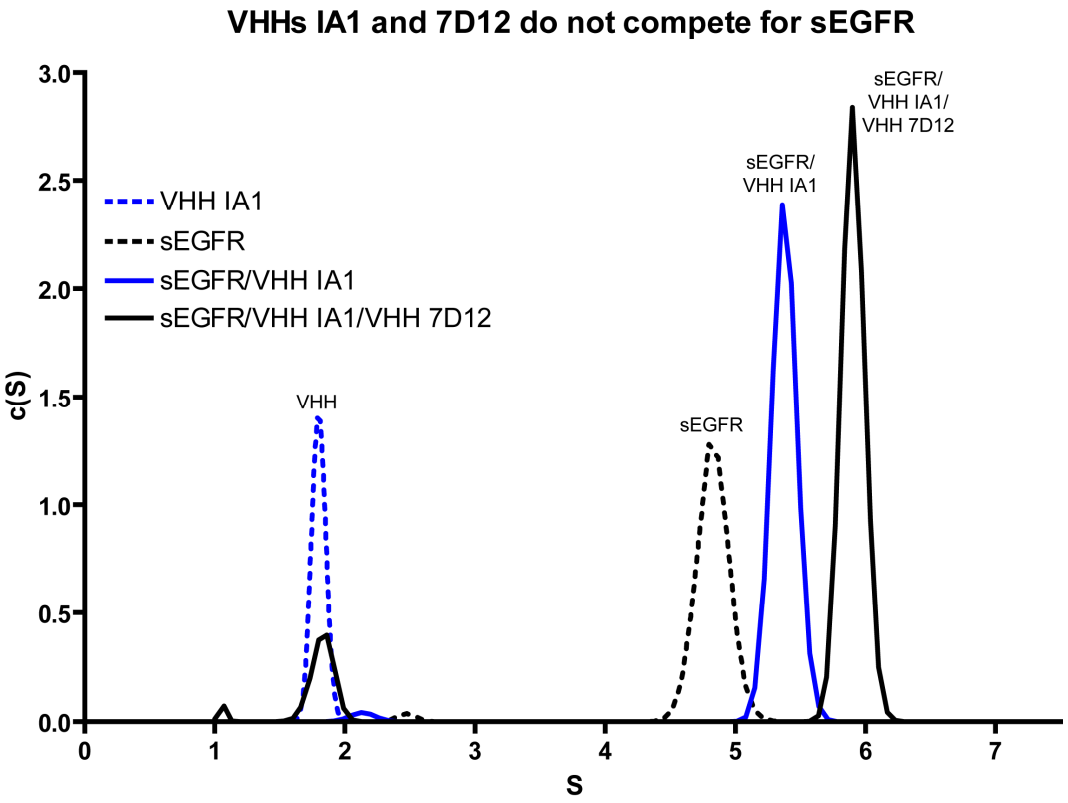

Figure 51. VHHs IA1 and 7D12 do not compete for binding to sEGFR.

Sedimentation velocity c(S) species analysis of 5 μM VHH IA1 (dashed blue), 5 μM sEGFR (dashed black), or stoichiometric mixtures of sEGFR and VHH IA1 (blue), or sEGFR, VHH IA1, and VHH 7D12 (black). A mixture of sEGFR with VHH IA1 and VHH 7D12 results in a single sedimenting species, corresponding to sEGFR bound to two VHH fragments. Thus, VHHs IA1 and 7D12 are able to simultaneously bind to sEGFR, indicating that they recognize distinct epitopes.

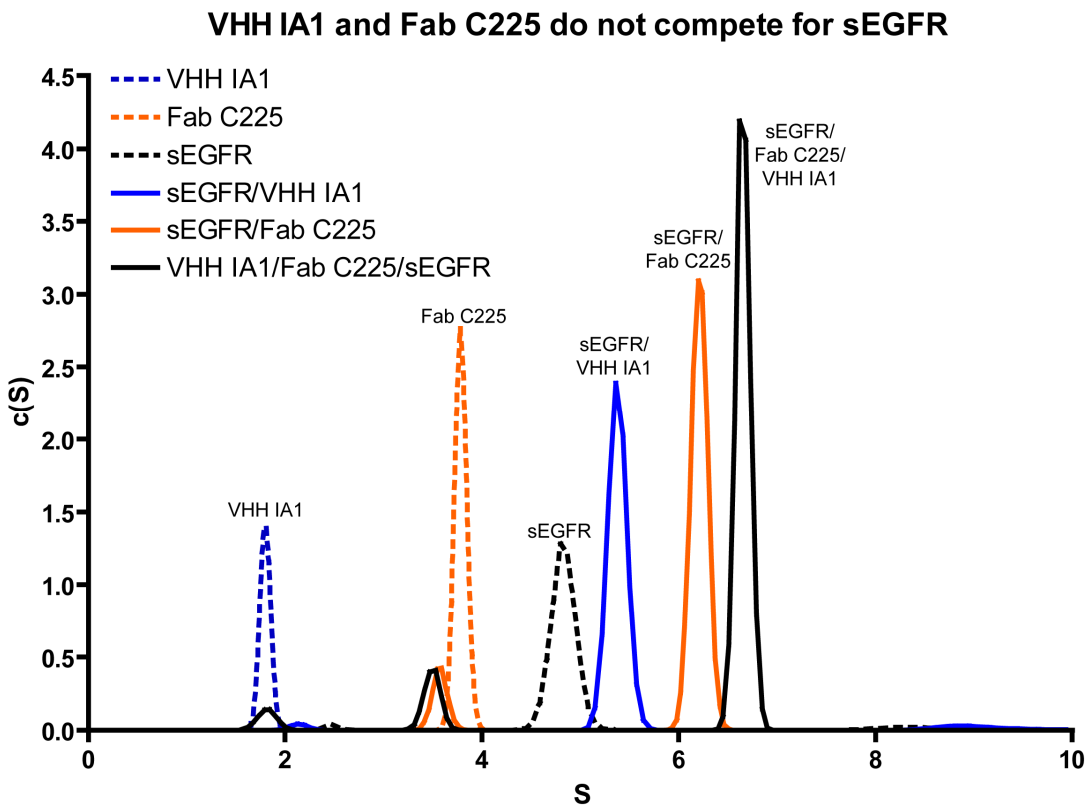

Figure 52. VHH IA1 does not compete with Fab C225 for binding to EGFR.

Sedimentation velocity c(S) species analysis of 5 μM VHH IA1 (dashed blue), 5 μM Fab C225 (dashed orange), 5 μM sEGFR (dashed black), or stoichiometric mixtures of these proteins. VHH IA1 and Fab C225 individually associate with sEGFR to form discrete sedimenting species (solid blue, solid orange, respectively). A mixture of sEGFR, VHH IA1, and Fab C225 (solid black) produces a single faster sedimenting species consistent with a ternary complex. VHH IA1 and Fab C225 can simultaneously bind sEGFR, and thus have distinct epitopes on the receptor surface.

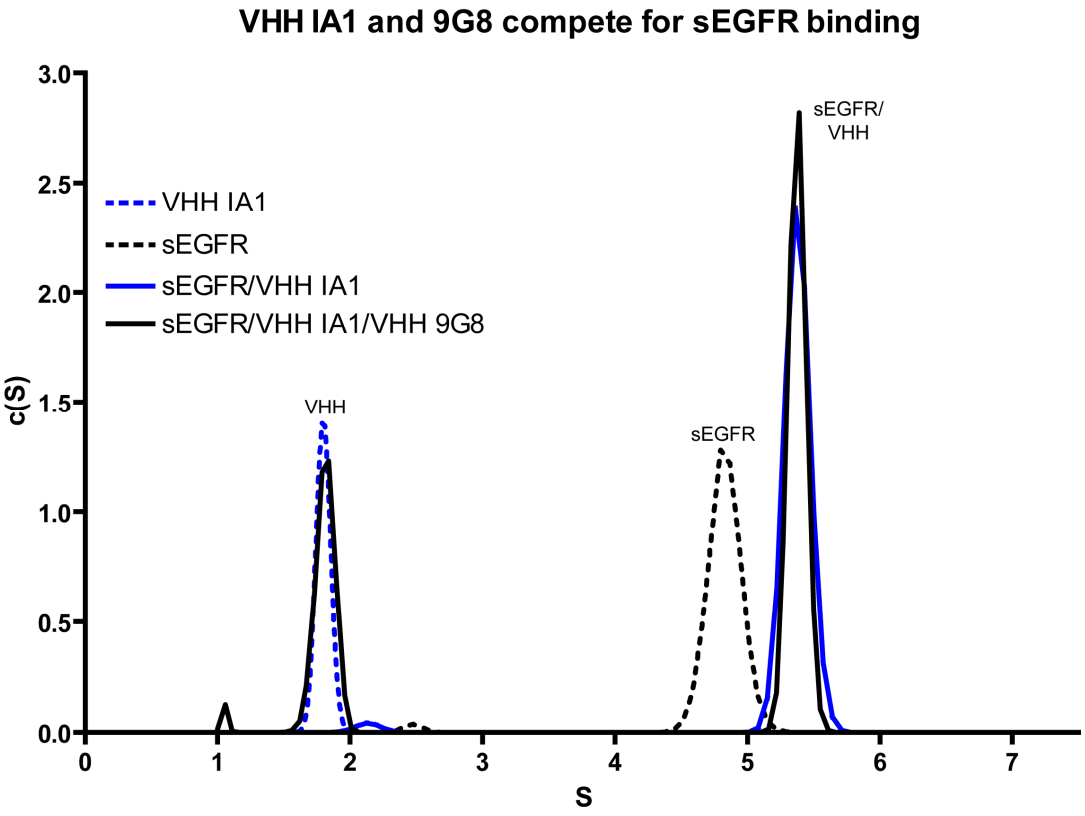

Figure 53. VHHs IA1 and 9G8 compete for binding to sEGFR.

Sedimentation velocity c(S) species analysis of 5 μM VHH IA1 (dashed blue), 5 μM sEGFR (dashed black), or stoichiometric mixtures of sEGFR and VHH IA1 (blue), or sEGFR, VHH IA1, and VHH 9G8 (black). A mixture of sEGFR with VHH IA1 and VHH 9G8 results in two sedimenting species; one corresponding to sEGFR bound to a single VHH, and one corresponding to VHH alone. Thus, VHHs IA1 and 9G8 are not able to simultaneously bind sEGFR, suggesting that these antibodies have similar epitopes.

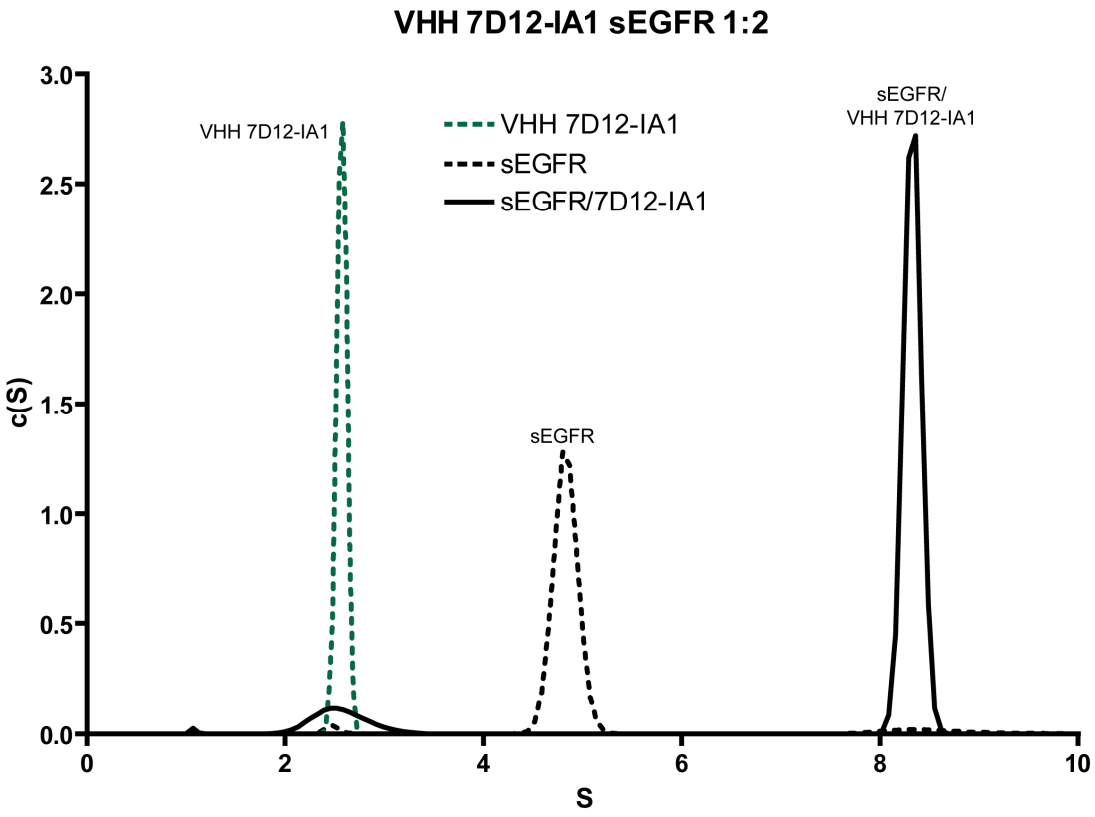

Figure 54. Bivalent VHH 7D12-IA1 binds to sEGFR with 1:2 stoichiometry.

Sedimentation velocity c(S) species analysis of 5 μM of the bivalent, bispecific VHH 7D12-IA1 (dashed teal), 5 μM sEGFR (dashed black), or a stoichiometric mixture of sEGFR and VHH 7D12-IA1 (black). A mixture of sEGFR and VHH 7D12-IA1 results in a single major sedimenting species that corresponds well with a 1:2 complex of VHH 7D12-IAI and sEGFR. There is no evidence of 1:1 complex. Formation of a complex where a distinct receptor molecule binds to each VHH of the biparatopic 7D12-IA1 suggests that simultaneous binding of both VHH modules to a single sEGFR is disfavored, at least in this assay.

distinct and spatially separate epitope from these antibody fragments. By contrast, VHHs IA1 and 9G8 both compete with mAb 425 [205], and are therefore expected to recognize similar epitopes on the EGFR extracellular region. SV AUC is consistent with these expectations, revealing that VHHs IA1 and 9G8 compete for binding to sEGFR (Fig. 53) and suggesting that overlapping epitopes or steric effects prevent simultaneous binding.

Additionally, we have characterized by SV AUC the binding to sEGFR of a bispecific VHH construct comprising concatenated 7D12 and IA1 VHH domains separated by a 10 amino acid glycine/serine linker (VHH 7D12-IA1) [205]. We find that VHH 7D12-IA1 binds to sEGFR with 1:2 (VHH:EGFR) stoichiometry, and does not form a 1:1 species (Fig. 54). If VHH 7D12-IA1 were capable of bivalent sEGFR binding, one would expect a 1:1 species to predominate. The absence of a 1:1 species in this experiment suggests that simultaneous binding of each paratope to sEGFR is compromised in the bispecific format, favoring formation of the 1:2 species. This further implies that VHH IA1 and 7D12 bind to nonadjacent epitopes.

3.3.3 The structure of an unliganded VHH

We have determined the X-ray crystal structure of one of the inhibitory VHHs in our panel, VHH IA1, in the absence of EGFR to a resolution of 1.55 Å (Figs. 55 & 56,Table 9). The framework region of VHH IA1 (i.e., residues outside of the CDRs defined by Chothia [26]) adopts a canonical VH Ig fold (Fig. 57), which aligns to Cab-Lys3 (a well studied VHH [208]) with RMSD ≈ 0.4 Å and to the VH domain of C225 (a conventional IgG1 antibody [133]) with RMSD ≈ 1.0 Å.

VHH IA1 also incorporates several features typical of VHHs but uncommon for conventional VH domains. The CDRs of VHHs are frequently longer than those of conventional VH domains; CDR3 of the VHHs in this study are 4 – 10 amino acids longer than the same region in cetuximab VH (Fig. 58). In VHH IA1, part of CDR3 projects outward from the Ig domain (Fig. 56A), which occurs in other VHH structures (Fig. 56B), but not in cetuximab VH (Fig. 56C). VHH CDR3 also packs against a portion of the VH/VL interface surface (Fig. 56) [189]. These CDR differences increase the area and potential variation of the antigen recognition surface.

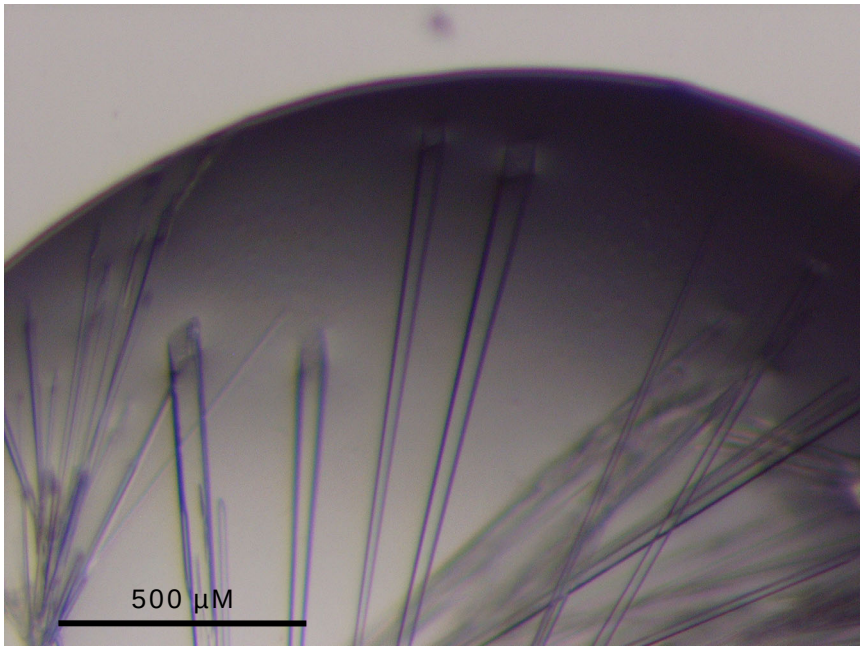

Figure 55. Crystals of VHH IA1 alone.

Representative crystals of the VHH IA1 are shown. Crystals formed with maximum dimensions of 100 × 100 × 2000 µm in 30% PEG8000, 0.2 M $(NH_4)_2SO_4$, 0.1 M MES, pH 6.0.

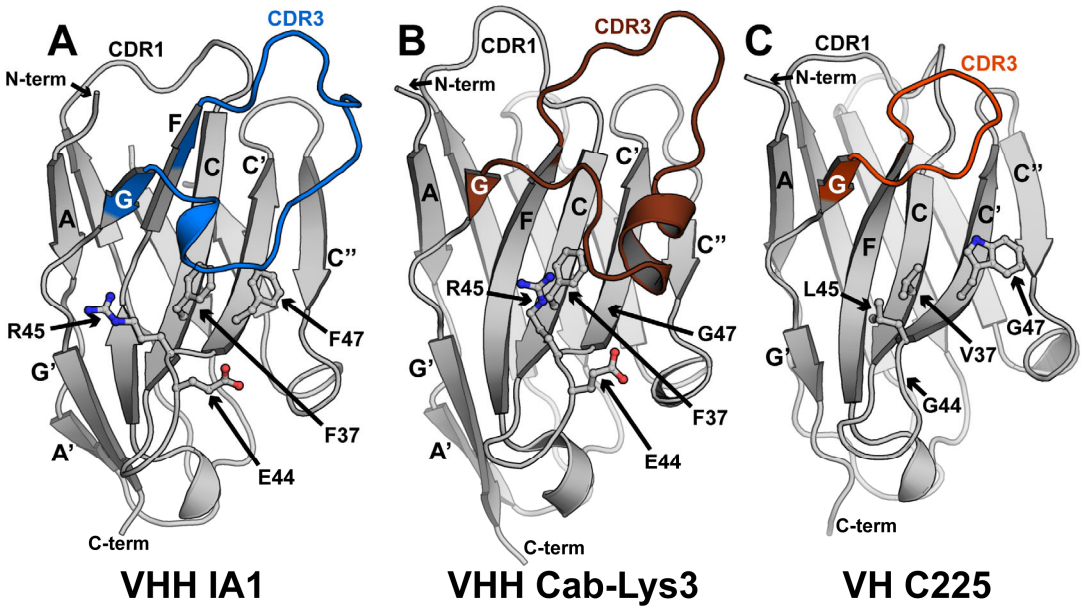

Figure 56. The structure of VHH IA1 alone.

Cartoons are shown of VHH IA1 in its unbound state (A), VHH Cab-Lys3 (B; from PDB: 1JTP), and the VH domain from antibody C225 (C; from PDB: 1YY9) are shown. Both VHH IA1 and Cab-Lys3 posses a relatively long CDR3 (colored), compared to the conventional CDR3 of VH C225. In both (A) and (B), part of CDR3 folds back against the Ig domain, occluding the upper portion of the VH/VL interface region. Additionally, several amino acids in the interface region are typically altered in the HCAbs, and are shown as sticks. The changes to positions 37 and 47 appear to accommodate packing of CDR3, while changes to 44 and 45 result in a more hydrophilic surface.

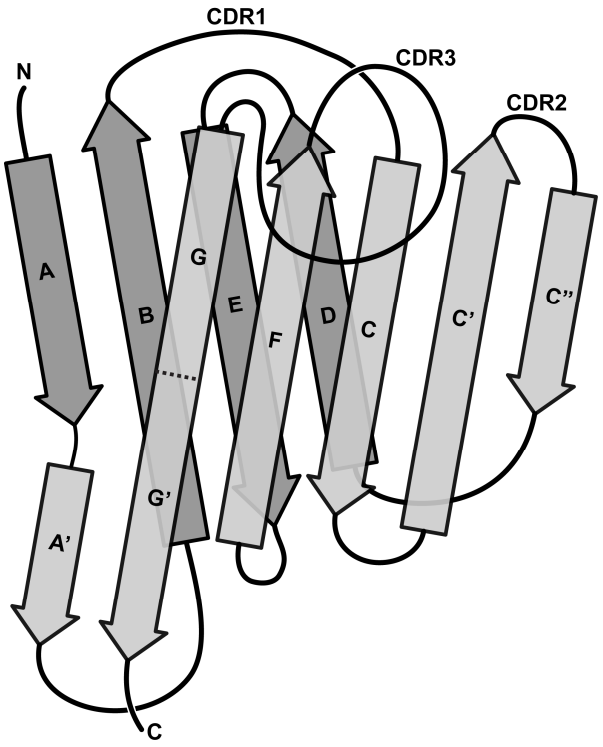

Figure 57. VHH secondary structure.

A cartoon representation is shown of the VHH Ig domain secondary structure. The Ig β-sandwich fold comprises two β-sheets (dark and light gray). The foreground β-sheet (light gray) contains the surface that packs against the VL domain in a conventional VH (including strands G, F, C, C', and C''). VHHs incorporate alterations to amino acids in C, C', and the C/C' loop that improve solubility and packing in the absence of a VL partner. VHHs also typically possess relatively long CDR1 and CDR3 regions. The elongated CDR3 often packs against part of the VH/VL interface surface. Standardized strand nomenclature is shown [20].

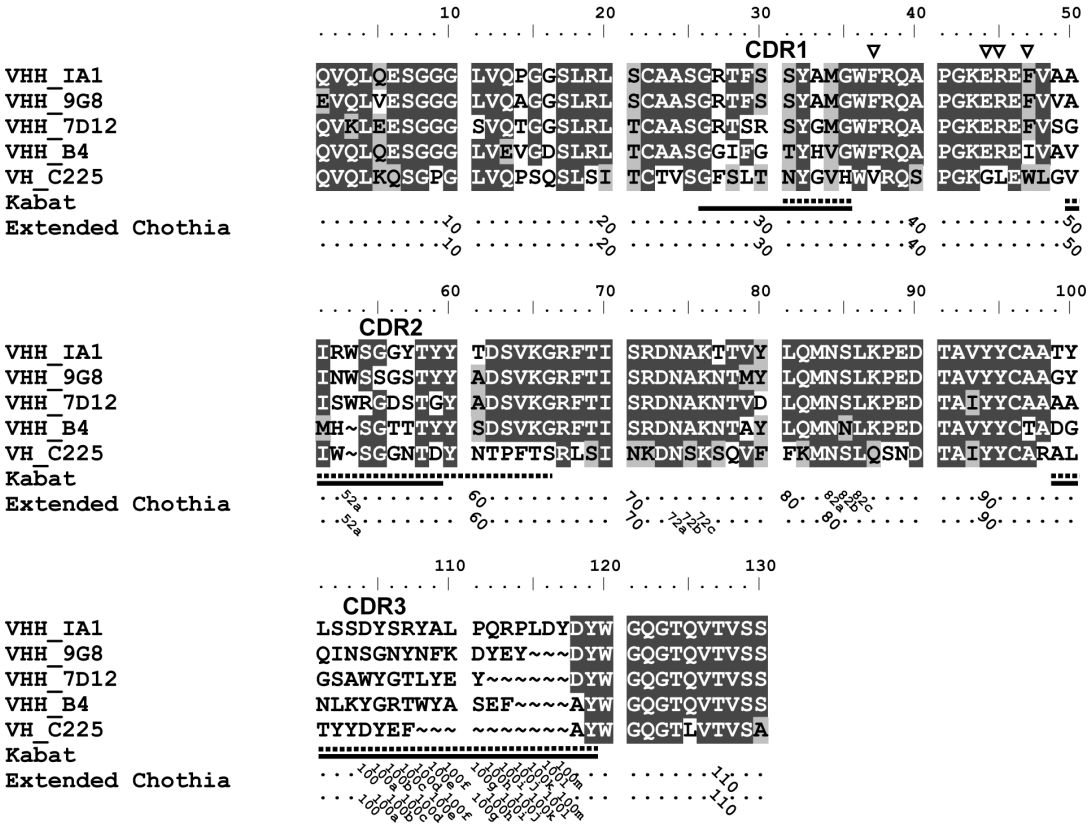

Figure 58. Alignment of VHH and VH sequences.

A multiple sequence alignment of the anti-EGFR VHH constructs used in this study, along with the VH region of mAb C225. Standardized Kabat numbering [7] and a modified Chothia numbering scheme [15] (both determined by the Abnum algorithm [6, 15]) are shown. CDRs defined by Kabat (dashed lines) and by Chothia (which correspond closely with observed loop regions; solid lines) are shown [7, 10]. Triangles indicate conserved amino acids in the VH/VL interface region that are substituted in VHHs (see Fig. 56). Residues in CDR3 have been left-justified, rather than aligned according to sequence similarity, to facilitate standardized antibody numbering. Darker shading indicates greater sequence similarity.

VHHs also contain point mutations to hydrophobic surface amino acids (typically V37F, G44E, L45R, and W47G/S/L/F) that form the VH/VL interface in conventional antibodies [189, 204, 210]. These changes are thought to improve solubility (in the case of G44E and L45R) and to stabilize packing in the absence of a VL partner [204]. Indeed, E44 and R45 in VHH IA1 are clearly more polar than the homologous residues of cetixumab VH (G44 and L45), while VHH IA1 F73 and F47 pack against CDR3 (Fig. 56A & C). Similar point mutations occur in the other VHHs in this study, and are apparent in a multiple sequence alignment against cetuximab VH in Fig. 58.

3.3.4 The VHH 7D12 epitope on domain III

To understand the molecular details of the interaction between VHH 7D12 and EGFR, we determined the crystal structure of VHH 7D12 bound to sEGFRd3. The VHH 7D12/sEGFRd3 complex formed crystals in two distinct crystallization conditions (Fig. 59). Hexagonal rod shaped crystals (100 × 50 × 50 μm) formed at pH 6.0 and diffracted to a resolution of 2.85 Å. Smaller angular crystals (100 × 100 × 200 μm) formed at a nominal reservoir pH of 3.5 and diffracted to 2.66 Å (Table 9). The low pH of the latter crystal form did not introduce any apparent disruptions in protein conformation or complex formation. The pH 6.0 and pH 3.5 crystal forms of the complex superimpose with RMSD < 1.0 Å, and the contacts stabilizing the interaction are identical in both structures.

The framework region of VHH 7D12 reveals a typical VH Ig fold, and aligns to the framework region of unliganded VHH IA1 with RMSD ≈ 0.44 Å, and to the same region of Cab-Lys3 with RMSD ≈ 0.52 Å. EGFR domain III (amino acids 311 – 514) does not exhibit any major changes in main chain position compared to previous reported crystal structures of tethered EGFR or sEGFRd3, and aligns with an RMSD < 1.0 Å (0.4 Å against 1YY9; 0.6 Å against 1NQL; 0.7 Å against 3B2U). The 4 amine terminal amino acids of sEGFRd3 are identical to the first four residues of mature EGFR, and are included in the sEGFRd3 construct to ensure secretion and signal sequence cleavage. This amine terminal region adopts an extended conformation in both

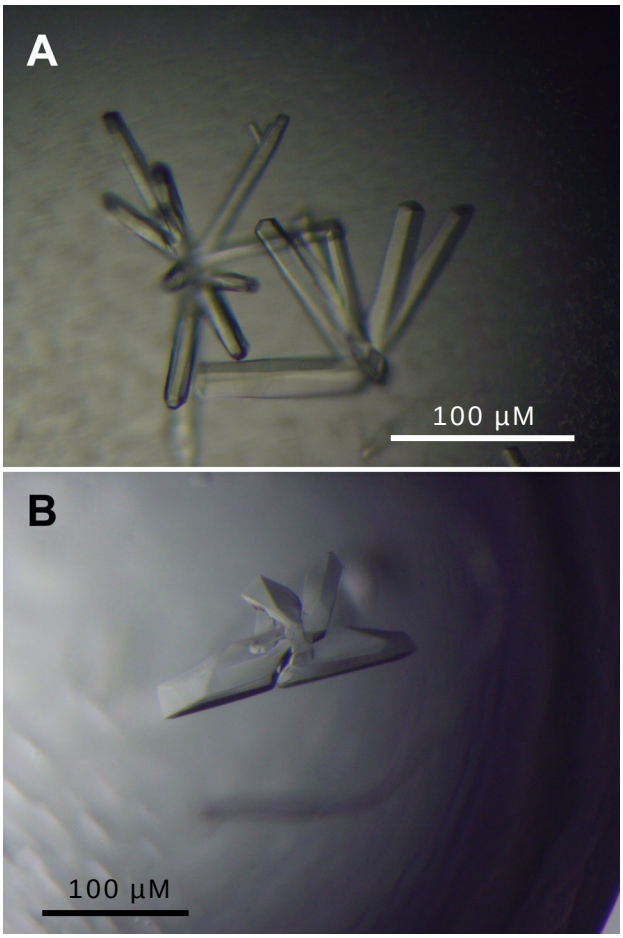

Figure 59. Crystals of VHH 7D12/sEGFRd3.

Representative crystals of the (A) pH 6.0 and (B) pH 3.5 VHH 7D12/sEGFRd3 complexes. (A) 100 × 50 × 50 µm grew in 22.5% PEG3350, 50 mM KI, 0.1M MES, pH 6.0. (B) 100 × 100 × 200 µm crystals formed in 22.5% PEG3350, 0.1M sodium citrate, pH 3.5.

Table 9. VHH/EGFR Crystallographic Statistics

Data Collection Statistics[a]	VHH IA1	VHH 7D12/ sEGFRd3 (pH 6.0)	VHH 7D12/ sEGFRd3 (pH 3.5)	VHH IA1/ Fab C225/ sEGFR	VHH 9G8/ Fab C225/ sEGFR
Space group	$P4_3$	$P6_122$	$P2_12_12_1$	$P2_1$	$P2_1$
Unique cell dimensions	a, b = 52.7 Å, c = 62.5 Å	a, b = 148.0 Å, c = 82.5 Å	a = 78.7 Å, b = 147.2 Å, c = 254.8 Å	a = 66.2 Å, b = 96.3 Å c =128.1 Å, β= 100.7°	a = 66.4 Å, b = 95.8 Å c =129.5 Å, β= 99.9°
X-ray source	APS 23-ID-B	APS 23-ID-B	APS 23-ID-B	CHESS F1	APS 23-ID-B
Wavelength (Å)	0.98	1.0332	0.98	0.918	0.98
Resolution limit (Å)	1.55	2.90	2.65	3.05	2.80
Observed/Unique	107,430/24,694	136,787/22,976	568,198/83,116	69,621/29,807	138,909/37,840
Redundancy	4.4 fold	6.0 fold	6.8 fold	2.3 fold	1.9 fold
Completeness (%)	99.2 (94.5)	97.0 (73.4)	96.6 (75.4)	98.8 (99.6)	99.0 (92.4)
R_{sym}[b]	0.037 (0.283)	0.117 (0.574)	0.108 (0.614)	0.099 (0.463)	0.096 (0.413)
$\langle I/\sigma \rangle$	47.5 (4.4)	21.9 (2.0)	20.7 (2.0)	10.7 (2.2)	14.8 (2.3)
Refinement Statistics					
Resolution limits (Å)	62.7 – 1.55	128 – 2.85	49.53 – 2.66	125.00-3.05	47.92 – 2.82
No. reflections/No. test set	23,430/1,253	12,067/617	78,793/4,233	28,282/1,507	35,947/1,889
R factor (R_{free})[c]	0.20 (0.23)	0.24 (0.31)	0.23 (0.28)	0.23 (0.31)	0.24 (0.30)
Model	VHH IA1, aa 1-129 1 sulfate ion 87 water molecules	VHH 7D12, aa 1-122 sEGFRd3, aa 307-511 1 MES 1 iodide ion 9 water molecules	6x VHH 7D12[d], aa 1-124 6x sEGFRd3[d], aa 307-503 33 water molecules	VHH IA1 sEGFR FabC225 21 water molecules	VHH 9G8 sEGFR FabC225 8 water molecules
Total number of atoms	1,074	2,455	14,727	7,620	7,759
RMSD bond lengths (Å)	0.007	0.015	0.017	0.013	0.015
RMSD bond angles (°)	1.111	1.905	1.949	1.659	2.049

a Numbers in parenthesis refer to last resolution shell.

b $R_{sym} = |I_h - \langle I_h \rangle| / I_h$, where $\langle I_h \rangle$ = average intensity over symmetry equivalent measurements.

c R factor = $|F_o - F_c| / F_o$, where summation is over data used in the refinement; R_{free} includes only 5% of the data excluded from the refinement.

d Number of missing amino acids varies by chain.

crystallization conditions and forms a strand/strand interaction with VHH 7D12 β-strand C'' from a crystallographically related copy of the complex, stabilizing crystal packing.

VHH 7D12 binds the "upper" surface of domain III, near the mAb C225 epitope and the ligand binding surface (Fig. 60). This interaction involves electrostatic and hydrogen bonds, as well as hydrophobic contacts between antibody and antigen (Table 10). VHH CDR1 and CDR3 contact the first two turns of the domain III β-helix (amino acids ≈ 310 – 375); CDR2 makes no contacts with sEGFR. The complex buries an average of ≈ 700 $Å^2$ on each surface, 48% of which is hydrophobic surface area (Table 10). The interface has a shape complementarity parameter [83] of 0.68, which is typical for antibody/antigen interactions and higher than that of other VHH/antigen interactions (Table 10).

As shown in Fig. 61, a cluster of polar and electrostatic interactions occur in the center of the VHH 7D12 epitope. EGFR D355 forms a salt bridge with R30 of VHH 7D12 (the only polar interaction to an amino acid of CDR1). EGFR R353 makes electrostatic contacts with both 7D12 D101 and E100f, as well as hydrogen bonds to the backbone of CDR3. Additional hydrogen bonds link the CDR3 main chain carbonyls to the side chain of EGFR Q384. These central polar interactions are flanked by two sets of apolar contacts. The aliphatic portion of 7D12 R30 packs against EGFR F357. Additionally, the phenol ring of 7D12 Y100e packs against the saccharide ring of an N-acetyl glucoseamine (NAG4201, linked to EGFR N420). NAG4201 has rotated ≈ 90° with respect to its typical position in EGFR domain III [2, 5, 17] to contact Y109.

3.3.5 Comparison of EGF, FabC225, and VHH 7D12 binding

As VHH 7D12 binds to domain III near the binding surfaces of FabC225 and ligand, we have compared the molecular details of these interactions (Fig. 62). EGF binding to domain III is largely stabilized by three hydrophobic interactions: L47 projects into a hydrophobic pocket on the surface of domain III, Y44 packs against a hydrophobic surface, and the phenol ring of Y13 packs flat against that of F357 (Fig. 62A). A salt bridge to EGFR D355 and polar contacts between the EGF main chain and Q387 further stabilize binding.

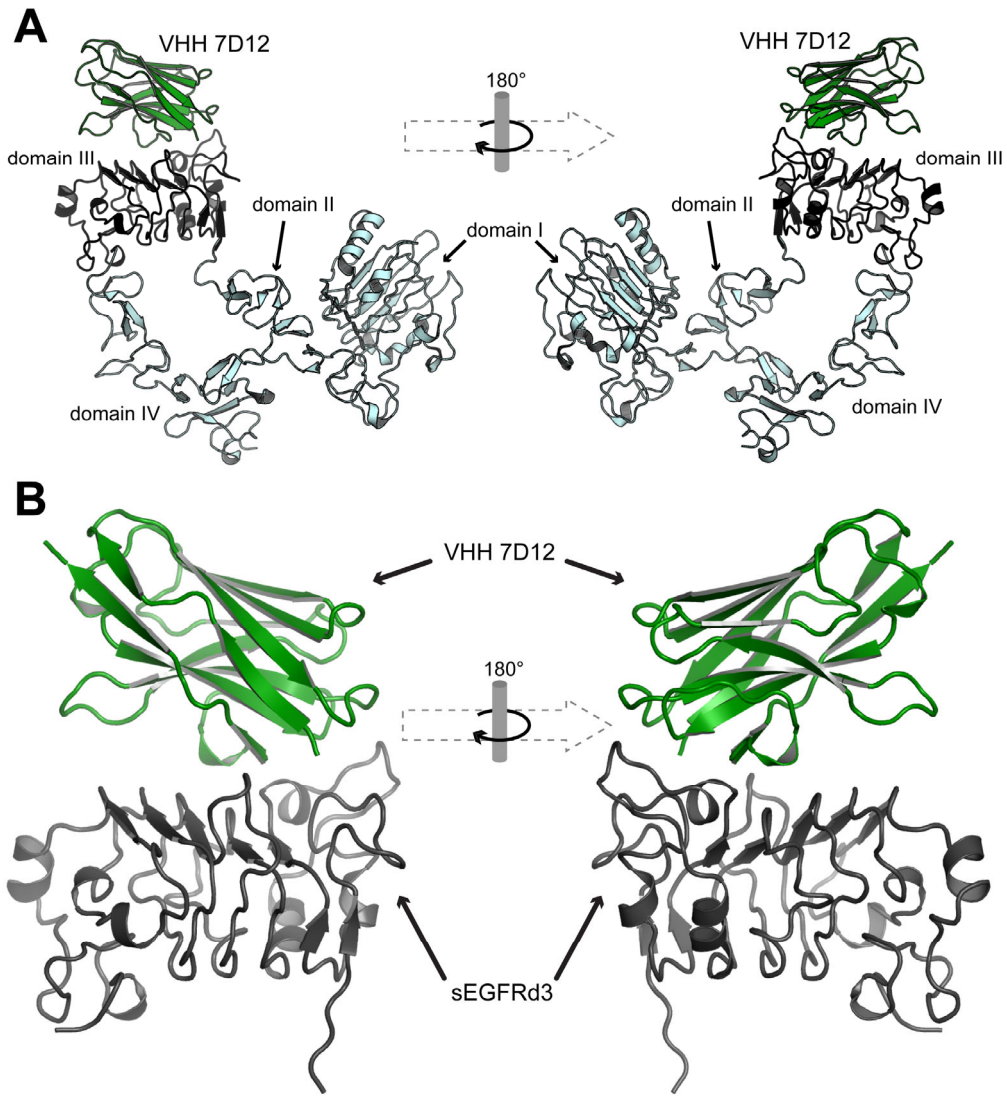

Figure 60. The crystal structure of the VHH 7D12/sEGFRd3 complex.

(A) Cartoons of VHH 7D12 (green) bound to sEGFRd3 (gray) are modeled on the structure of tethered sEGFR (light blue; PDB: 1NQL). Two orientations are shown, related by a 180° rotation. (B) A closer view of the VHH 7D12/sEGFRd3 complex. The VHH 7D12 epitope overlaps that of FabC225 and the ligand binding site (Fig. 64). Binding of VHH 7D12 is stabilized by interactions from CDR1 and 3 (Figs. 61 & 63). The N-terminal LEEK sequence projects away from domain III (downward, in this orientation), and adopts an extended conformation.

Table 10. Antibody/antigen interaction details.

	VHH/EGFR				
	VHH 7D12/ sEGFRd3 (pH 3.5)	VHH 7D12/ sEGFRd3 (pH 6)	VHH IA1/sEGFR	VHH 9G8/sEGFR	
Avg. Excl. Area (Å^2)	692	705	755 (647)	683 (589)	
Apolar fraction	0.482	0.488	0.393 (0.343)	0.385 (0.374)	
H. Bonds	6	6	16	8	
E. Contacts	3	3	4	5	
S. C.	0.716	0.677	0.644 (0.655)	0.680 (0.670)	
	Fab/EGFR			Ligand/EGFR	
	Cetuximab/ sEGFR	Matuzumab/ sEGFR	IMC-11F8/ sEGFRd3	EGF/ sEGFR	TGFα/ sEGFR
Avg. Excl. Area (Å^2)	898	745	914	1600 (861)	1438 (746)
Apolar fraction	0.492	0.674	0.481	0.581 (0.597)	0.579 (0.648)
S. C.	0.752	0.655	0.645	0.693 (0.684)	0.760 (0.692)
	VHH/RNaseA	VHH/β-amylase			VHH/Bacterial Receptor
	CAB-RN05/ RNaseA (2P45)	CAB-AMD9/ β-amylase (1KXQ)	(1KXT)	CAB-10/ β-amylase (1KXV)	(2BSE)
Avg. Excl. Area (Å^2)	582	1120	828	848	577
Apolar fraction	0.564	0.509	0.488	0.538	0.485
	VHH/lysozyme				
	CAB-Lys3/ lysozyme (1JTO)	CAB-Lys3/ lysozyme (1JTP)	1DL19/ lysozyme (1RI8)	CAB-Lys2/ lysozyme (1RJC)	D2-L24/ lysozyme (1ZVH)
Avg. Excl. Area (Å^2)	848	813	708	813	572
Apolar fraction	0.582	0.604	0.397	0.432	0.518

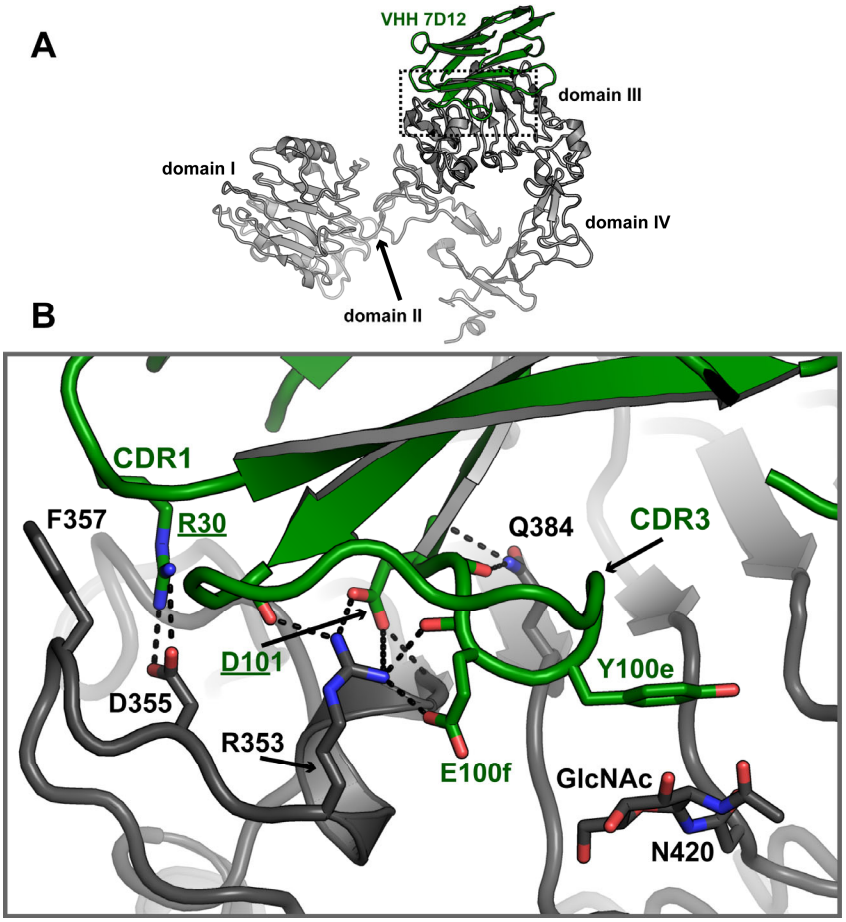

Figure 61. Detail of the interaction between VHH 7D12 and domain III of EGFR.

(A) A cartoon of VHH 7D12 (green) bound to EGFR domain III (gray) is shown, modeled over the structure of tethered sEGFR (PDB: 1NQL). The orientation is similar to that shown in the left side of Fig. 60A. (B) A detailed view of the region boxed in (A); VHH strands C' and C'' have been hidden for clarity. VHH 7D12 binds to its epitope via CDRs 1 and 3. A central network of interactions to EGFR R353 stabilizes binding, along with additional charged and polar interactions, and apolar contacts at the periphery of the epitope. Dashed lines indicate hydrogen bonds (≤ 3.5 Å) or electrostatic interactions (≤ 4.5 Å). Underlined VHH amino acids (Kabat numbering [7]) were altered by site directed mutagenesis to validate the interactions observed in the crystal structure.

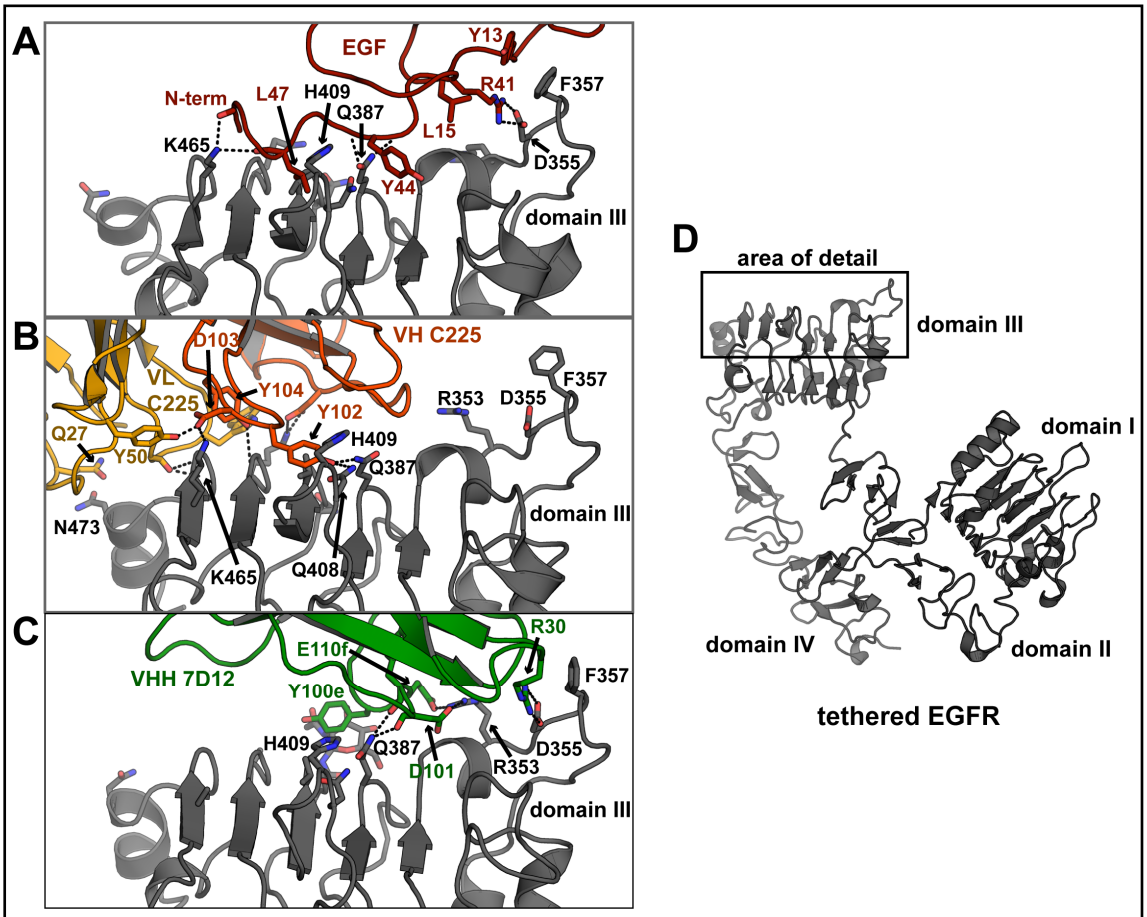

Figure 62. Comparison of the interactions between EGF, FabC225, and VHH 7D12 with EGFR domain 3.

Cartoons of EGFR domain III in complex with (A) EGF (PDB: 1IVO), (B) FabC225 (1YY9), and (C) VHH 7D12 are shown. EGF binding is stabilized by apolar contacts to EGFR F357 and burial of EGF L47 in a hydrophobic pocket; interactions between EGFR Q387 and K456 and the EGF main chain; and a salt bridge to EGFR D355. FabC225 binding involves similar hydrophobic burial of heavy chain Y102, and polar contacts to EGFR Q387. VHH 7D12 also makes polar contacts to EGFR Q387, and makes apolar contacts to F357. FabC225 and VHH 7D12 interact with different EGFR motifs involved in EGF binding. (The notable exception is polar interaction with Q387.) (D) The region of EGFR shown in (A – C) is highlighted on tethered sEGFR (1NQL).

The interaction of FabC225 to its epitope utilizes some of the features on domain III as EGF. Heavy chain amino acid Y102 occupies the domain III hydrophobic pocket into which EGF L47 projects, and makes hydrogen bonds to Q384 (Fig. 62B). Additional electrostatic contacts are made by the heavy chain and extensive hydrogen bonds by the light chain also contribute to the interaction, but these occur mostly C-terminal to the EGF binding.

VHH 7D12, unlike FabC225 and EGF, is not stabilized by the insertion of an apolar side chain into the domain III hydrophobic pocket. Instead, the polar and charge contacts form the core of the epitope (Fig. 62C). VHH 7D12 does contact residues on domain III that are critical for EGF binding (VHH main chain contacts to Q387; a salt bridge between VHH R30 and EGFR D355; van der Waals packing between VHH R30 and EGF F357; Fig. 63), but, in contrast to FabC225, these occur on the N-terminal side of domain III. These differences are apparent when comparing the binding footprint and epitopes of EGF, FabC225, and VHH 7D12 (Fig. 64).

Alteration of EGFR amino acids F357, D355, and Q384 has been shown to disrupt EGF binding (Table 7, Figs. 43 & 44), but not binding of mAb C225. We find that an sEGFR variant with alterations to D355 and F357 binds immobilized VHH 7D12 with ≈ 5 fold weaker affinity (Table 11). This supports the structural observation that VHH 7D12 more closely occludes this portion of the ligand binding surface than FabC225.

To further validate the interactions observed in the crystal structure, two altered forms of VHH 7D12 were generated incorporating mutations to amino acids R30 and D101 (VHH 7D12(R30A) and VHH 7D12(D101A)), which make stabilizing contacts in the complex structure. Disruption of R30 abolishes the interaction with sEGFR, while mutation of D101 weakens binding to receptor by three-fold (Table 11).

3.3.6 VHH IA1 binds to a novel epitope on domain III, adjacent to domain II

We next sought to understand the structural basis of EGFR binding and inhibition by VHH IA1, which competes with the inhibitory antibody matuzumab [205]. While VHH IA1 and sEGFR stably associate in solution (Fig. 50), this complex did not yield crystals under any of the

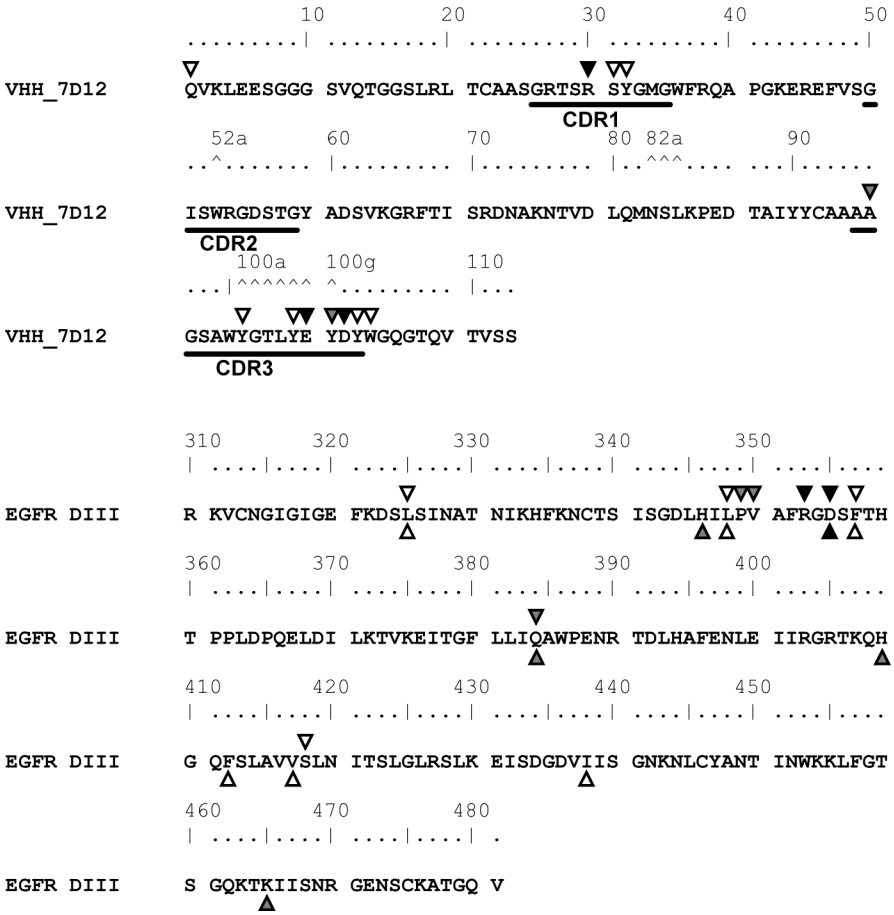

Figure 63. Details of the interaction between VHH 7D12 and EGFR domain III.

The sequences of VHH 7D12 (Kabat numbering [7]) and domain III of EGFR are shown. Amino acids marked by black triangles above the sequence mediate electrostatic contacts in the VHH/sEGFRd3 crystal structure; gray triangles indicate hydrogen bonds; white triangles denote van der Waals contacts. Triangles below the sequence indicate these types of interactions in the EGF/EGFR crystal structure (PDB: 1IVO). CDRs (Chothia definition [26]) are underlined. VHH 7D12 makes apolar contacts with sEGFR L325, L348, and F357; a salt bridge with D355; and a hydrogen bond with Q384. Several of these residues (notably D355 and Q384) also play a key role in stabilizing the interaction between sEGFR domain III and ligands EGF and TGFα [1, 2].

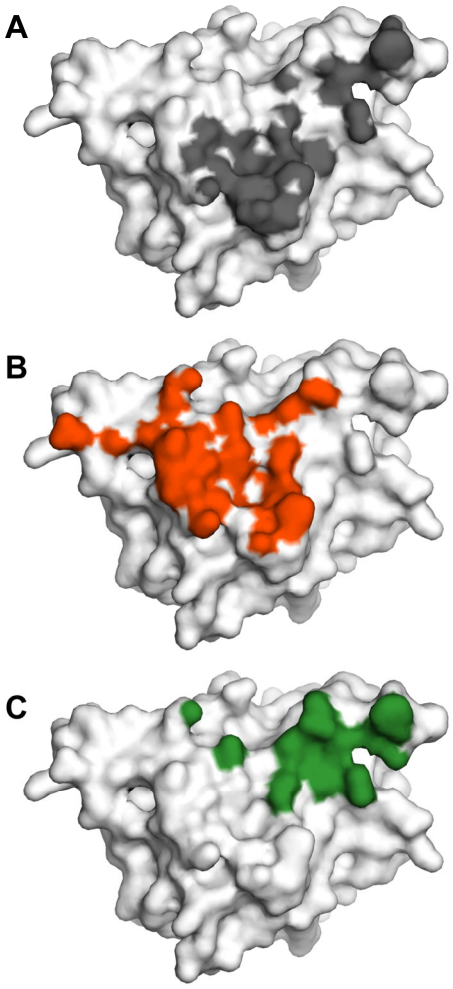

Figure 64. Comparison of the binding sites of EGF, FabC225, and VHH 7D12 on EGFR domain 3.

The EGF binding footprint (gray, A), FabC225 epitope (orange, B), and VHH 7D12 epitope (green, C) are shown on a surface representation of EGFR domain III, which is rotated 90° about a horizontal axes compared to its orientation in Figure 62D. The highlighted areas indicate surface exposed atoms within 4 Å of bound ligand or antibody. The VHH 7D12 epitope overlaps with both the ligand binding site and the C225 epitope.

Table 11. Equilibrium binding constants of sEGFR or variants to immobilized VHHs with epitope/paratope alterations.

	sEGFR
VHH 7D12(R30A)[1]	no binding
VHH 7D12(D101A)[1]	656 ± 22 nM
VHH IA1(R27A)[1]	> 5.0 μM
VHH IA1(D101A)[1]	> 3.5 μM
	sEGFR(D355T/F357A)
VHH 7D12	928 ± 106 nM
VHH IA1	342 ± 6.5 nM
VHH 9G8	303.5 ± 12 nM

1. Kabat numbering [7] is used for VHH residues.

condition screened. We reasoned that crystallization may be hindered by a lack of surface area suitable for crystal packing on the extensively glycosylated sEGFR, as well as by EGFR conformational flexibility. By contrast, the complex between Fab C225 and sEGFR possesses a larger fraction of un-glycosylated surface area, restricts the conformational freedom of the EGFR extracellular region, and forms well diffracting crystals [5]. Thus, we utilized the stable ternary complex formed by VHH IA1, Fab C225, and sEGFR (Fig. 52) as a target for crystallization.

Crystals of the IA1/C225/sEGFR complex (75 × 75 × 1000 μm; Fig. 65) diffracted to 3.05 Å (Table 9). A single copy of the ternary complex occupies the asymmetric unit, and all three protein components are observed (Fig. 66). EGFR adopts a tethered conformation [17], and FabC225 makes identical contacts with EGFR domain III as in the reported FabC225/sEGFR complex structure [5] (RMSD ≈ 0.5 Å). An elbow angle [211] of 131° is observed between the constant and variable regions of FabC225, which is similar to the angle found in the structure of FabC225 alone (PDB: 1YY8; 135°) and smaller than that in the FabC225/sEGFR complex (PDB: 1YY9; 150°) (Fig. 67). Crystal packing is stabilized predominantly by contacts between Fab and domain III. FabC225 and sEGFR domain III are well ordered, as are portions of domains II, domain IV, and the VHH IA1 paratope region. However, EGFR domain I, the N-terminal portion of domain II, and the distal portion of VHH IA1, which makes crystal packing contacts with EGFR domain I, are poorly resolved and these regions are absent or present as backbone atoms only in the refined model (Table 9).

VHH IA1 binds sEGFR in the cleft formed between domains II and III (Fig. 68). However, ordered contacts are only observed between VHH IA1 and domain III of sEGFR, suggesting that VHH binding is primarily stabilized by these interactions. An average surface area of 755 $Å^2$ is buried on each partner, only 39 % of which is hydrophobic (Table 10). VHH CDRs 1 and 3 contact sEGFR, as do the first two amino-terminal residues of the VHH (Figs. 69 & 70). CDR2 does not participate in the interaction.

As reflected by the relatively low proportion of buried apolar surface area (34% for the interaction with domain III; Table 10)), the interaction between VHH IA1 and EGFR domain III is

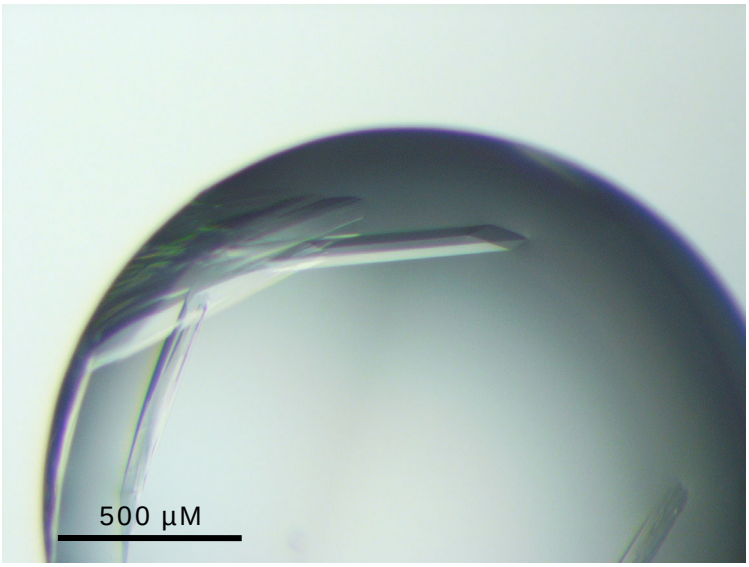

Figure 65. Crystals of the VHH IA1/FabC225/sEGFR complex.

Representative crystals of the VHH IA1/FabC225/sEGFR complex. 75 × 75 × 1000 µm crystals formed in 17.5% P3350, 1.5M NaCl, 5% glycerol, 0.1 M MES, pH 6.5.

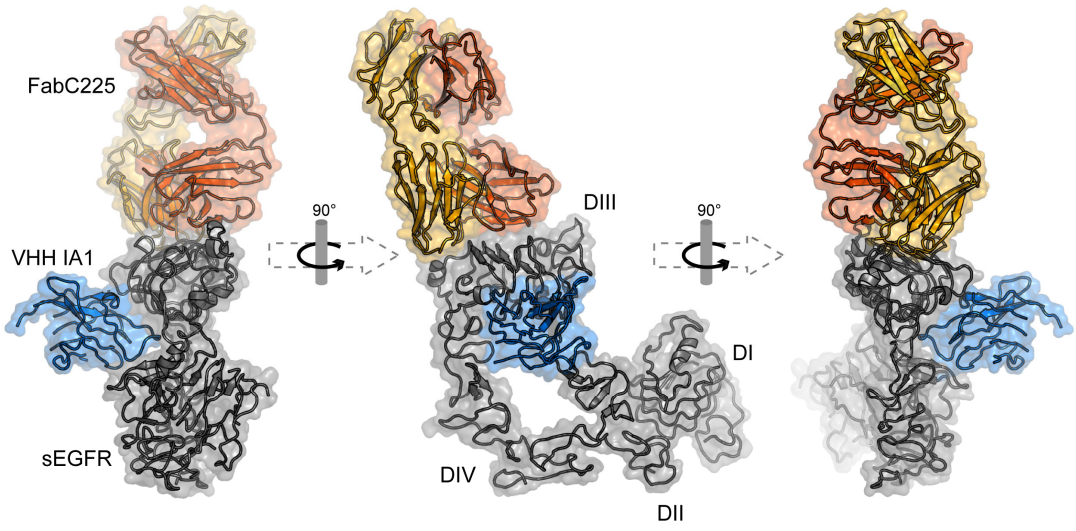

Figure 66. The crystal structure of the VHH IA1/FabC225/sEGFR ternary complex.

A cartoon of the VHH IA1/FabC225/sEGFR ternary complex crystal structure is shown, with a transparent molecular surface. Three orientations are shown, related by 90° rotations. EGFR is colored gray, VHH IA1 is blue, FabC225 light chain is yellow, and FabC225 heavy chain is orange. Part of the C-terminal portion of VHH IA1 (opposite the CDRs) is not observed in the structure. VHH IA1 binds on the "side" of EGFR domain III, adjacent to domain II (appreciated in the left- and rightmost views). The VHH IA1 epitope does not overlap with that of FabC225.

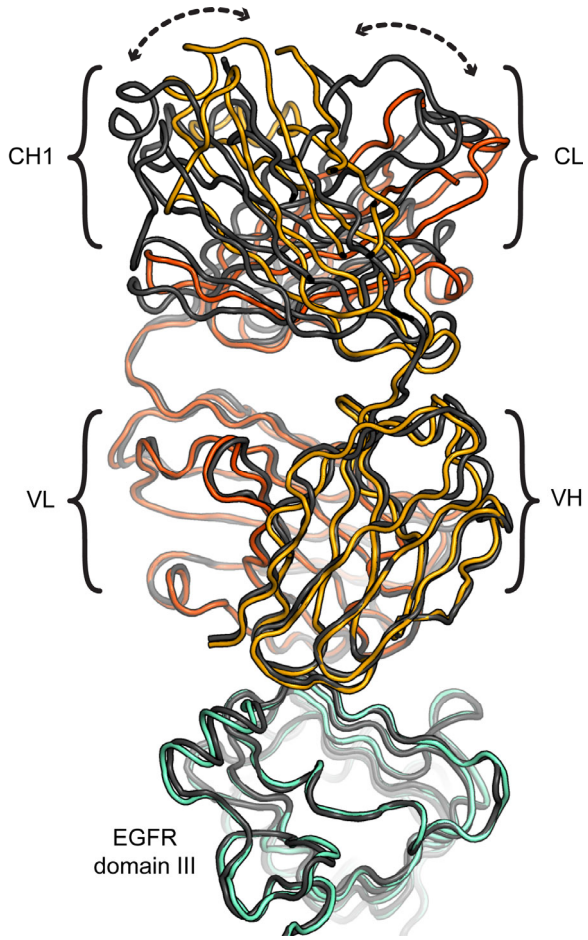

Figure 67. Conformational flexibility in FabC225 is apparent by comparing the FabC225/sEGFR and VHH IA1/FabC225/sEGFR crystal structures.

A backbone ribbons representation of FabC225 (heavy chain: yellow; light chain: orange) and sEGFR domain III (teal) from the VHH IA1/FabC225/sEGFR complex structure is shown aligned to the same regions from the FabC225/sEGFR complex structure (black; PDB ID: 1YY9), with respect to sEGFR domain III. The relative orientations of the constant and variable portions of FabC225 differ: an elbow angle of 131° is observed in the VHH IA1 complex, compared to 150° in the FabC225/sEGFR structure. VHH IA1 is omitted for clarity.

stabilized primarily by polar and electrostatic contacts. In addition to hydrogen bonds between the backbone of the first two amino acids of VHH IA1 and the main chain of domain III (Figs. 69 & 70), VHH binding is stabilized by three clusters of interactions. The first involves an electrostatic contact from VHH R27 and predicted hydrogen bonds from Y102 and Y32 to EGFR E431. A second cluster is formed by a salt bridge between VHH D101 to EGFR R405, likely polar contact between VHH T95 and EGFR R403, and hydrogen bonds from main chain atoms of CDR3 to both R403 and R405. A third, and more decentralized, cluster consists of interactions between CDR3 and residues at the N-terminal end of domain III. These include predicted hydrogen bonds from VHH Y96, S99 and the CDR3 backbone to EGFR E400; polar contacts between VHH S98 and E376, and between S99 and K375; as well as formation of a salt bridge between VHH D100 and R310. We also note that the charged side chains in this region of domain III form a network of salt bridges (R403 to E376 to R301). These contacts occur in all crystals structures of sEGFR in the tethered state, where they may contribute to the orientation of domains II and III. VHH IA1 makes contacts to each of these domain III side chains.

To validate the interactions observed in the VHH IA1 complex crystal structure, two altered forms of VHH IA1 were generated incorporating mutations to amino acids R27 and D101 (VHH IA1(R27A) and VHH IA1(D101A)). Purified sEGFR binds VHH IA1(R27A) with a K_D of > 5 μM and VHH IA1(D101A) with a K_D > 3.5 μM, a decrease in affinity of approximately 20- and 14-fold, respectively (Table 11).

The crystal structure reveals no well ordered interactions between VHH and sEGFR domain II. It is possible that discrete contacts between occur, but are not resolved due to the noted disorder in the distal portion of VHH IA1. We note that the ordered side chain of VHH IA1 Y100d is within 4.2 Å of the β-carbon of the disordered side chain of EGFR M294. This is close enough for van der Waals contacts between Y100d and M249. Even with this side chain not modeled, we calculate that an average surface area of ≈ 100 $Å^2$ is excluded from solvent in the VHH IA1/domain II interaction (Table 10).

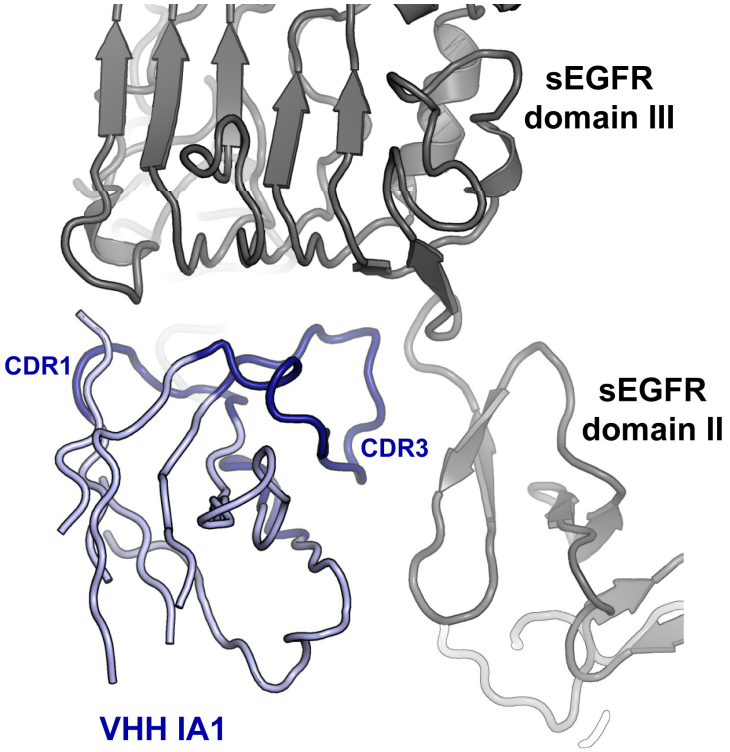

Figure 68. VHH IA1 binds between domains II and III of sEGFR.

Cartoons of VHH IA1 (blue) and sEGFR (gray) from the VHH IA1/FabC225/sEGFR complex structure are shown. Ordered CDR regions indicated are colored darker blue. VHH IA1 lies in a cleft formed by sEGFR domains II and III, though ordered contacts are only observed to domain III. Compare to the position of bound VHH 9G8, Fig. 74.

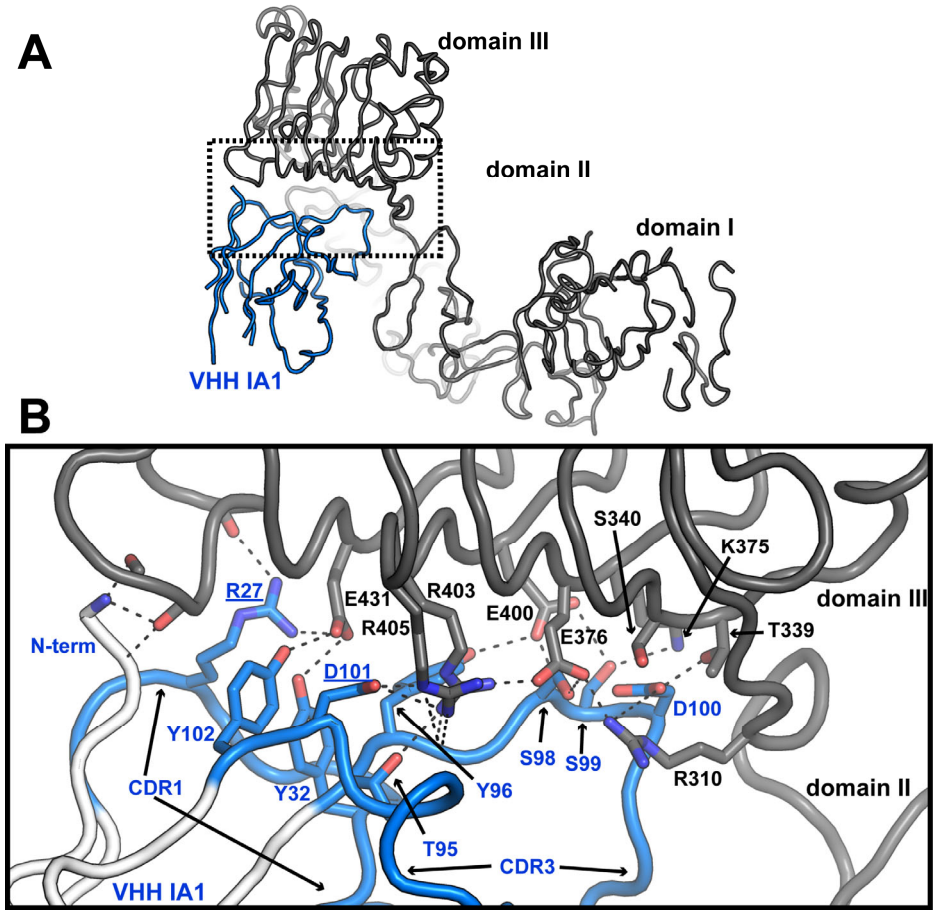

Figure 69. Detail of the interaction between VHH IA1 and domain III of EGFR.

(A) Cartoons of VHH IA1 (blue) bound to EGFR (gray) from the VHH IA1/FabC225/sEGFR complex are shown (FabC225 has been omitted in this view). Domain IV projects into the page behind domains II and III. (B) A detailed view of the boxed region in (A) is shown. VHH framework regions are colored white. VHH IA1 binds to domain III primarily via its N-terminus, and by three clusters of polar interactions made by CDRs 1 and 3. No ordered contacts are observed between VHH IA1 and EGFR domain II. Dashed lines indicate expected hydrogen bonds (≤ 3.5 Å) or electrostatic interactions (≤ 4.5 A). The underlined VHH amino acids (Kabat numbering [7]) were altered by site directed mutagenesis to validate the interactions observed in the crystal structure.

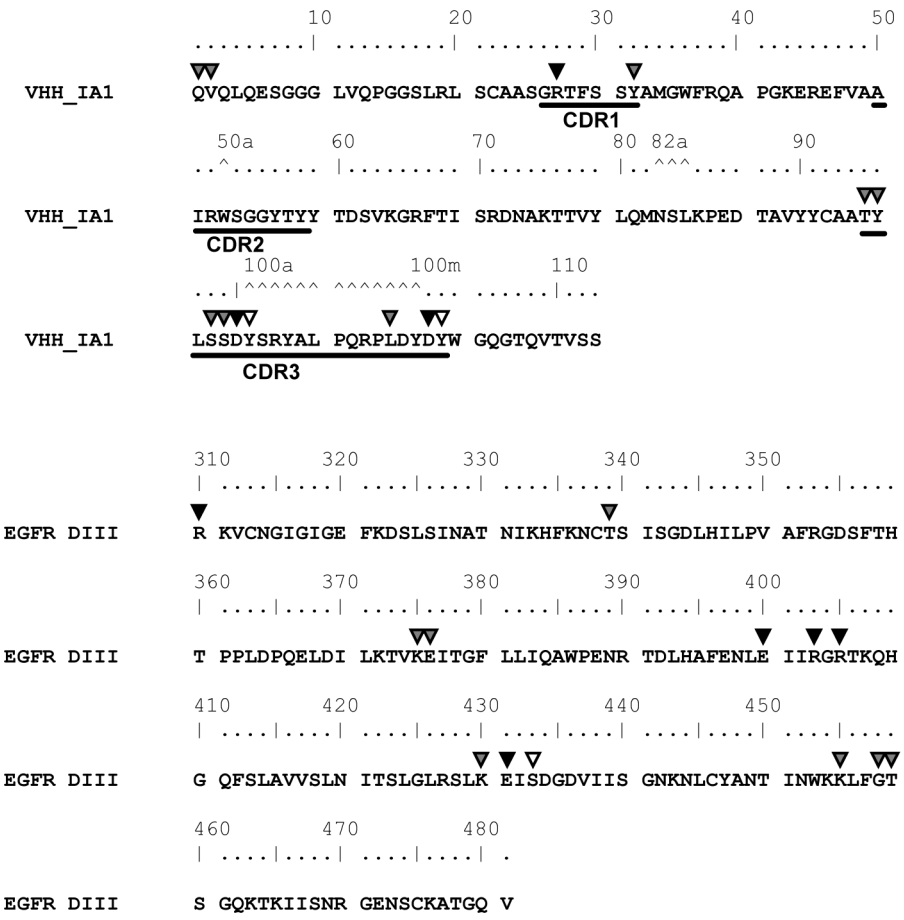

Figure 70. Details of the interaction between VHH IA1 and domain III of EGFR.

The sequences of VHH IA1 (Kabat numbering [7]) and domain III of EGFR are shown. Amino acids marked by filled triangles mediate electrostatic interactions in the VHH/sEGFR crystal structure; gray triangles indicate predicted hydrogen bonds; white triangles denote van der Waals contacts. The interaction is stabilized by contacts to the amine terminus, CDR1, and CDR3 of VHH IA1.

3.3.7 The conformations of free and bound VHH IA1

The availability of free and bound structures of VHH IA1 allows us to compare the conformational changes that occur in this VHH upon epitope binding. The framework region of the Ig fold is essentially identical in bound and free forms, which superimpose with a backbone atom RMSD < 1.0 Å. However, differences are apparent in the paratope. In free IA1, the N-terminal region lies between CDRs 1 and 3, such that the side chain of V2 projects into a hydrophobic crevice lined in part by Y32 (Fig. 71A). In the bound structure, the N-terminal region is oriented away from the globular Ig core and makes polar contacts with domain III (Fig. 69 & 71A). CDR1 in the bound form adopts a slightly different conformation (backbone RMSD: 1.98 Å), in which T28 has moved up and F29 has moved down towards the hydrophobic core, occupying part of the vacancy left by V2 (Fig. 71B). M34 and Y32 side chains rotate upward. The rotation by Y32 brings this side chain in position to make polar contacts with EGFR E431 (Fig. 69).

Less extensive rearrangements occur between CDRs 2 and 3. The C-terminal region of CDR3 undergoes a slight conformational change in the bound form, including a small shift by Y100d (Fig. 71C). Y56 of CDR2 rotates upward ≈ 180° in the bound form, and packs against Y100d. Aside from this change, CDR2 (which makes no contacts to sEGFR) and CDR3 adopt essentially the same orientation in both structures (RMSD < 0.6 Å).

3.3.8 The conformation of sEGFR in the VHH IA1 complex

EGFR domain III in the VHH IA1 complex adopts an identical conformation to domain III in the crystal structures of unliganded sEGFR (PDB: 1NQL; RMSD = 0.55 Å) and the FabC225/sEGFR complex (PDB: 1YY9; RMSD = 0.38 Å). However, conformational differences are apparent in domain II that slightly alter the overall orientation of domains in the tethered sEGFR. Comparison of unliganded sEGFR, the FabC225/sEGFR complex, and the VHH IA1 ternary complex reveals two locations of flexibility, or “pivot points.” These occur just prior to the C240/C267 disulfide bond (Fig. 72A) and just prior the C305/C309 disulfide bond [14] (Fig.

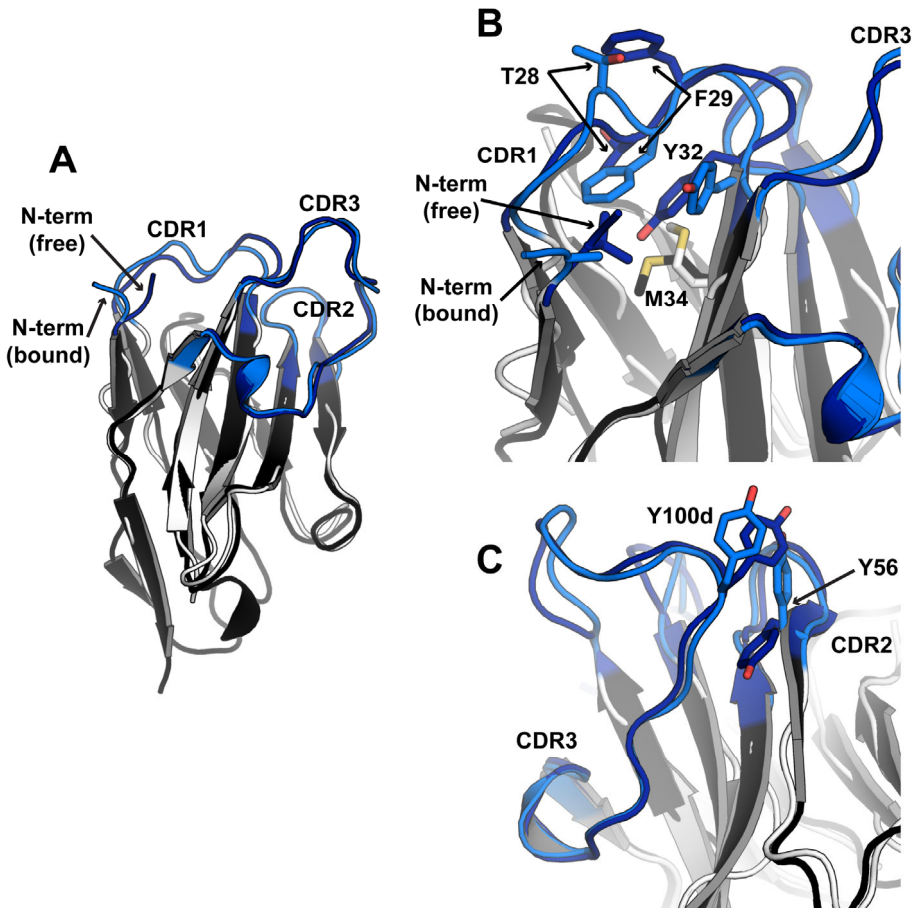

Figure 71. Comparison of free and bound conformations of VHH IA1.

Binding to domain III induces slight conformational changes in VHH IA1. (A) In free IA1 (black; CDRs dark blue) the N-terminus packs against CDR1 and the hydrophobic core of the Ig fold. In the bound structure (white; CDRs light blue), this segment is oriented away from the VHH and appears to makes polar contacts with domain III. (B) Movement of the N-term removes a valine side chain from the core. Several hydrophobic side chains change position. M34 and Y32 side chains rotate, while T28 and F29 essentially switch positions. Rotation of Y32 brings this side chain in position to form a hydrogen bond with domain III. (C) Motions are also noted in CDR2 and CDR3, most notably a large rotation by Y56.

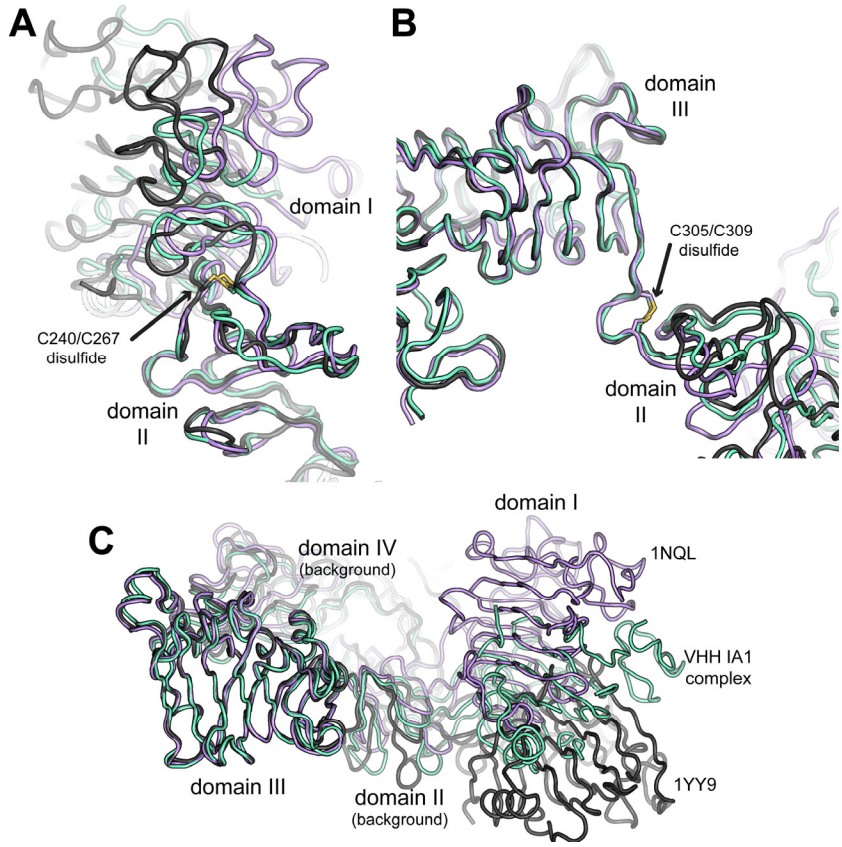

Figure 72. The relative positions of domains I and III in several EGFR structures indicates points of flexibility in domain II.

Ribbons representations from the structures of tethered sEGFR at low pH (violet; PDB ID: 1NQL), the FabC225/sEGFR complex (black; 1YY9), and VHH IA1/FabC225/sEGFR (teal). (A) Portions of domains I and II are shown, aligned with respect to residues 268 – 304 of domain II. This reveals a site of conformational flexibility, or "pivot point," immediately prior to the C240/C267 disulfide bond. (B) Portions of domains II and III are shown aligned with respect to domain III. A second pivot point precedes disulfide bond C305/C309. (C) A different view of sEGFR, aligned with respect to domain III, reveals different relative positions of domains I and III due to flexibility at the two domain II pivot points. Domain I in the VHH IA1 complex adopts a position between that observed in 1NQL and 1YY9.

72B), suggesting that regions 1-239 and 240-304 move as rigid bodies in unliganded sEGFR. Domain I in the VHH IA1 complex occupies a position, relative to domain III, between the positions observed in the unliganded (PDB: 1NQL) and FabC225 complex (1YY9) structures (Fig. 72C).

3.3.9 VHH 9G8 binds EGFR similarly to VHH IA1, in spite of differences in CDR3

A similar crystallization strategy was employed to determine the structure of bound VHH 9G8, which competes with VHH IA1 for binding to sEGFR (Fig. 53). Rod-like crystals of VHH 9G8/FabC225/sEGFR (100 × 100 × 1000 μm; Fig. 73) diffracted to 2.8 Å. The crystal form of the VHH 9G8 complex is essentially identical to that of the VHH IA1 complex (overall backbone RMSD = 0.54 Å). FabC225 and sEGFR adopt identical conformations, VHH 9G8 occupies a similar position between EGFR domains II and III, and the same regions of the complex are less well ordered as in the VHH IA1 complex structure (see Section 3.3.6).

As in the VHH IA1 complex, ordered contacts are only observed between VHH 9G8 and EGFR domain III, although VHH 9G8 is adjacent to domain II (Fig. 74). The interaction with sEGFR buries an average surface area of 638 $Å^2$ on each partner, ≈ 100 $Å^2$ of which is on domain II (Table 10). A relatively low proportion of this area, 39%, involves apolar surface. Binding of VH 9G8, like VHH IA1, is stabilized largely by polar and electrostatic contacts. N-terminal residues, CDR1, and CDR3 contribute to the interaction; CDR2 makes no stabilizing contacts (Figs. 75 & 76).

In addition to a backbone hydrogen bond between the VHH 9G8 V2 and domain III, VHH 9G8 binding is stabilized by two of the three contact clusters observed in the VHH IA1/sEGFR complex (Fig. 75 & 76). The side chains of VHH 9G8 amino acids R27, Y32, and Y102 interact with EGFR E431, and R27 makes additional backbone contacts to domain III. These VHH residues are identical in VHHs IA1, and this first interaction cluster is essentially identical in both structures. The second cluster is formed by electrostatic contacts from VHH 9G8 D101 and E100i to both EGFR R403 and R405, and a polar contact between the CDR3 backbone and

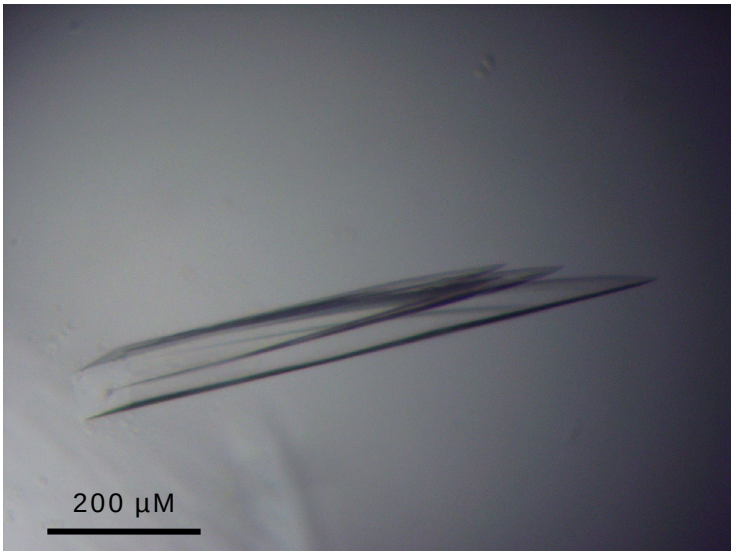

Figure 73. Crystals of the VHH 9G8/Fab C225/sEGFR complex.

Representative crystals of the VHH 9G8/Fab C225/sEGFR complex. 100 × 100 × 1000 µm crystals formed in 10% P3350, 0.1 M HEPES, pH 7.0.

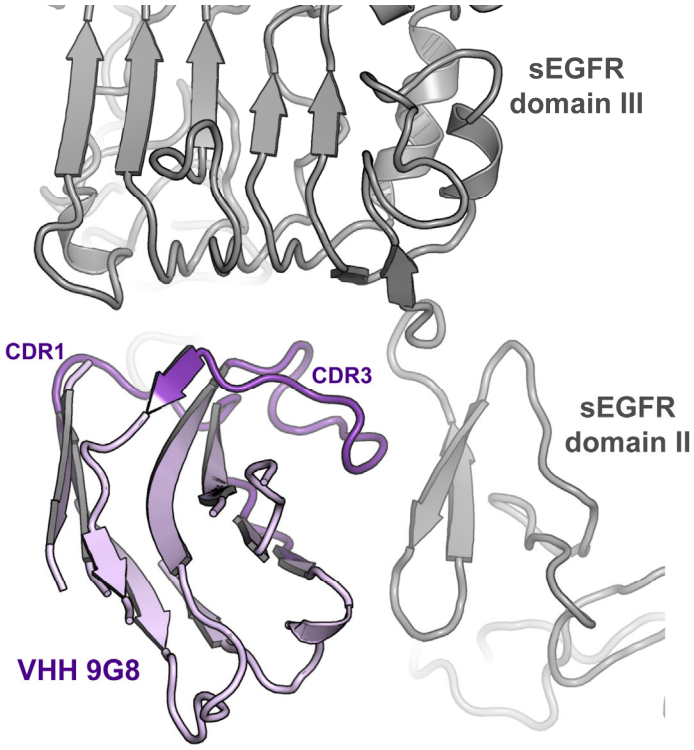

Figure 74. VHH 9G8 binds between domains II and III of sEGFR.

Cartoons of VHH 9G8 (purple) and sEGFR (gray) from the VHH 9G8/FabC225/sEGFR complex structure are shown. VHH 9G8 lies in a cleft formed by sEGFR domains II and III, though ordered contacts are only observed to domain III.

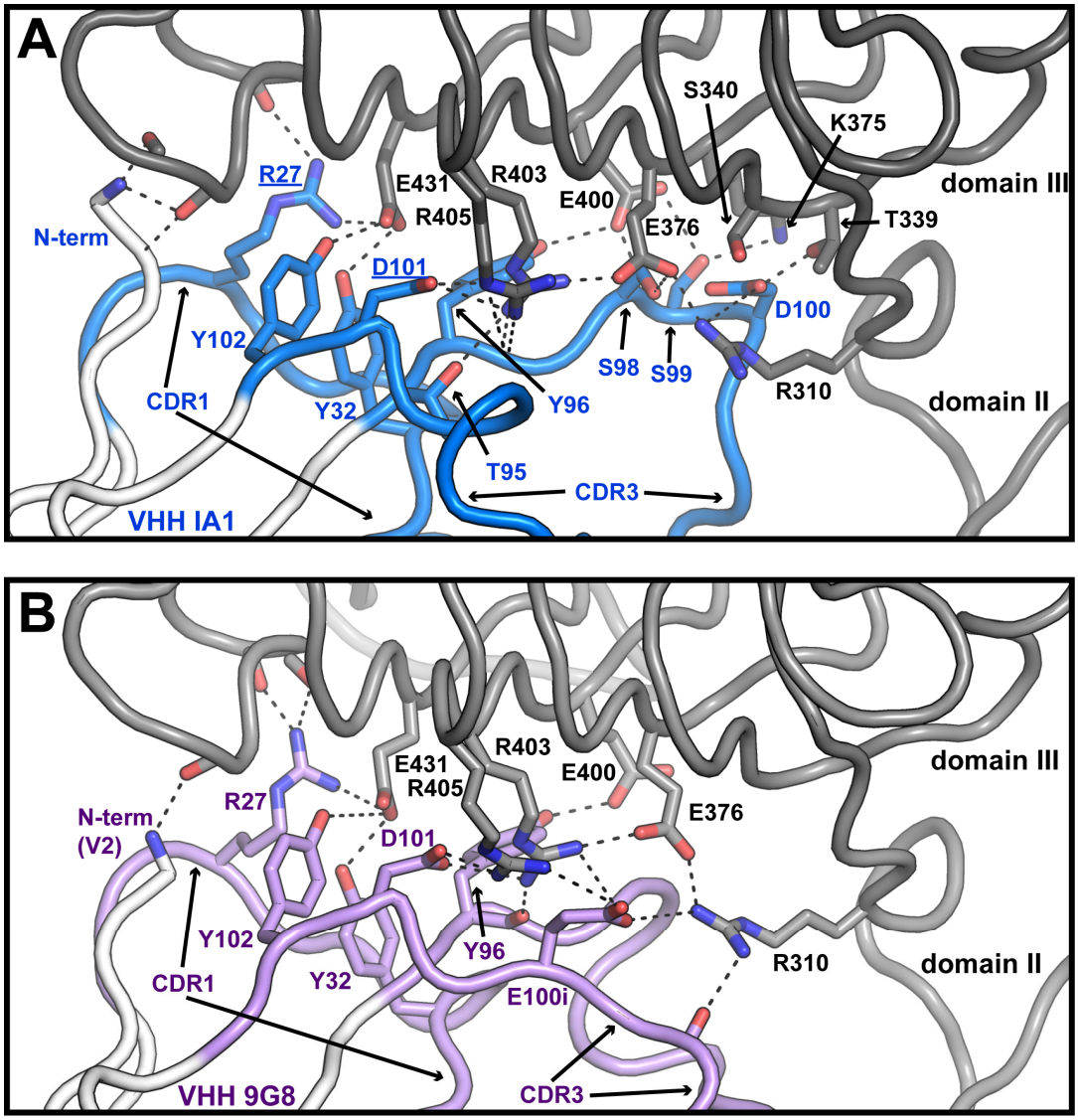

Figure 75. Detail of the interaction between VHH 9G8 and domain III of EGFR.

(A) Cartoons of VHH IA1 (blue) interacting with domain III (gray), reproduced from Fig. 69, and (B) VHH 9G8 (purple) binding domain III (gray). Interactions mediated by the VHH 9G8 N-term, CDR1, and the C-terminal portion of CDR3 (D101) are very similar to those stabilizing bound VH IA1. The shorter CDR3 of VHH 9G8 makes fewer stabilizing contacts do domain III, but does make important electrostatic interactions. Dashed lines indicate predicted hydrogen bonds (≤ 3.5 Å) or electrostatic interactions (≤ 4.5 Å).

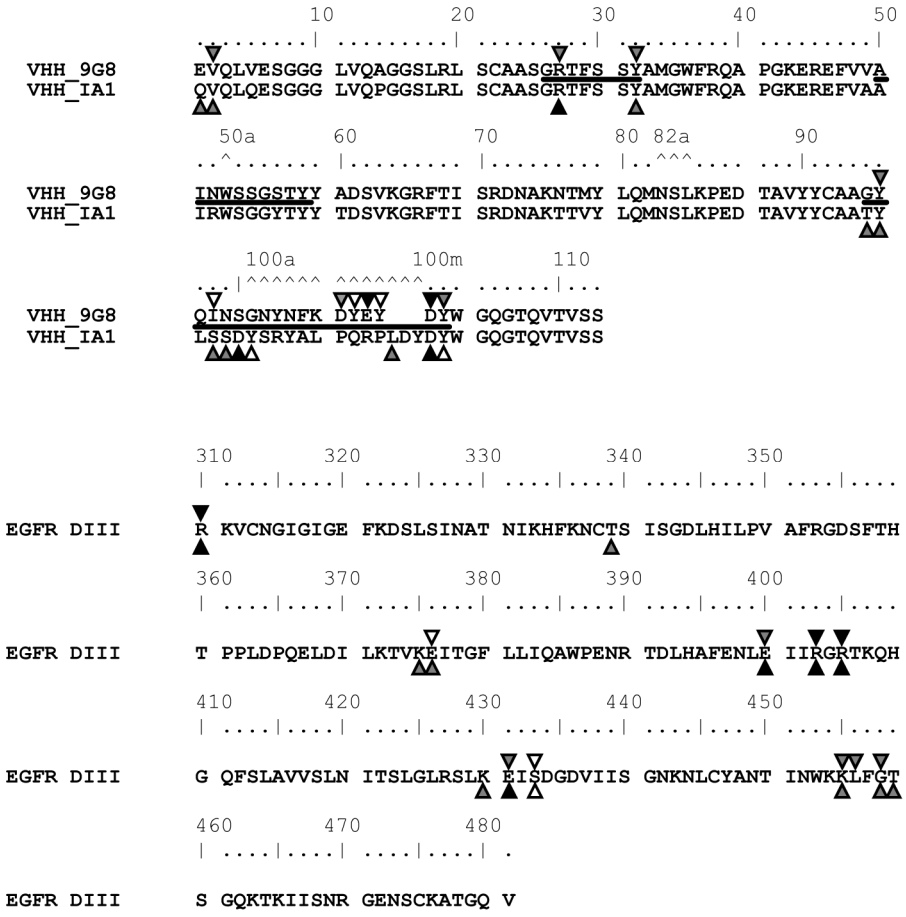

Figure 76. Details of the interaction between VHH 9G8 and domain III of EGFR.

The sequences of VHH 9G8, VHH IA1, and domain III of EGFR are shown. Amino acids marked by triangles above the sequence denote interactions between VHH 9G8 and EGFR; triangles below the sequence denote interactions between VHH IA1 and EGFR. Filled triangles indicate electrostatic interactions; gray triangles indicate predicted hydrogen bonds; white triangles denote van der Waals contacts. VHH 9G8 binding is stabilized by interactions with the amine terminus, CDR1, and CDR3. Contacts made by the N-terminal region of VHH 9G8, CDR1, and the C-terminal region of CDR3 are similar to those by VHH IA1.

R403. A salt bridge between D101 and EGFR is observed in the second VHH IA1 interaction cluster as well, but E100i is unique to VHH 9G8. The third cluster of the VHH IA1 complex is not formed by VHH 9G8, as 9G8 CDR3 is shorter and of different composition. Instead, in this N-terminal region of domain III, VHH Y96 contacts EGFR E400 and VHH E100i forms a salt bridge to EGFR R310 (replacing the D100/R310 interaction observed with VHH IA1). As with VHH IA1, a network of electrostatic contacts is created by VHH 9G8 binding: VHH D101 contacts EGFR R403 & R405, which interact with both EGFR E376 and VHH E100i, both of which interact with EGFR R310.

3.3.10 Comparison of VHH IA1, VHH 9G8, and matuzumab binding to EGFR

The VHH IA1 epitope partially overlaps that of the inhibitory antibody matuzumab [6] (Figs. 77 & 78), and, like matuzumab, does not overlap the ligand binding region of domain III. Matuzumab binding is stabilized by extensive contacts to the domain III loop 458-462, to which VHH IA1 makes hydrogen bonds. Matuzumab also forms a salt bridge with EGFR D431, which interacts with VHH IA1 R27. This is consistent with binding and competition experiments, indicating that VHH IA1 competes for receptor binding with matuzumab, but does not directly compete for ligand binding to EGFR [205]. However, aide from this overlap, the matuzumab and VHH IA1/9G8 epitopes are quite distinct. The VHH epitopes are located further towards the N-terminal end of domain III (Fig. 78), and further towards the “bottom” (the opposite side of domain III as the ligand binding site).

4. Discussion

4.1 Antibody inhibition of sEGFR

4.1.1 Discussion of Antibody Epitopes

The antibodies in this study, with the exception of 2E9, all bind to domain III of EGFR. The epitopes of several other inhibitory antibodies of clinical significance have been studied. The X-ray crystal structure of Fab11F8 has been described, in complex with domain III [147].

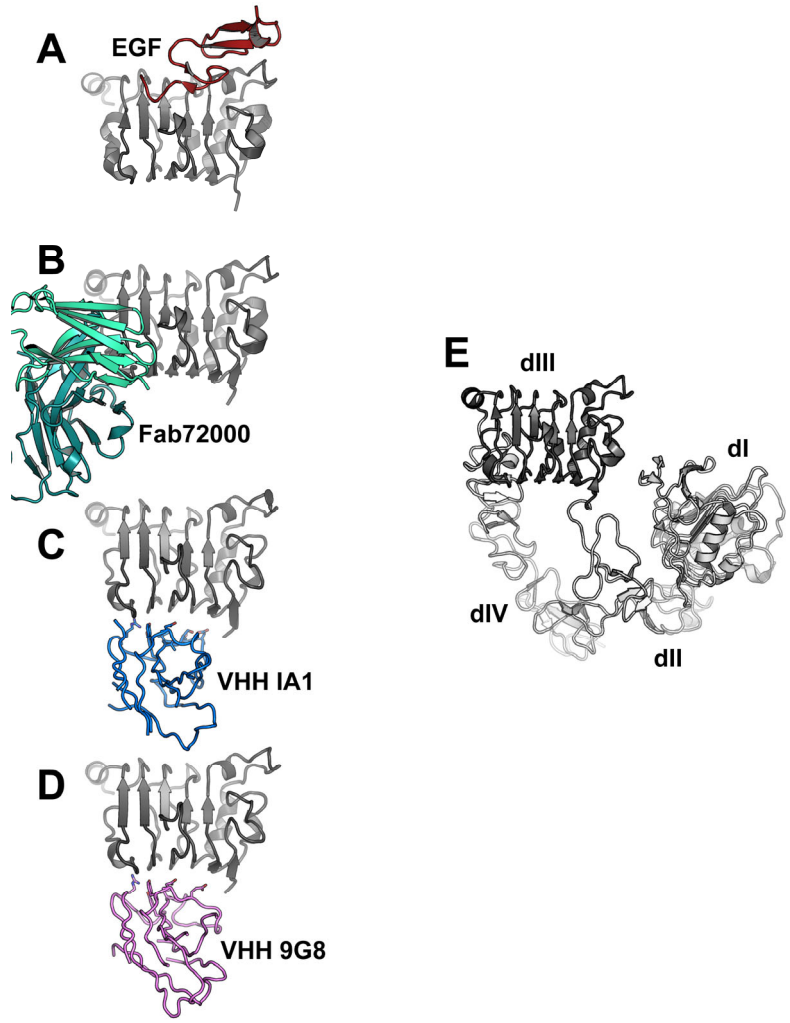

Figure 77. VHH IA1 and VHH 9G8, like matuzumab, do not compete with ligand.

(A-D) Cartoons of EGF (dark red, A), matuzumab Fab (teal, B), VHH IA1 (blue, C) and VHH 9G8 (violet, D) bound to domain III of EGFR (gray). While EGF binds to the "top" of domain III, matuzumab and VHHs IA1 and 9G8 bind to the "side" or "bottom," and clearly do not sterically clash with ligand binding. Matuzumab makes extensive contacts to the C-terminal end of domain III. (E) A cartoon of tethered sEGFR in the same orientation as in (A-D).

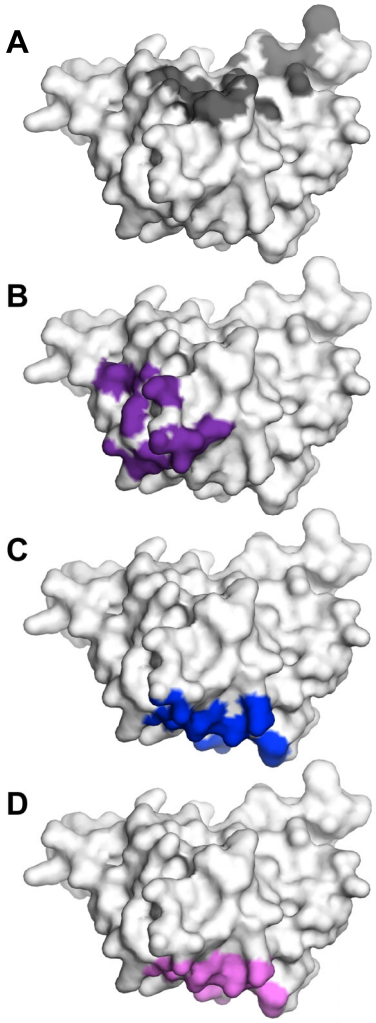

Figure 78. Comparison of the binding sites of EGF, Fab72000, VHH IA1, and VHH 9G8.

The EGF binding footprint (gray, A), Fab72000/matuzumab epitope (orange, B), VHH IA1 epitope (blue, C), and VH 9G8 epitope (pink, D) are shown on surface representations of EGFR domain III, which is rotated 90° about a horizontal axes compared to its orientation in Figure 62D. The highlighted areas indicate surface exposed atoms within 4 Å of bound ligand or antibody. VHH IA1 and 9G8 epitopes overlap closely, and each overlaps with the Fab72000 epitope, but not the ligand binding site. Compare to the epitopes of FabC225 and VHH 7D12 in Fig. 64.

The binding of HumaxEGFr (zalutumumab) has been studied by cryo-EM and mutational analysis [10]. Fine epitope analysis has been used to investigate the binding of ABX-EGF (panitumumab) [9]. The epitopes of each of these antibodies has been found to overlap the 225 epitope and the ligand binding region. Zalutumumab and panitumumab have been further characterized by antibody competition assays. They have been shown to compete with each other for binding to cell surface EGFR, and both compete with cetuximab and matuzumab [212].

These reports, combined with the results presented here, allow us to roughly map out the epitopes of these six inhibitory antibodies on the surface of domain III (Fig. 79). Antibodies C225, 11F8, and 13A9 bind to similar regions of domain III, overlapping (C225 and 11F8) or adjacent to (13A9) the ligand binding region, but not overlapping with 425. Antibodies 425 and 108 bind further from the ligand binding region on domain III, spatially nearer domain IV. Panitumumab and zalutumumab bind intermediate epitopes, overlapping the epitopes of C225 and 425, as well as the ligand binding site [9, 10].

It is notable that most characterized inhibitory EGFR antibodies bind to domain III, even though inhibition via domain III is not the only route towards inhibition of receptor activation. Antibody 2E9 inhibits activation via domain I, and trastuzumab and pertuzumab inhibit activation of the related RTK Her2/ErbB2 via domains IV and II, respectively. This domain III bias in EGFR antibodies may partly arise from similarities in the antibody generation and screening strategies. The five antibodies characterized in this study were created by immunizing mice with intact cells or preparations from high EGFR expressing cells lines (A431 or transfected CHO), typically followed by screening for inhibition of ligand binding [25, 133, 136, 139, 140]. Panitumumab and zalutumumab were generated by similar methods using mice that produce fully human antibodies [177, 180]. Antibody IMC-11F8 was derived from a phage display system, but was identified by screening for competition with cetuximab binding to EGFR [155]. The bias towards domain III may also be a result of the apparent antigenic nature of the domain. Domain III presents a large, relatively charged surface, free of glycosylation, and accessible to solution in the autoinhibited conformation of the receptor [147].

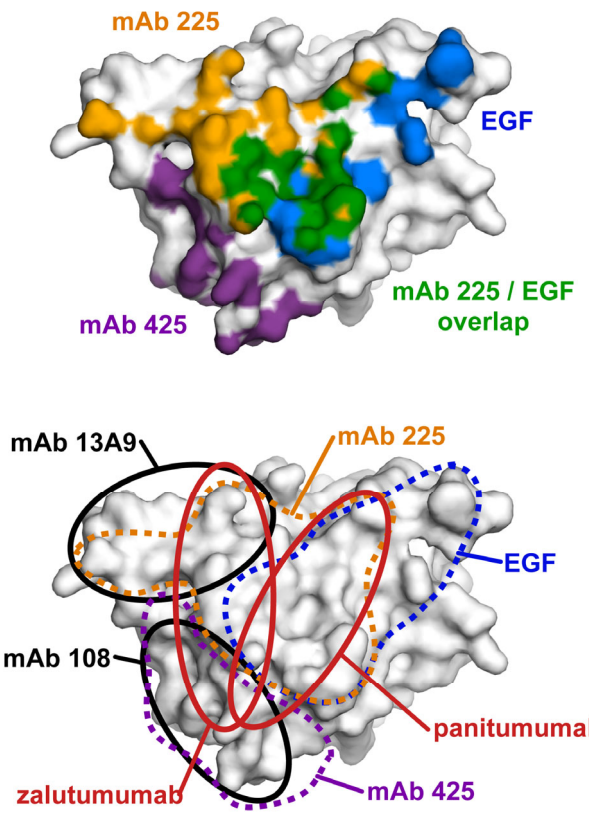

Figure 79. Binding surfaces of ligand and inhibitory antibodies on EGFR domain III. (A) The epitopes of inhibitory antibodies cetuximab (gold) and matuzumab (purple), and the binding footprint of ligand EGF (blue), are shown on a surface representation of EGFR domain III. Domain III surface area involved in both cetuximab and EGF binding is shown in green. (B) The approximate epitopes of six inhibitory antibodies, and the binding footprint of EGF, are shown on EGFR domain III. Binding surfaces determined from crystallographic studies are outlined by dashed lines [2, 5, 6]. Epitopes of mAbs 13A9 and 108 (black) are inferred from surface mutation (Figs. 43, 44) and antibody competition studies (Fig. 45). The epitopes of panitumumab and zalutumumab (red) are inferred from biochemical studies [9, 10].

4.1.2 mAb 108, mAb 2E9, and negative cooperativity

We find that antibodies 108 and 425 have at least partially overlapping epitopes and directly compete for binding to sEGFR. However, mAb108, but not mAb425, inhibits "high-affinity" EGF binding to receptors at the cell surface [25, 136]. Multiple hypotheses have been presented to explain the phenomena of heterogeneous receptor binding to cells, including binding to extended and inactivated receptor populations, heterogeneous receptor density [125] at the cell surface, the involvement of an external interaction partner [126], and negative cooperativity [127]. Different explanations may be proposed to explain this inhibitory behavior of mAb 108, to address these different hypotheses. It is difficult to extrapolate from results observed in the context of purified EGFR extracellular region to events at the cell surface, since "heterogeneity" in ligand binding is not observed in solution with purified sEGFR [124].

If the effects of mAb 108 and 2E9 arise in the context of negative cooperativity, they may function by stabilizing different conformations or states of the EGFR dimer. Antibody 2E9, which recognizes domain I of the receptor, might prevent or disfavor formation of a symmetric 2:2 dimer, but not an asymmetric 1:2 dimer (ligand:EGFR). Conversely, mAb 108 might restrict formation of an asymmetric 1:2 dimer, but not a symmetric 2:2 dimer. The lack of molecular understanding of the details of ligand binding at the cell surface, and the absence of a crystal structure of Fab108 or Fab2E9 bound to receptor, makes it difficult to test these hypotheses.

4.1.3 Partial inhibition of ligand binding by antibodies 13A9, 108, and 425

Biacore competition studies indicate that antibodies C225 and 11F8 essentially abolish sEGFR binding to ligand, while binding is only partly diminished by antibodies 13A9, 108, and 425. For binding to ligand to occur at all in the presence of saturating amounts of high affinity ligand indicates that these latter antibodies do not directly compete with ligand binding to EGFR. Rather, the observed decrease in binding is likely due to steric constraints experienced by the binding of antibody:sEGFR complex to immobilized ligand. While the concentration of sEGFR in these competition experiments (600 nM) is nominally above the K_D for the EGF/sEGFR

interaction (130 nM), the antibody:sEGFR complex likely binds immobilized ligand with diminished affinity.

We also observe that the presence of excess Fab fragments of 108, 425, and 13A9 results in a greater fold decrease in receptor binding than caused by intact mAbs. Rather than attribute this difference to enhanced competition by the Fab, we suggest that it arises from the stoichiometric nature of the assay. The magnitude of the SPR response is generally proportional to the average mass of the material bound to the chips surface. Thus, binding of a Fab:sEGFR complex would be expected to produce ≈ 57% the SPR response of a mAb:sEGFR complex, and ≈ 42% the response of a mAb:2·sEGFR complex. This correlates with the relative response values observed for Fab and mAb in Figure 46B, though we cannot conclusively determine whether a 1:1 or 1:2 mAb to sEGFR complex binds in the mAb competition experiments.

4.1.4 Selective competition of TGFα by 13A9

Activation of EGFR by the ligands EGF and TGF has been shown to elicit different cellular responses [140]. The selective inhibition by mAb 13A9 of TGFα, but not EGF, binding has long been an interesting observation in the literature [140]. Our data place the 13A9 epitope near the surface of domain III that binds the C-terminal peptide regions of mature EGF and TGFα (Figs. 44 & 79). This concurs with studies utilizing chimeric TGFα/EGF ligands, which implicate the C-terminal region of mature ligand as most responsible for the differential competition [130]. The reason for this is not immediately apparent from the sequences of the ligands, as the terminal region of EGF is actually longer than that of TGFα by four amino acids.

Based on the X-ray crystal structures of sEGFR in complex with EGF [2] and TGFα [1], we note three key differences that may account for the differential selectivity of mAb 13A9 (Fig. 80). First, domains I and III occupy slightly different relative positions in the EGF bound and TGFα bound receptor structures (Fig. 80A), resulting in a RMSD of 1.27Å in the first repeat of domain I (E2-V36 of mature EGFR). Steric clashes between bound 13A9 and domain I may contribute to the ability of 13A9 to selectively block high-affinity bivalent binding to TGFα but not

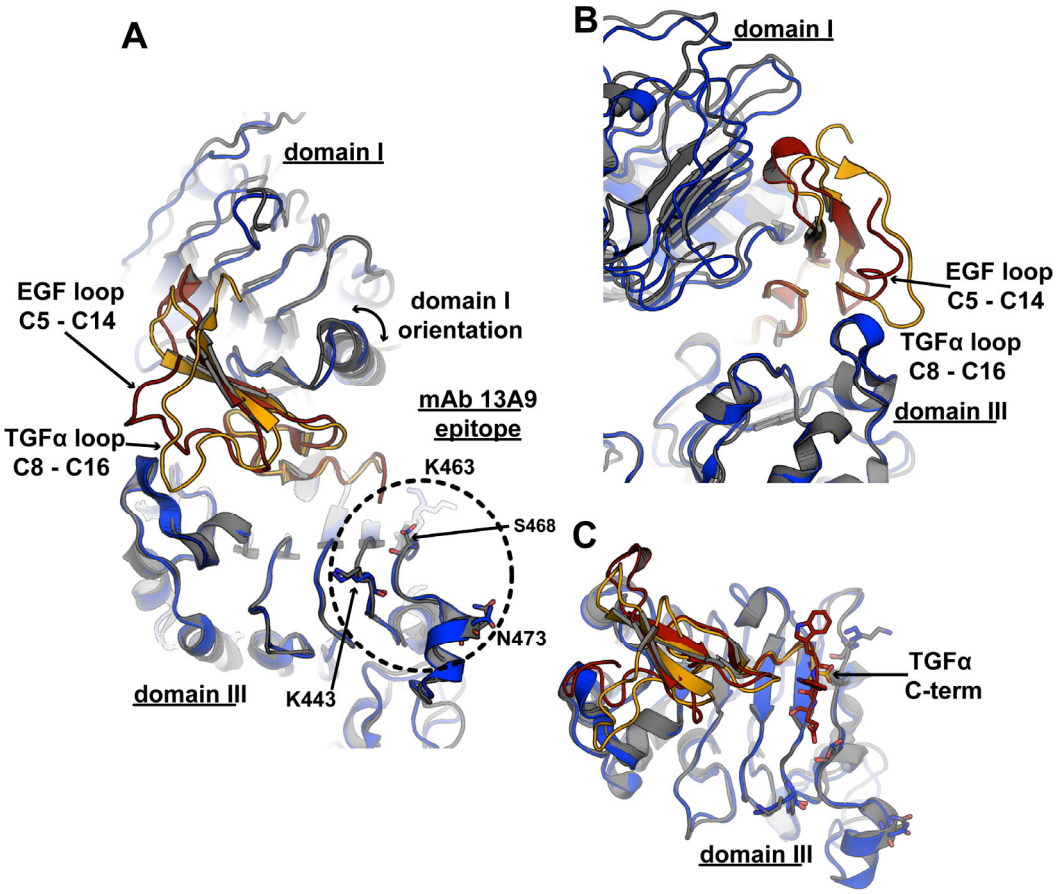

Figure 80. Differences between the structures of EGF and TGFα bound to sEGFR may account for the selectivity of mAb 13A9.

(A) A cartoon of the overlaid structures of EGF (red) bound to EGFR (gray) (PDB: 1IVO) and TGFα (yellow) bound to EGFR (blue) (1MOX). The approximate epitope of mAb 13A9 is circled (Fig. 44). A domain I α-helix located above the mAb 13A9 epitope occupies different positions in EGF and TGFα bound structures, and is nearer the epitope in the TGFα bound structure. (B) The loop between the first two cysteines of ligand packs against EGFR in the structure of bound EGF, but projects into solvent in the TGFα structure. (C) EGF possesses a longer C-terminal region than TGFα. The charged C-terminus of TGFα lies near the 13A9 epitope, and may interfere with 13A9 binding.

to EGF. This hypothesis could be tested by assessing the effects of bound 13A9 on the binding of ligand to domain III alone. Second, the loop between the first and second cysteines (C6 – C14) of mature EGF packs against the receptor, while the corresponding segment of bound TGFα (C8 – C16) projects outward into solvent (Fig. 80B). This region is within ≈ 30 Å of the approximate 13A9 epitope. The "length" of a typical Fab fragment (from the paratope surface to the C-termini) is ≈ 60 Å. If 13A9 binds domain III at an angle, with respect to the paratope surface, steric effects between bound TGFα and 13A9 may be possible. Finally, the terminal residue of TGFα, A50, and the corresponding residue of EGF, W49, are both ordered [1, 2]. The negatively charged carboxy-terminus of TGFα may repulse an acidic region on the 13A9 Fab region, destabilizing simultaneous binding (Fig. 80C).

These arguments suppose that the mode of interaction between receptor and ligand observed in the crystal structures of EGF and TGFα bound to EGFR are the only ones that occur in solution. It is conceivable that binding modes other than these occur, and that the competitive selectivity of mAb 13A9 arises from these ligand bound states.

4.1.5 VHH vs Fab binding to EGFR

Crystal structures have been reported of VHHs bound to proteins, including complexes with lysozyme [33, 195, 213-216], RNase A [207, 217], pancreatic β amylase [34] , and the bacteriophage P2 receptor binding protein [209] (in which VHH was utilized as a crystallization aid). To our knowledge, the complexes described here of VHH bound to sEGFR represent the first crystal structures of inhibitory VHHs bound a clinically relevant target.

Inhibitory VHHs 7D12, IA1, and 9G8 all recognize EGFR domain III. This region of EGFR has been noted as particularly antigenic [147], and is recognized by most of the characterized conventional inhibitory mAbs (Table 6) [128]. This is partly because domain III provides a large and relatively flat epitope, against which the Fab region may engage much of its relatively flat antigen binding surface. Planar interfaces are observed in the complex structures of FabC225, FabIMC-11F8, and FabEMD72000.

In contrast to Fab fragments, the available antigen recognition surface of VHHs is convex or wedge shaped, due both to the typically elongated CDR3 and to the absence of the VL domain. While VHHs exhibit variability in their mode of antigen binding [34], there is speculation that these fragments are especially suited to binding pockets or clefts. Such interactions would maximize the contact area between VHH and antigen, and may contribute to the nanomolar affinity achieved by some VHHs in spite of a smaller variable region [218]. For example, VHH cAb-Lys3 binds a cleft shaped epitope on lysozyme, resulting in an interaction surface area values on par with Fabs that bind this antigen [208].

In spite of the expected preference for a concave epitope, each of the VHHs characterized here makes ordered contacts with EGFR over a relatively planar surface. Rather than project outward as observed in the cAb-Lys3/lysozyme complex, the CDR3 of each anti-EGFR VHH folds back against the antibody, stabilized by van der Waals contacts (though not by disulfide bonds, as often occurs in camel HCAbs [189]). VHH 7D12 binds to a relatively planar epitope on EGFR domain III, stabilized largely by contacts to the folded portion of CDR3. VHHs IA1 and 9G8 bind EGFR in a cleft formed between domains II and III. However, ordered contacts are only observed between VHH and sEGFR domain III, also involving residues from the folded portion of CDR3. In none of the VHH/EGFR structures are ordered contacts made along a cavity-shaped epitope.

We observe that these VHHs interact with domain III at an angle, engaging less than the entire theoretical interaction surface. Consequently, VHH binding buries a smaller surface area, with weaker associated K_D values, than binding of conventional antibodies to ErbB receptors [5, 6, 24, 31, 147] (Table 10). Structures of Fab fragments bound to EGFR reveal excluded surface area values between 758 and 932 Å^2 (around the typical range for antibody/antigen interactions, ≈ 750 – 900 Å^2 [219]), compared to 683 – 755 Å^2 for VHH/EGFR binding.

4.1.6 Possible changes in EGFR conformation induced by VHH IA1 and 9G8

We note a difference in the position of domain I in the VHH IA1 and VHH 9G8 ternary complexes versus the FabC225/sEGFR complex (Fig. 72). This difference could be the result of

steric effects of bound VHH IA1, or of inherent flexibility in sEGFR. The tether region in the VHH IA1 complex structure is too poorly ordered to determine if tether stabilizing contacts are maintained. However, the same region of sEGFR is better ordered in the VHH 9G8 complex structure, where it is apparent that at least some tether contacts are preserved (Fig. 81).

The crystallographic data for the VHH 9G8 complex yields interpretable electron density in the region of the EGFR domain II/IV tether interaction. These contacts consist of four hydrogen bonds between the domain II tether and domain IV, which partially stabilize the unliganded receptor conformation [14, 17]. In the VHH 9G8 complex we clearly observe two of the four stabilizing hydrogen bonds (one between the backbone carbonyl of Y251 and the side chain of H566, and another between the side chains of Y246 and D563) indicating that the tether is not sterically disrupted in this complex structure. The conformational similarity between the VHH IA1 complex and VHH 9G8 complex leads us to conclude that tether contacts occur in both structures.

To enter the extended conformation, a series of interactions between must be broken between the globular portion of domain III and the region 305-312 of EGFR. This includes R310, which breaks its interaction with E376 and faces away from the receptor. VHH IA1 interacts with R310 in the tethered conformation, increasing the energetic cost of leaving the tethered state. Moreover, bound VHH IA1 sterically prevents domain II from adopting the position observed in the extended state of the receptor. Thus VHH IA1 noncompetitively inhibits EGFR activation. VHH IA1 stabilizes the tethered conformation and, like matuzumab, sterically blocks adoption of the extended conformation.

4.1.7 Comparison of VHH inhibitors to cetuximab and matuzumab

The VHH fragments in this study can be classified into two categories of inhibitory antibody, typified by cetuximab and matuzumab. Like cetuximab, VHH 7D12 is a competitive inhibitor. Its epitope overlaps the ligand binding site on domain III of EGFR, and bound VHH 7D12 sterically prevents adoption of the extended conformation. VHHs IA1 and 9G8, by

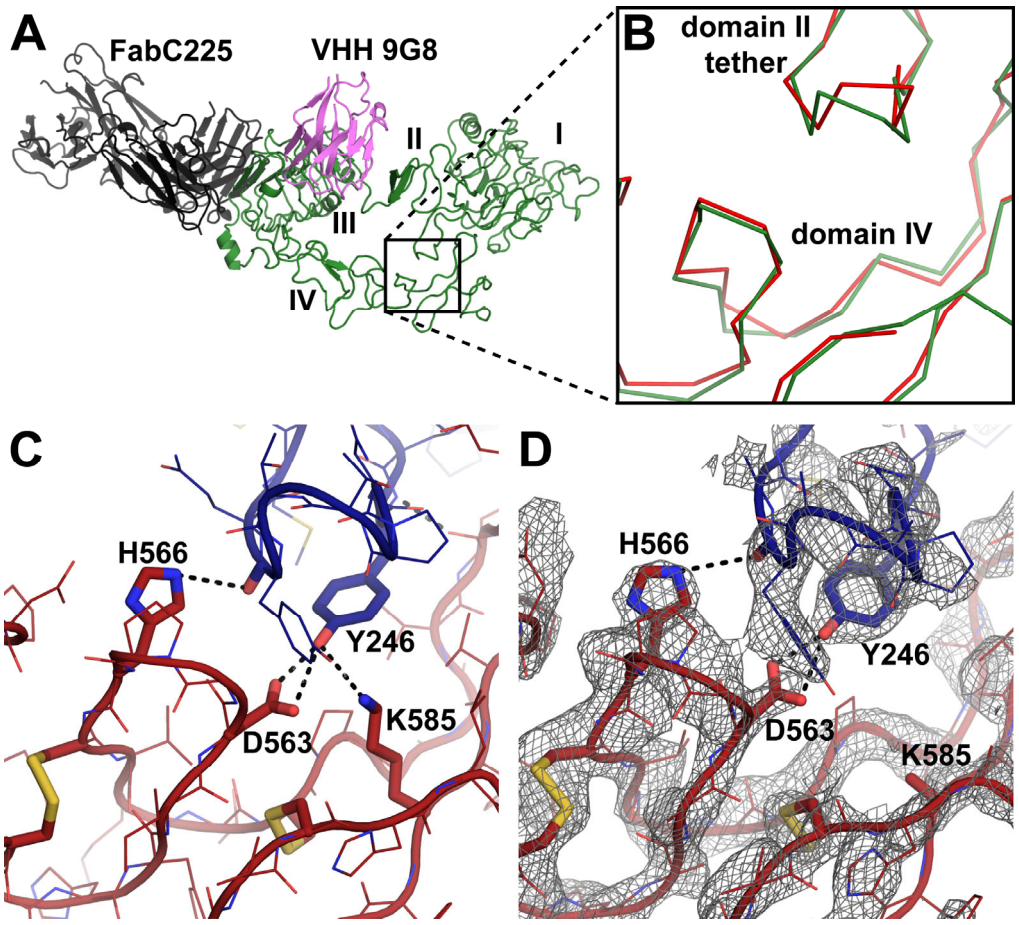

Figure 81. Tether contacts occur in the VHH 9G8 complex structure.

(A) A cartoon of the VHH 9G8 (violet)/FabC225 (black)/sEGFR (green) structure is shown (B) A zoomed in view in the same orientation as (A) shows a ribbons trace of the EGFR tether region from the 9G8 (green) and IA1 (red) complex structures. The overall conformations of these two structures overlap closely. (C) The tether region of sEGFR from structure 1NQL. The tether is stabilized by hydrogen bonds bewteen domain IV (red) and domain II (blue). (D) A 1.5 σ $2f_O$-f_C map is shown over a cartoon of the tether region of the VHH 9G8 complex. Hydrogen bonds are predicted between EGFR H566 and the domain II tether backbone, and between D563 and Y246. The side chain of K585 is not modeled in this structure.

contrast, sterically block the extended conformation and dimerization without directly competing for ligand binding to domain III. As therapeutic agents, VHHs IA1 and 9G8 have a potential advantage, in that their activity is not diminished in the presence of excess ligand.

We have not characterized the inhibitory characteristics of these antibodies on ligand binding to the apparent high- and low-affinity receptor populations at the cell surface. Epitope mapping studies indicate that mAb 108, which specifically blocks high-affinity binding, binds EGFR domain III near the VHH IA1 and 9G8 epitopes (Fig. 79). These VHH inhibitors may exhibit similar effects, which would make them useful tools in understanding the origin of high- and low-affinity receptor populations. Cell based studies of receptor binding will help to further characterize these VHHs.

It is not clear to what extent EGFR domain II stabilizes binding of VHH IA1 and 9G8. These VHHs do not tightly bind sEGFRd3 in solution, which may be due to the presence of the LEEK sequence in sEGFRd3, or to the absence of critical interactions to domain II (or a combination of both effects). These possibilities may be addressed by measuring binding to VHH by alternative sEGFR truncation constructs, incorporating domain III and portions of domain II. If binding of VHH IA1 and 9G8 is not stabilized by interactions to domain II, there may be the potential for a more extensive contact surface. It is possible that structurally guided alterations of VHH framework residues may promote stable contacts with domain II, which in turn may increase the affinity of the interaction and stabilize the tethered conformation of sEGFR.

4.1.8 VHHs as next-generation anti-EGFR therapeutics

As touched upon in the introduction, there are dissatisfactions with conventional mAbs as therapeutic agents. Monoclonal antibodies are often associated with immune mediated side effects (often rashes and inflammation), require expensive and elaborate eukaryotic expression systems, and are laborious to clone and modify. Therapies based on VHH fragments offer advantages over mAbs in some respects, as well as certain inherent limitations.

The glomerular filtration barrier restricts the passage of plasma components into the urine on the basis of charge, shape, and size. The upper size limit for passage through this barrier is approximately 70 kDa (roughly the mass of serum albumen) [220, 221]. VHH fragments (≈15 kDa) are cleared relatively rapidly by the kidneys, whereas mAbs (≈150 kDa) are cleared slowly by metabolically pathways. The elimination half-life ($t_{1/2}$) of a VHH in mouse is on the order of 90 min [218]. By contrast, $t_{1/2}$ of mAb is on the order of days (≈2 days for cetuximab in mice [222]; ≈1 – 5 days for cetuximab in humans [223]; ≈4 – 10 days for matuzumab in humans [224]; ≈20 days for trastuzumab in humans [225]).

The much shorter VHH $t_{1/2}$ detracts from their utility as cancer therapeutics. An efficacious course of treatment with inhibitory anti-EGFR mAbs attempts to achieve bioavalable levels of antibody over weeks. There has been investigation into ways to alter VHH pharmacokinetics. The mass and hydrodynamic size of antibody fragments, and thus their elimination $t_{1/2}$, may be increased by conjugation to polyethylene glycol (PEGylation [226]). VHH fragments may be concatenated, which increases the fragment mass [199]. Multivalent VHH constructs have also been created that bind album in addition to their primary epitope, which was found to increase the circulating $t_{1/2}$.

This potential to "tune" VHH $t_{1/2}$ may be a benefit. While a longer $t_{1/2}$ may be desirable in some applications (for instance, conventional mAb based cancer therapies), relatively rapid renal clearing may be preferable in others. One such application may be the targeted delivery of radionuclides or toxins to tumor cells, in which case high tumor affinity and specificity but rapid elimination may be useful.

4.1.9 Conclusions

The data presented here serve to illustrate the heterogeneity in binding found among inhibitory antibodies directed against EGFR. These antibodies exhibit multiple modes of receptor inhibition and different degrees of ligand competition, and have proved to be valuable tools for dissecting the details of receptor function. Our work supports efforts to generate antibodies with unique inhibitory properties, or antibodies directed against unutilized regions of

the receptor. The development of novel antibodies may prove useful as improved therapeutics, in combination therapies with existing drugs (i.e., to elicit ADCC [212]), and as biochemical tools to probe the details of EGFR activation.

5. Experimental Methods

5.1 Cloning and molecular biology

Standard cloning techniques were used. All sequencing and PCR primers were checked and modified, where necessary, to avoid self-dimers and strong single strand hairpin formation [227]. Typical PCR conditions included 50 ng of template DNA, 250 ng of each primer (Integrated DNA Technologies), 2 µL 10 mM dNTP mix (Invitrogen), 3.5 U of PfuUltra polymerase (Stratagene), 10 µL 10× reaction buffer, in a reaction volume of 100 µL. Restriction enzymes were obtained from NEB. DNA ligations were performed with the DNA Ligation Kit 2.1 (Takara), and incubated ON at 16°C. The sequences of all PCR-generated DNA were confirmed.

5.2 Expression and purification of proteins

5.2.1 Vps74

The *VPS74* locus was amplified by PCR and cloned into pET-28a (EMD), encoding a hexa-histidine (H_6) tag and a T7 epitope tag at the N-terminus of the protein (T7-Vps74), or into modified pET-21a (EMD) containing a tobacco etch virus (TEV) protease cleavage site [228] between the N-terminal hexa-histidine tag and the multiple cloning site (H_6-Vps74). PCR methods were used to generate vectors to express altered forms of Vps74.

To produce bacterial lysates, T7-Vps74 was expressed in *E. coli* BL21 (DE3) (EMD) by induction with IPTG for three hours at 37°C. Cells were lysed in PBS (phosphate buffered saline, 25 mM NaH_2PO_4/Na_2HPO_4, 150 mM NaCl, pH 7.5) containing 1X Complete Mini protease inhibitor cocktail (Roche) using lysozyme and sonication. Clarified lysate was stored at -80º C.

To produce pure Vps74, H_6-Vps74 or T7-Vps74 (or relevant variants) were expressed in BL21-CodonPlus (DE3) RP *E. coli* (Stratagene) in lysogeny broth (LB). Overexpression was induced by 0.5 mM IPTG (isopropyl β-D-1-thiogalactopyranoside, Acros Organics) at an optical density of 0.6, followed by 3 h growth at 37ºC. Selenomethionine-derivitized protein (SeMet-Vsp74) was expressed in B834 (DE3) *E. coli* (Novagen) in simplified minimal medium [229]. Briefly, cells were grown to an OD of 0.6 in LB, pelleted, resuspended in PBS, pelleted again,

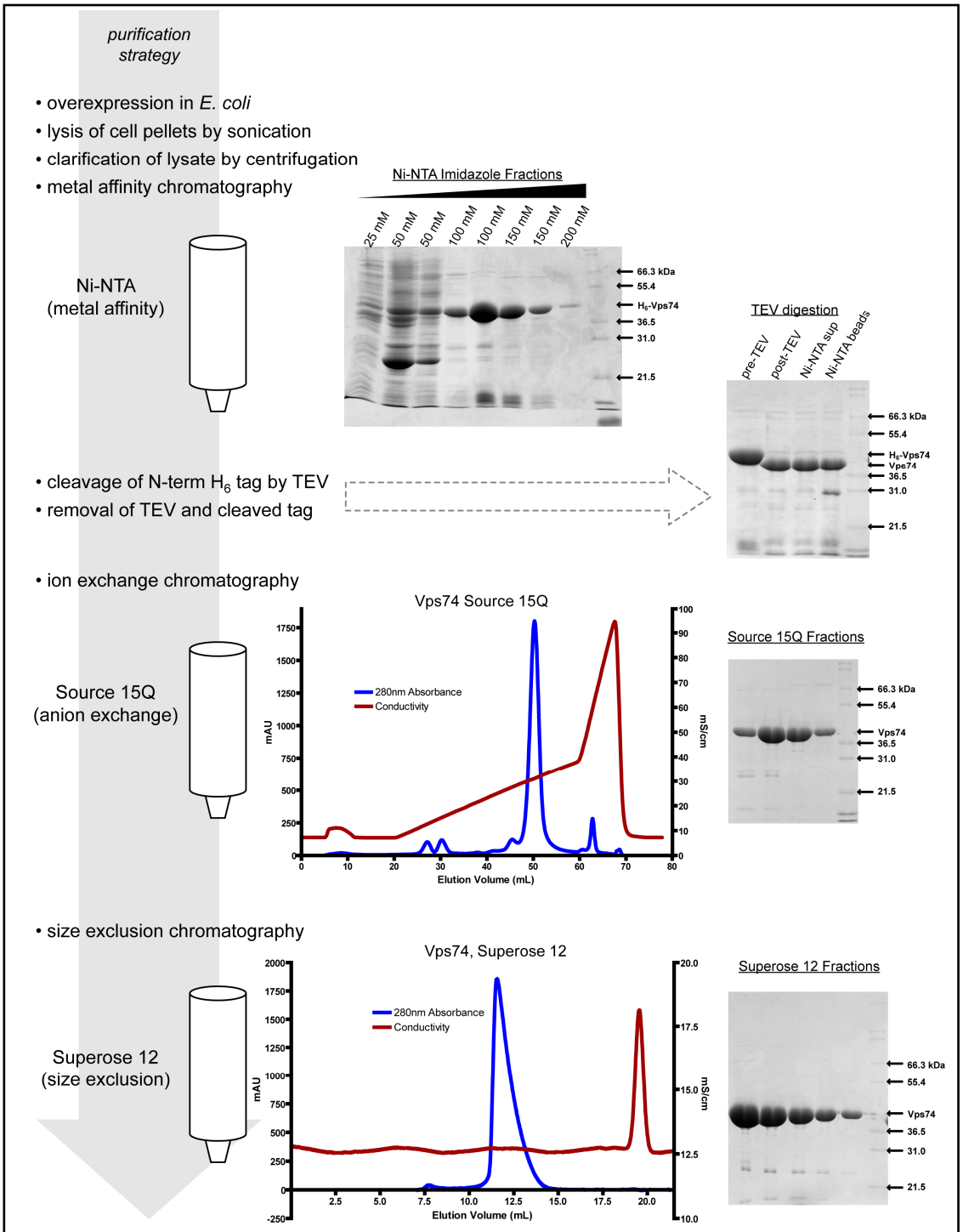

Figure 82. Vps74 purification.

Representative FPLC traces and Coomassie stained SDS-PAGE gels from the purification of full length Vps74.

resuspended in simplified minimal medium (2% glucose, 2.5 mM $MgSO_4$, 19 mM NH_4Cl, 22 mM KH_2PO_4, 22 mM Na_2HPO_4, 10 mg thiamine, 10 mg $Fe_2(SO_4)_3$, 50 mg L-selenomethionine, pH 7.4), and overexpressed overnight at 20°C by the addition of 0.5 mM IPTG.

Cells were lysed by sonication (3 rounds of 25 s total, 50 % amplitude, 1 s on/2 s off) in 25 mM Na_2HPO_4/NaH_2PO_4, 250 mM NaCl, 10% glycerol, pH 8.0. Whole lysates were clarified by 45 min centrifugation at 35,000 rcf in a SS-34 rotor. Clarified lysates were applied to a nickel-nitrilotriacetic acid (Ni-NTA, Qiagen) affinity column by gravity flow. Protein was eluted by a step gradient of increasing imidazole concentration (fractions of 25, 50, 100, 150, 200 mM imidazole in 25 mM Na_2HPO_4/NaH_2PO_4, 250 mM NaCl, pH 8.0). TEV protease (150 µg / 5 mg Vps74) was added to eluted H_6-Vps74 and the proteins dialyzed overnight against 50 mM Tris, 50 mM NaCl, 1 mM DTT, pH 8.0 [228]. The cleaved tag, uncleaved H_6-Vps74, and TEV (which is also his-tagged) were removed by incubation with 50 µL of Ni-NTA resin. TEV-digested Vps74 or undigested T7-Vps74 were purified by anion exchange chromatography (Source 15Q, GE Healthcare) and were eluted by a 5 column volume linear gradient of 50 mM to 300 mM NaCl in 10 mM HEPES, 1 mM DTT, pH 7.5. Proteins were further purified by size exclusion chromatography (Superose 12, GE Healthcare), in 10 mM HEPES, 150 mM NaCl, pH 7.5. Purified protein was concentrated to 10 mg/ml and stored in 10 mM HEPES, 150 mM NaCl, 1 mM TCEP, pH 7.5 at 4°C. Incorporation of selenomethionine into SeMet-Vps74 was assessed by matrix assisted laser desorption ionization mass spectrometry (MALDI MS), and was determined to be near 100%.

5.2.2 Golgi protein GST-fusions

Golgi protein-GST fusion proteins were generated by the introduction of the first 15 codons of *KRE2*, *MNN2*, *MNN5*, and *OCH1* upstream of the GST open reading frame of pGEX-KG (GE Healthcare), and a stop codon before the multiple cloning site. Use of an NcoI cloning site altered codon 2 of *MNN2* and *MNN5* (Leu to Val), and of *OCH1* (Ser to Ala). Dr. David Katzmann (Mayo Research Foundation) provided the Cps1-GST fusion plasmid.

GST fusion proteins were expressed in *E. coli* BL21 (DE3) (Novagen) by induction with IPTG overnight at 22ºC. Cell were lysed in PBS containing 1X Complete Mini protease inhibitor cocktail (Roche) using lysozyme and sonication. Clarified lysate was stored at -80º C.

5.2.3 GOLPH3

GOLPH3 was expressed either from a pET-28a vector (EMD) to produce N-terminally H_6 and T7 tagged proteins (T7-GOLPH3) or from a pET-21a vector (EMD) modified to introduce an N-terminal H_6 tag followed by a tobacco etch virus (TEV) protease cleavage site [228] (H_6-GOLPH3). PCR methods were used to generate vectors to express altered forms of GOLPH3.

To produce pure GOLPH3, the relevant protein was expressed in BL21(DE3) (EMD; T7-GOLPH3) or Rosetta2(DE3)pLysS *E. coli* (EMD; H_6-GOLPH3). Overexpression was induced by the addition of 0.5 mM IPTG at a culture density of 0.6 OD, for 3 h at 37°C. Selenomethionine labeled H_6-GOLPH3Δ51 (SeMet-GOLPH3Δ51), lacking the N-terminal 51 amino acids of GOLPH3, was expressed in B834(DE3) *E. coli* (EMD) in simplified minimal medium [229], as described for Vps74.

GOLPH3 exhibited a tendency to precipitate at various points during its purification, particularly at high protein concentration. Solubility was improved by the use of buffers ≥ pH 8.5, > 150 mM NaCl, and ample reducing agent where practical throughout the purification.

Cells were lysed by sonication (3 rounds of 25 s total, 50 % amplitude, 1 s on/2 s off) in 25 mM Na_2HPO_4/NaH_2PO_4, 250 mM NaCl, 10% glycerol, pH 8.0. Whole lysates were clarified by 45 min centrifugation at 35,000 rcf in a SS-34 rotor. Clarified lysates were applied to a Ni-NTA (Qiagen) affinity column by gravity flow. Protein was eluted by a step gradient of increasing imidazole concentration (fractions of 25, 50, 100, 150, 200 mM imidazole in 25 mM Na_2HPO_4/NaH_2PO_4, 250 mM NaCl, pH 8.0). β-mercaptoethanol (BME) was added to each imidazole fraction upon elution, to a final concentration of 5 mM. TEV protease (150 µg / 5 mg GOLPH3) was added to eluted H_6-GOLPH3, and the proteins were dialyzed for 2 h at RT, then overnight at 4°C, against 50mM Tris, 250mM NaCl, 5 mM BME, 1 mM EDTA, pH 8.0. The

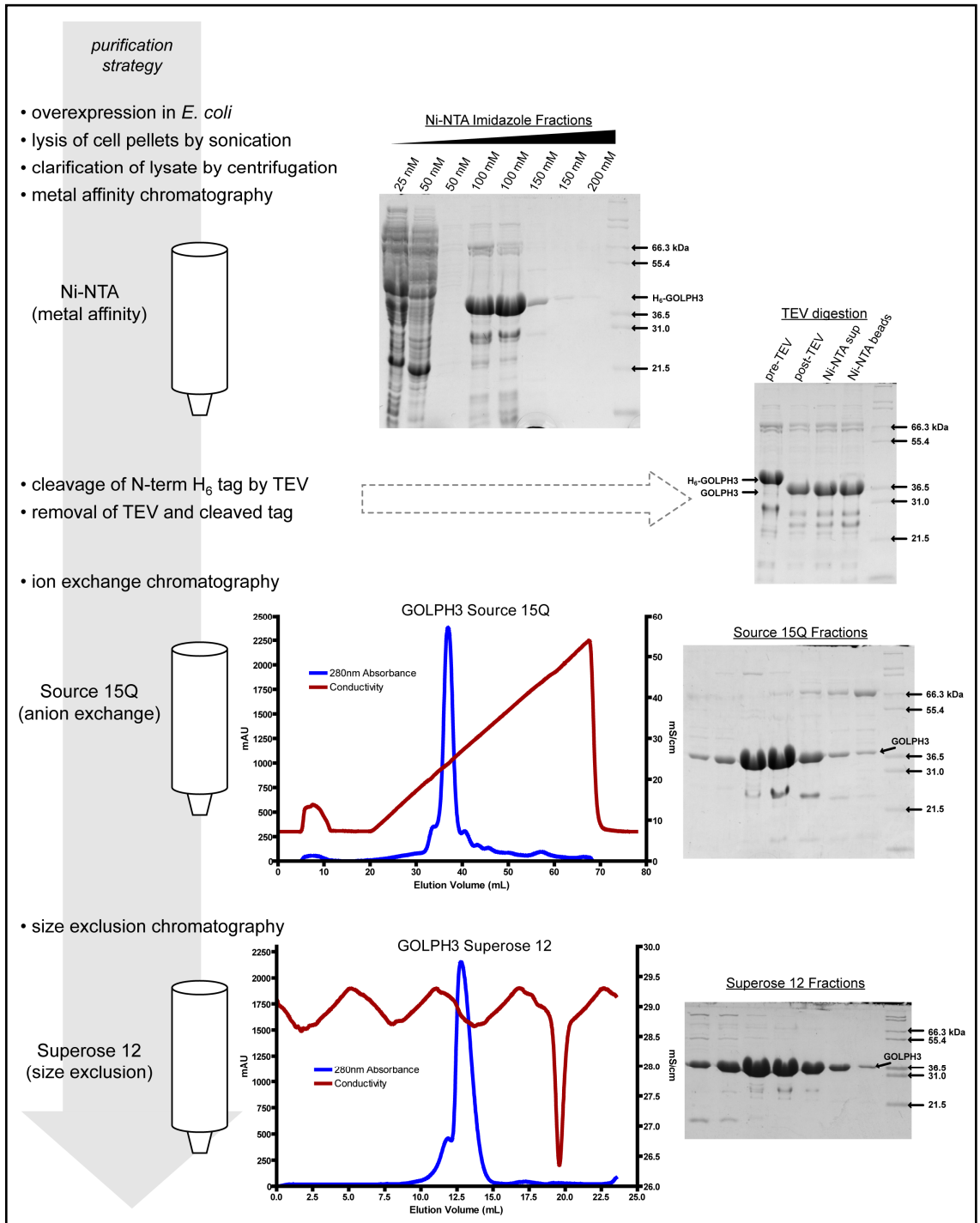

Figure 83. GOLPH3 purification.

Representative FPLC traces and Coomassie stained SDS-PAGE gels from the purification of full length GOLPH3.

cleaved H_6-tag, uncleaved H_6-GOLPH3, and TEV (which is also his-tagged) were removed by incubation with 50 µL of TALON resin (Clontech). TEV-digested GOLPH3 or undigested T7-GOLPH3 were purified by anion exchange chromatography (Source 15Q, GE Healthcare) and were eluted by a 5 column volume linear gradient of 50 mM to 300 mM NaCl in 10 mM TAPS, 5 mM BME, pH 9.0. Proteins were further purified by size exclusion chromatography (Superose 12, GE Healthcare) in 10 mM bicene, 300 mM NaCl, 5 mM BME, pH 8.5. Purified protein was concentrated to 10 mg/mL, buffer exchanged in a spin concentrator (Vivaspin; Sartorius) into buffer lacking reducing agent, and stored in a buffer of 10 mM bicene, 300 mM NaCl, 5 mM TCEP, pH 8.5.

5.2.4 sEGFR

DNA encoding the soluble extracellular region of EGFR, comprising upstream signal sequence as well as codons 1-618 of the mature protein, was cloned into pFastBac I (Invitrogen) as described [230]. The following sEGFR variants were generated (residue numbers refer to those of mature EGFR): sEGFR, amino acids 1-618; sEGFR501, amino acids 1-501; sEGFRd1, amino acids 1-183; sEGFRd3, amino acids 1-4 (LEEK) fused to amino acids 310-514. sEGFR point mutants were generated by standard PCR methods.

sEGFR proteins were expressed as secreted soluble protein by baculovirus infection of Sf9 cells (clonal isolates of fall armyworm, *Spodoptera frugiperda*, ovarian cells). Except where indicated, cells were maintained as suspension cultures at 27°C in Sf-900 II serum free medium (Invitrogen) supplemented with 50 U/mL penicillin, 50 µg/mL streptomycin (Gibco). Baculovirus and Sf9 methods were adapted from the BacToBac system (Invitrogen).

Vectors were transformed into DH10Bac *E. coli* (Invitrogen) by heat shock, followed by 5 h incubation in LB at 37°C. Transformants were incubated on LB agar plates containing 50 µg/mL kanamycin, 7 µg/mL gentamicin, 10 µg/mL tetracycline, 100 µg/mL Bluo-gal (Invitrogen), and 40 µg/mL IPTG (Acros Organics) for 24 h at 37°C, followed by >48 h at room temperature. Bacmid recombination was identified by white colony color. Positive colonies were subcultured

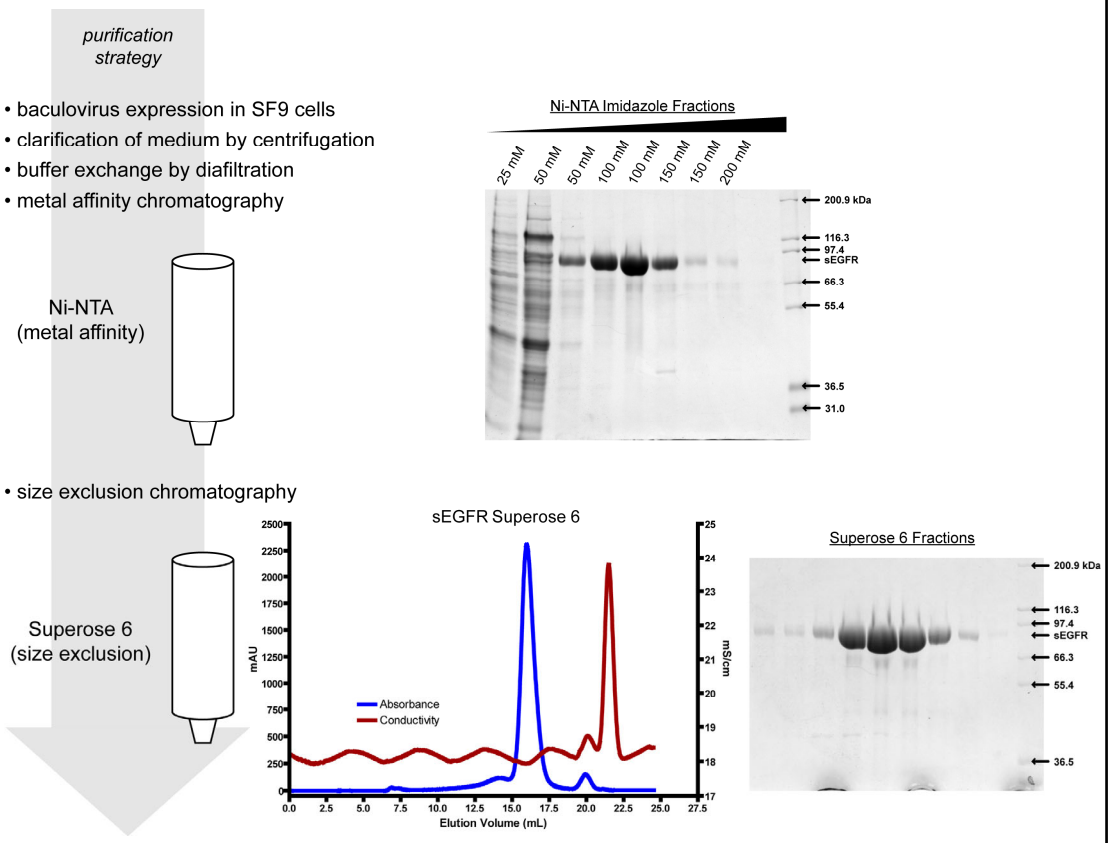

Figure 84. EGFR purification.

Representative FPLC trace and Coomassie stained SDS-PAGE gels from the purification of wild type sEGFR.

into 5 mL LB with antibiotics, and incubated overnight. Bacmid was purified with a Plasmid Mini kit (Qiagen). The presence of the expression sequence was confirmed by PCR.

6-well plates were seeded with 9×10^5 Sf9 cells per well and incubated for 1 h at 27°C to promote adhesion. For each bacmid construct, a bacmid dilution (10 µL of purified bacmid, 100 µL unsupplemented medium) was mixed with a dilution of Cellfectin (Invitrogen) (6 µL Cellfectin, 100 µL unsupplemented medium), and incubated for 30 min at room temperature. Wells were washed with 2 mL unsupplemented medium. 800 µL unsupplemented medium was added to each bacmid/Cellfectin mixture, and the resulting solution (≈ 1 mL) was applied to well of cells. Plates were incubated overnight, after which the medium was replaced with Sf-900 II medium containing penicillin/streptomycin. Plates were incubated for 6 – 7 days at 27°C.

Medium containing baculovirus was harvested and clarified by 5 min 1000 rpm centrifugation. Fetal bovine serum (2% final concentration) was added to the recovered supernatant (P0 virus isolate), and the resulting primary virus isolate was stored at 4°C. Cell debris from centrifugation were analyzed by immunoblot with Tetra-His antibody (Qiagen) to confirm the presence of expressed protein. Virus was amplified by the infection of a T75 flask (TPP) containing 10 million cells in 15 mL of medium with 150 µL of primary virus isolate. After 6 days of infection, medium was clarified by 5 min centrifugation at 1000 rpm. Fetal bovine serum (2% final concentration) was added to the clarified medium, and the resulting P1 virus stock was stored at 4°C.

For large scale protein production, P2 virus was generated similarly to P1 stocks, but by infection of T75 flasks with 150 µL of P1 stock. Sf9 suspension cultures in sidearm bioreactor flasks (0.5 L culture volume per 1 L flask; Cytostir, Kontes) were infected at $2.5 – 3\times10^6$ cells/mL with 15 mL P2 virus per 500 mL culture. Four days after infection, medium was clarified by 20 min 5000 rcf centrifugation. 0.5 mM phenylmethanesulphonylfluoride (PMSF) was added, and clarified medium was diafiltered against > 3 volumes of 25 mM Tris, 150 mM NaCl, pH 8.0, using a 30 kDa or 10 kDa (for sEGFRd3, sEGFRd1, and sEGFR239) Prep-Scale-TFF cartridge (Millipore). Conditioned medium was applied to a Ni-NTA column, and protein was eluted by a

step gradient of imidazole (25, 50, 100, 150, 200 mM) in 25 mM Na_2HPO_4/NaH_2PO_4, 250 mM NaCl, pH 8.0. sEGFRd3 only was dialyzed into 10 mM Hepes, 25 mM NaCl, pH 8, and purified over a Source 15S cation exchange column (GE Healthcare) by a 5 column volume linear gradient of 25 – 500 mM NaCl, 10 mM HEPES, pH 8. All sEGFR constructs were purified by size exclusion chromatography over a Superose 6 (GE Healthcare; sEGFR, sEGFR501) or Superose 12 (GE Healthcare; sEGFRd3, sEGFRd1, sEGFR239) column in 10 mM HEPES, 150 mM NaCl, pH 8. Purified proteins were concentrated to approximately 10 mg/mL in size exclusion chromatography buffer.

5.2.5 VHH antibody fragments

DNA of VHH antibody constructs 7D12, IA1, and 9G8 was obtained from R. Roovers, PhD (Utrecht University, Netherlands) and cloned into pET-22b (EMD). Protein for crystallization included a C-terminal hexa-histidine tag (VHH-H_6) immediately preceding the stop codon. Protein for binding experiments included a C-terminal biotinylation sequence (GLNDIFEAQKIEWH) [231] recognized by *E. coli* biotin holoenzyme synthetase, BirA, followed by a hexa-histidine tag (VHH-AVI). An upstream pelB leader sequence directed periplasmic expression. Variants wit site directed alterations were generated by standard PCR methods.

VHH proteins were overexpressed by overnight lactose auto-induction [232, 233] at 25°C in LB supplemented with 12.5 mM Na_2HPO_4, 12.5 mM KH_2PO_4, 6.25 mM $(NH_4)_2SO_4$, 5 mM Mg_2SO_4, 10 mM MOPS, 0.5% glycerol, 0.5% lactose, 0.5% disodium succcinate, and 0.05% glucose. Cells were pelleted, resuspended in 25 mM Na_2HPO_4/NaH_2PO_4, 150 mM NaCl, 10% glycerol, pH 8.0, and frozen. Resuspened pellets were thawed, cells were repelleted, and the supernatant was further clarified by 30 min 15000 rcf centrifugation. Clarified supernatant was applied to a Ni-NTA column and eluted by a step gradient of imidazole (25, 50, 100, 150, 200 mM) in 25 mM Na_2HPO_4/NaH_2PO_4, 250 mM NaCl, pH 8.0. Protein was further purified over a Superose 12 size exclusion column (GE Healthcare) in 10 mM HEPES, 150 mM NaCl, pH 8.0, at room temperature. Purified protein was concentrated to 5 – 20 mg/mL. VHH proteins exhibit

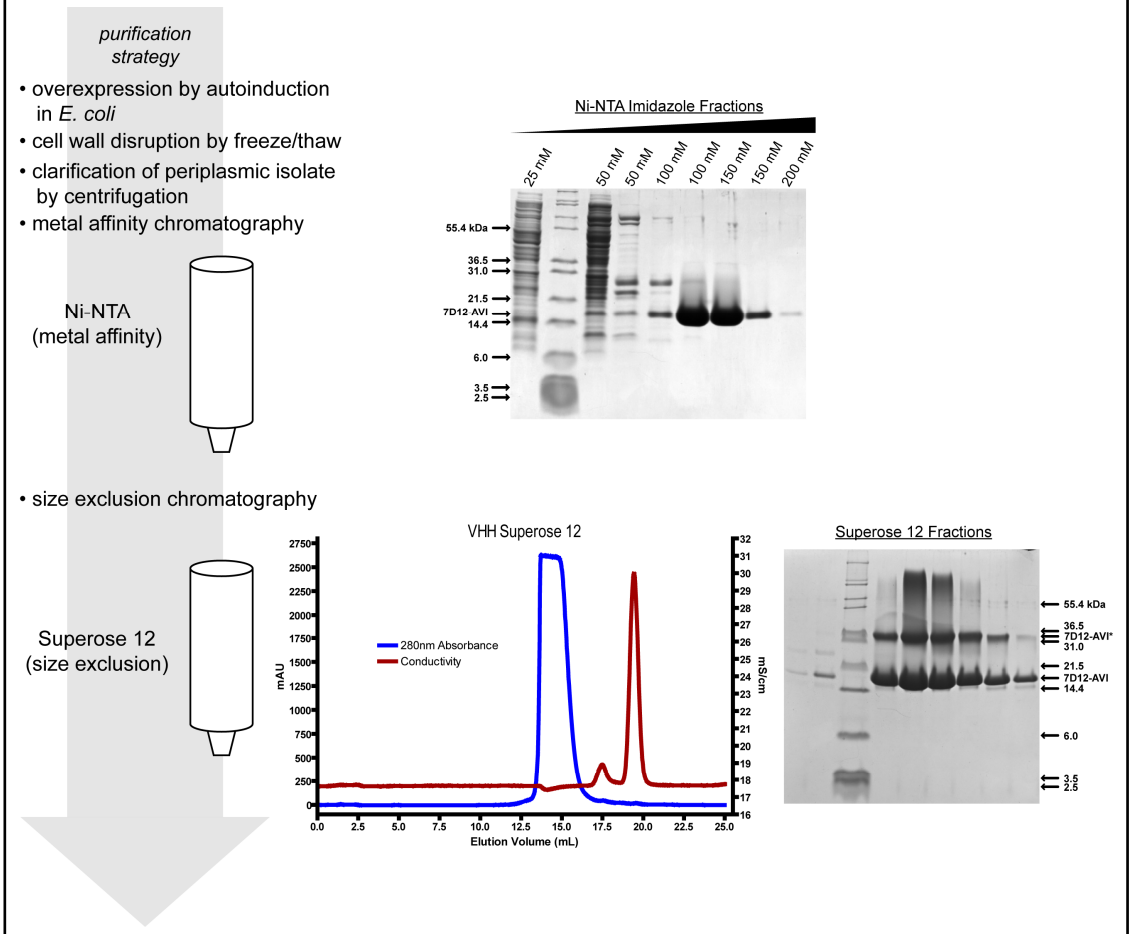

Figure 85. VHH purification.

Representative FPLC trace and Coomassie stained SDS-PAGE gels from the purification of VHH 7D12-AVI. VHH fragments exhibit high thermal stability. VHH constructs frequently produced high molecular weight smears and disulfide bonded artifacts on gels (7D12-AVI* on the lower gel).

cold-induced insolubility at high concentrations, and purified protein was stored at approximately 20°C.

5.2.6 Monoclonal antibody fragments

Monoclonal antibody C225 and FabC225 were provided by Imclone, Inc. Monoclonal antibody 13A9 was provided by Genentech, Inc. Monoclonal antibody 425 was a generous gift from Prof. Ulrich Rodeck (Jefferson University), and monoclonal antibody 108 from Profs. Joseph Schlessinger and Irit Lax (Yale University). The Fab fragments from mAb 425, mAb 108 and mAb 13A9 were prepared by papain cleavage and protein A purification using the Pierce Fab Preparation Kit (Thermo Scientific) and used without further purification. For C225 and 425, where the precise amino acid sequence is known, Fab concentrations were determined using calculated molar extinction coefficients. For 108 and 13A9, molar extinction coefficients were estimated based on protein concentrations determined using Bradford protein assays (BioRad) calibrated with standard curves determined using bovine gamma globulin.

5.3 Crystallographic methods

All crystals were grown by hanging drop vapor diffusion at 22°C.

5.3.1 Vps74

0.5 µL of 10 mg/mL purified Vps74 or SeMet-Vsp74 was mixed with an equal volume of 8% PEG3350, 5% ethylene glycol, 50 mM MES, 50 mM NaCl, 50 mM $CaCl_2$, 10mM EGTA, pH 6.0 (crystallization buffer), and suspended over 300 µL crystallization buffer. Hexagonal crystals (75 × 75 × 300 µm) (Fig. 4) formed and grew to full size in approximately two weeks, and were cryo-stabilized by brief exposure to crystallization buffer supplemented with 15% ethylene glycol, and flash frozen in liquid nitrogen. Vps74Δ59 crystals were generated similarly, but grown in 5% glycerol, 15% PEG3350, 100 mM citrate, 0.75 M KCl, 50 mM EGTA, pH 6.0, and stabilized in this buffer containing 15% glycerol.

X-ray diffraction data were collected at CHESS F1 and APS ID-23 D beamlines. The presence of selenium in SeMet-Vps74 was verified at APS ID-23 D by an excitation scan (intensity as a function of emission wavelength). Selenium peak and inflection wavelengths were chosen, and anomalous-scattering factors f' and f'' were estimated, on the basis of a fluorescence scan (fluorescence intensity as a function of excitation wavelength), using the program CHOOCH [234]. Low beam attenuation and relatively long exposure times were required to collect highest resolution data, which led to radiation induced damage of crystals and loss of diffraction. Consequently, we were unable to collect three wavelength anomalous data on a single SeMet-Vps74 crystal. Peak and remote datasets were collected from a single crystal and the inflection dataset from a second crystal.

Datasets were indexed, integrated, and scaled with HKL-2000 [235]. The three anomalous datasets were additionally scaled in the program Fshcal [8, 236]. The protein substructure was determined using MAD methods in the program SHELX C/D/E [237-239]. All but the first 61 amino acids of Vps74 could be traced in the resulting experimentally phased map using the program Coot [240]. The structure was refined against data from native Vps74 and Vps74Δ59 crystals using the programs Refmac [240] and CNS [27]. Crystallographic statistics are reported in Table 1. The coordinates have been submitted to the PDB: Vps74 ID 2ZIH, Vps74Δ59 ID 2ZII.

5.3.2 GOLPH3

Drops containing 0.5 µL of 10 mg/mL SeMet-GOLPH3Δ51 and an equal volume of crystallization buffer (0.9 M $(NH_4)_2SO_4$, 0.1 M MES, pH 6.0) were suspended over 300 µL of crystallization buffer. Crystals (200 × 200 × 200 µm) (Fig. 18) were transferred to crystallization buffer containing 15% ethylene glycol and were flash frozen in liquid nitrogen.

X-ray diffraction data were collected at APS beamline 23-ID B. Data were processed with HKL-2000 [235]. MAD data were collected and an initial low resolution (3.6 Å) experimental electron density map was determined. The structure was solved by molecular replacement with a complete dataset collected to 2.9 Å resolution (Table 3) in the program Phaser [241], using

chain A from the structure of Vps74 (PDB ID: 2ZIH) as the search model. The GOLPH3Δ51 model was built in Coot [240], and refined using REFMAC [8] and CNS [27]. Correct tracing of features of GOLPH3Δ51 distinct from the yeast homolog was confirmed by the low resolution experimental map (data not shown). Crystallographic statistics are summarized in Table 3. The coordinates have been submitted to the PDB (ID 3KN1).

5.3.3 VHH IA1

0.5 µL of 18 mg/mL purified VHH IA1-H_6 was combined with 1.0 µL drops of crystallization buffer (30% PEG3350, 0.2 M $(NH_4)_2SO_4$, 0.1 M MES, pH 6.0), and was suspended over 300 µL crystallization buffer. Rod shaped crystals (100 × 100 × 2000 µm) (Fig. 55) were flash frozen in liquid nitrogen.

X-ray diffraction data were collected at APS beamline 23-ID B. Data were processed with HKL-2000 [235]. The structure was solved by molecular replacement in the program Phaser [241], using the framework region of a hapten binding camelid VHH domain (PDB ID: 1I3V) as the search model. The VHH IA1 model was built in Coot [240], and refined using REFMAC [8] and CNS [27]. Data collection and refinement statistics are summarized in Table 9.

5.3.4 VHH 7D12/sEGFRd3

The complex between VHH 7D12-H_6 and sEGFRd3 crystallized in two crystal forms, in distinct growth conditions.

The first form developed in drops containing 0.5 µL of 10 mg/mL protein complex and 0.5 µL of crystallization buffer (22.5% PEG3350, 50 mM KI, 0.1M MES, pH 6.0), suspended over 300 µL crystallization buffer. Hexagonal rod shaped crystals (100 × 50 × 50 µm; Fig. 59) were transferred to crystallization buffer containing 12.5% glycerol and were flash frozen in liquid nitrogen.

X-ray diffraction data were collected at APS beamline 23-ID B. Data were processed with HKL-2000 [235]. The structure was determined by molecular replacement in the program Phaser [241], using as search models sEGFRd3 from the complex structure of EGFR domain III and Fab11F8 (PDB ID: 3B2U, chain A) and the VHH IA1 structure framework region. The protein

model was built in Coot [240], and refined using REFMAC [8] and CNS [27]. The presence of an iodide ion was confirmed by inspection of an anomalous difference map.

A second crystal form grew in drops containing 1.0 µL of 10 mg/mL protein complex and 1.5 µL of crystallization buffer (22.5% PEG3350, 0.1M sodium citrate, pH 3.5), suspended over 300 µL of crystallization buffer. Crystals (100 × 100 × 200 µm; Fig. 59) were transferred to crystallization buffer containing 5% ethylene glycol and were flash frozen in liquid nitrogen.

X-ray diffraction data were collected at APS beamline 23-ID B. Data were processed with HKL-2000 [235]. The structure was determined by molecular replacement in the program Phaser [241], using the pH 6.0 VHH 7D12/sEGFRd3 structure as the search model. The protein model was built in Coot [240] and refined using REFMAC [8] and CNS [27]. Six copies of the VHH 7D12/sEGFR complex are observed in the asymmetric unit. Data collection and refinement statistics for both crystal forms are summarized in Table 9.

5.3.5 VHH IA1/Fab C225/sEGFR

The ternary complex VHH IA1-H_6/FabC225/sEGFR formed in drops containing 0.5 µL of 11 mg/ml protein and 0.5 µL of crystallization buffer (17.5% P3350, 1.5M NaCl, 5% glycerol, 0.1 M MES, pH 6.5) suspended over 300 µL of crystallization buffer. Rod shaped crystals (75 × 75 × 1000 µm; Fig. 65) were transferred to crystallization buffer containing 20% PEG3350 and 7.5% glycerol and were flash frozen in liquid nitrogen.

X-ray diffraction data were collected at CHESS beamline F1. Data were processed with HKL-2000 [235]. The structure was determined by molecular replacement in the program Phaser [241], using as independent search models the structure of VHH IA1, the C225 Fv region, the C225 CH2/CL region, EGFR domain I - II, and EGFR domains III - IV (cetuximab and sEGFR models were derived from PDB ID: 1YY9). The protein model was built in Coot [240] and refined using REFMAC [8] and CNS [27]. Crystal packing is extensively stabilized through contacts made to Fab C225 and EGFR domain III. Consequently, Fab C225 and EGFR domain III are observed to be very well ordered, as are the C-terminal portion of domain II, most of domain IV, and the regions of VHH IA1 that contact EGFR domain III (particularly CDRs 1 and 3).

EGFR domain 1 projects into a solvent cavity in the crystal lattice and makes contacts only with VHH IA1 from a symmetry related copy of the complex. As a result, EGFR domain 1, the N-terminal portion of domain II, and the distal portion of VHH IA1 are poorly ordered in the crystal structure. These regions have been built as predominantly backbone atoms or are entirely omitted in the model. Data collection and refinement statistics are summarized in Table 9.

5.3.6 VHH 9G8/Fab C225/sEGFR

The ternary complex VHH 9G8-H_6/Fab C225/sEGFR formed in drops containing 0.5 uL of 7 mg/ml protein and 2.5 µL of crystallization buffer (10% P3350, 0.1 M HEPES, pH 7.0) suspended over 300 µL of crystallization buffer. Rod shaped crystals (100 × 100 × 1000 µm; Fig. 73) were transferred to crystallization buffer containing 15% PEG3350 and 15% ethylene glycol and were flash frozen in liquid nitrogen.

X-ray diffraction data were collected at APS beamline ID 23-B. Data were processed with HKL-2000 [235]. The structure was determined by molecular replacement in the program Phaser [241], using the VHH IA1/Fab C225/sEGFR structure as a search model. The protein model was built in Coot [240] and refined using REFMAC [8] and CNS [27]. The space group, unit cell, and crystal packing of the 9G8 complex are essentially identical to the IA1 ternary complex. As in the IA1 complex, Fab C225 and EGFR domain III are particularly well ordered while EGFR domain 1 and the distal portion of VHH 9G8 are poorly ordered or absent in the final model. Data collection and refinement statistics are summarized in Table 9.

5.3.7 Structure analysis

Shape complementarity values were determined by the SC module in CCP4, excluding solvent molecules. Interactions between chains were identified by NCOTACT in CCP4, using a distance cutoff of 3.5 Å for hydrogen bonds and van der Waals contacts, and 4.5 Å for electrostatic interactions. Average excluded surface areas were calculated by CONTACT in CCP4. The fraction of buried hydrophobic surface area was determined by dividing the surface area of buried carbon atoms by the total. Structural alignments and RMSD values were

generated within the programs Pymol or Coot, and reflect main chain atoms only, unless otherwise specified.

5.4 Analytical ultracentrifugation

Sedimentation velocity analytical ultracentrifugation (SV AUC) experiments were conducted in a Beckman Optima XL-A instrument in a An-Ti 60 rotor at 20°C. Sample absorbance was monitored at 280 nm, unless specified.

Transient rotor temperature fluctuations were occasionally observed at the beginning of a run. The resulting convection currents manifest as residual fitting errors near the top of the solution column for the first few scans, which is correlated with an artifact at the minimum of the c(S) distribution. This artifact has been truncated from the reported c(S) distributions, for clarity. Otherwise, only fits producing low residuals with no systematic error are reported.

5.4.1 Oligomeric state of Vps74 and GOLPH3

The oligomeric state of Vps74 or GOLPH3 proteins was assayed by SV AUC at 40,000 or 50,000 rpm, in a buffer of 150 mM NaCl, 20 mM HEPES, pH 7.5, with and without 10 mM EGTA and 10 mM $CaCl_2$. The absorbance of 50 μM Vps74 (Fig. 12) was monitored at 250 nm, rather than 280 nm, in order to acquire signal in the linear range of the instrument. Readings were scaled by a factor of 2.1 to normalize the signal to readings taken at 280 nm (scale factor determined from UV absorption spectra).

Buffer density, buffer viscosity, and the partial specific volume of sedimenting species were estimated by the program Sednterp [117]. Size distribution c(S) analysis of SV AUC was performed in the program Sedfit [242]. Approximate sedimentation coefficients were determined from crystal structures by the program Hydropro [243].

The heterogeneity of Vps74 and GOLPH3 oligomeric species, apparent from SV AUC (Fig. 13) and other studies (DLS and size exclusion chromatography, data not shown), precluded the use of sedimentation equilibrium methods to determine an accurate K_D for oligomer assembly.

5.4.2 sEGFR interactions with antibodies

VHH fragments, Fab fragments, sEGFR constructs, and combinations of these proteins were analyzed by SV AUC at speeds of 35,000 to 50,000 rpm, depending of the mass of the expected species, in 10mM HEPES, 150 mM NaCl, pH 8.0.

Buffer density and buffer viscosity were estimated in the program Sednterp [117]. The partial specific volume of VHHs and Fab fragments was estimated by Sednterp on the basis of amino acid sequence. The partial specific volumes of sEGFR constructs were estimated by Sednterp on the basis of amino acid sequence, with the additional assumption that the protein carries 8 N-linked glycans of 8 saccharides each. For mixtures of species with markedly different estimates of partial specific volumes, an intermediate value was chosen for analysis. Size distribution c(S) analysis was performed with the program Sedfit [242].

5.5 Surface plasmon resonance (Biacore) studies

Surface plasmon resonance (SPR) binding experiments were performed with a Biacore 3000 instrument at 25°C.

5.5.1 Vps74/GOLPH3 binding to lipid

Lipid vesicles containing 3% PtdIns3*P*, PtdIns4*P*, PtdIns5*P*, or PtdIns(4,5)P_2 (Cell Signaling Technology) in a DOPC (Sigma-Aldrich) background were prepared using cycles of freeze/thaw with sonication followed by extrusion through a 0.1 μm membrane (Mini-Extruder, Avanti). Lipid vesicles were immobilized on a primed L1 biosensor chip (Biacore) [91]. Experiments were conducted in HBS buffer without surfactant (25 mM HEPES, 150 mM NaCl, pH 7.5). Binding response values are reported after reference (DOPC only surface binding) subtraction. Data were analyzed using Prism (version 4; GraphPad Software, Inc.).

5.5.2 sEGFR binding to immobilized antibodies

Monoclonal antibodies and Fab fragments were diluted to 50 μg/mL in acidic buffer and were immobilized to CM5 Biacore sensor chips by amine coupling, at a flow rate of 5 μL/min

with a 5 – 10 minute contact time. Optimal coupling was obtained in 10 mM sodium acetate at pH 4.0 for Fab425 and Fab2E9, pH 5.0 for all mAbs and for Fab13A9, and pH 5.5 for FabC225 and Fab108. Immobilization typically resulted in a response increase of 1000 – 1500 for mAbs and 1500 – 5000 for Fab fragments.

Binding of sEGFR to antibody surfaces was determined as described [5, 230]. A buffer containing 10 mM Hepes, 150 mM NaCl, 3 mM EDTA, 0.005% Tween 20, pH 8.0 (HBS) was used. 250 µL samples were passed over immobilized antibody surfaces to ensure equilibration. Response values at equilibrium were recorded, after subtraction of reference signal. Surfaces were regenerated between data points with two 5 µL injections of 1 M NaCl at low pH (2.5 – 5.5) to rapidly remove residual sEGFR. For each combination of surface and regeneration buffer, it was verified that multiple cycles of regeneration did not impair sEGFR binding. A minimum of 12 samples of a two-fold dilution series of sEGFR or variants was analyzed in order of increasing concentration. The concentration range of samples was chosen to bracket the expected K_D of the interaction, up to a maximum concentration of 4 µM. Data were analyzed using Prism 4 (GraphPad Software, Inc.).

5.5.3 sEGFR binding to immobilized VHHs

Purified VHH-AVI constructs (not exogenously biotinylated) were diluted to 100 µg/ml in 10 mM acetate pH 5.0 and were coupled to CM5 Biacore sensor chips by standard amine coupling, according to the manufacturer's recommendations. Alternatively, purified VHH-AVI constructs were exogenously biotinylated by BirA [231, 233] and immobilized to streptavidin coated SA Biacore sensor chips. Binding of sEGFR to immobilized VHHs was performed as for binding to monoclonal antibodies and Fab fragments, except for regeneration conditions. VHH surfaces were found to be damaged by repeated regenerations below pH 3.5, so bound sEGFR was either left to dissociate from VHH surfaces during 1 h dissociation periods, or surfaces were briefly regenerated by a 5 µL pulse of 10 mM glycylglycine pH 3.5, 1 M NaCl. Data were analyzed as for binding to immobilized antibody.

5.5.4 sEGFR binding to immobilized ligands

Ligands EGF and TGFα were diluted to 200 µg/ml in 10 mM acetate pH 4.0 and coupled to CM5 Biacore sensor chips by standard amine coupling, according to the manufacturer's recommendations. Typical response increase after immobilization was 1000 – 1500. Binding of sEGFR to these immobilized ligands was performed and analyzed as described [230]. Data were analyzed as for binding to immobilized antibody.

Bibliography

1. Garrett, T.P., et al., *Crystal structure of a truncated epidermal growth factor receptor extracellular domain bound to transforming growth factor alpha.* Cell, 2002. **110**(6): p. 763-73.
2. Ogiso, H., et al., *Crystal structure of the complex of human epidermal growth factor and receptor extracellular domains.* Cell, 2002. **110**(6): p. 775-87.
3. Garrett, T.P., et al., *Antibodies specifically targeting a locally misfolded region of tumor associated EGFR.* Proc Natl Acad Sci U S A, 2009. **106**(13): p. 5082-7.
4. Chenna, R., et al., *Multiple sequence alignment with the Clustal series of programs.* Nucleic Acids Res., 2003. **31**(13): p. 3497-500.
5. Li, S., et al., *Structural basis for inhibition of the epidermal growth factor receptor by cetuximab.* Cancer Cell, 2005. **7**(4): p. 301-11.
6. Schmiedel, J., et al., *Matuzumab binding to EGFR prevents the conformational rearrangement required for dimerization.* Cancer Cell, 2008. **13**(4): p. 365-73.
7. Kabat, E.A., et al., *Sequences of proteins of immunological interest*. 1991, Bethesda, MD: National Institutes of Health.
8. CCP4, *The CCP4 Suite: Programs for Protein Crystallography.* Acta Crystallogr., 1994. **D50**: p. 760-763.
9. Freeman, D., et al., *Panitumumab and cetuximab epitope mapping and in vitro activity.* J Clin Oncol (Meeting Abstracts), 2008. **26**(15_suppl): p. 14536-.
10. Lammerts van Bueren, J.J., et al., *The antibody zalutumumab inhibits epidermal growth factor receptor signaling by limiting intra- and intermolecular flexibility.* Proc Natl Acad Sci U S A, 2008. **105**(16): p. 6109-14.
11. May, A.C.W., *Percent Sequence Identity: The Need to Be Explicit.* Structure, 2004. **12**(5): p. 737-738.
12. DeLano, W.L., *The PyMOL Molecular Graphics System*. 2004, DeLano Scientific Palo Alto, CA, USA.

13. Baker, N.A., et al., *Electrostatics of nanosystems: application to microtubules and the ribosome.* Proc. Natl. Acad. Sci. U.S.A., 2001. **98**: p. 10037-10041.

14. Dawson, J.P., Z. Bu, and M.A. Lemmon, *Ligand-induced structural transitions in ErbB receptor extracellular domains.* Structure, 2007. **15**(8): p. 942-54.

15. Abhinandan, K.R. and A.C. Martin, *Analysis and improvements to Kabat and structurally correct numbering of antibody variable domains.* 2008(0161-5890 (Print)).

16. Westhead, D.R., et al., *Protein structural topology: Automated analysis and diagrammatic representation.* Protein Sci, 1999. **8**(4): p. 897-904.

17. Ferguson, K.M., et al., *EGF activates its receptor by removing interactions that autoinhibit ectodomain dimerization.* Mol Cell, 2003. **11**(2): p. 507-17.

18. Macdonald, J.L. and L.J. Pike, *Heterogeneity in EGF-binding affinities arises from negative cooperativity in an aggregating system.* Proc Natl Acad Sci U S A, 2008. **105**(1): p. 112-7.

19. Schmitz, K.R., et al., *Golgi localization of glycosyltransferases requires a Vps74p oligomer.* Dev Cell, 2008. **14**(4): p. 523-34.

20. Potapov, V., et al., *Protein--protein recognition: juxtaposition of domain and interface cores in immunoglobulins and other sandwich-like proteins.* J Mol Biol, 2004. **342**(2): p. 665-79.

21. Dawson, J.P., et al., *Epidermal growth factor receptor dimerization and activation require ligand-induced conformational changes in the dimer interface.* Mol Cell Biol, 2005. **25**(17): p. 7734-42.

22. Cho, H.S. and D.J. Leahy, *Structure of the extracellular region of HER3 reveals an interdomain tether.* Science, 2002. **297**(5585): p. 1330-3.

23. Bouyain, S., et al., *The extracellular region of ErbB4 adopts a tethered conformation in the absence of ligand.* Proc Natl Acad Sci U S A, 2005. **102**(42): p. 15024-9.

24. Cho, H.S., et al., *Structure of the extracellular region of HER2 alone and in complex with the Herceptin Fab.* Nature, 2003. **421**(6924): p. 756-60.

25. Bellot, F., et al., *High-affinity epidermal growth factor binding is specifically reduced by a monoclonal antibody, and appears necessary for early responses.* J Cell Biol, 1990. **110**(2): p. 491-502.

26. Al-Lazikani, B., A.M. Lesk, and C. Chothia, *Standard conformations for the canonical structures of immunoglobulins.* J Mol Biol, 1997. **273**(4): p. 927-48.

27. Brunger, A.T., et al., *Crystallography & NMR system: A new software suite for macromolecular structure determination.* Acta Crystallogr., 1998. **D54**: p. 905-921.

28. Wood, C.S., et al., *PtdIns4P recognition by Vps74/GOLPH3 links PtdIns 4-kinase signaling to retrograde Golgi trafficking.* J Cell Biol, 2009. **187**(7): p. 967-75.

29. Chao, G., J.R. Cochran, and K.D. Wittrup, *Fine epitope mapping of anti-epidermal growth factor receptor antibodies through random mutagenesis and yeast surface display.* J Mol Biol, 2004. **342**(2): p. 539-50.

30. Dereeper, A., et al., *Phylogeny.fr: robust phylogenetic analysis for the non-specialist.* Nucleic Acids Res, 2008. **36**(Web Server issue): p. W465-9.

31. Franklin, M.C., et al., *Insights into ErbB signaling from the structure of the ErbB2-pertuzumab complex.* Cancer Cell, 2004. **5**(4): p. 317-28.

32. Beznoussenko, G.V. and A.A. Mironov, *Models of intracellular transport and evolution of the Golgi complex.* The Anatomical Record, 2002. **268**(3): p. 226-238.

33. De Genst, E., et al., *Chemical basis for the affinity maturation of a camel single domain antibody.* J Biol Chem, 2004. **279**(51): p. 53593-601.

34. Desmyter, A., et al., *Three camelid VHH domains in complex with porcine pancreatic alpha-amylase. Inhibition and versatility of binding topology.* J Biol Chem, 2002. **277**(26): p. 23645-50.

35. Tereshko, V., et al., *Toward chaperone-assisted crystallography: protein engineering enhancement of crystal packing and X-ray phasing capabilities of a camelid single-domain antibody (VHH) scaffold.* Protein Sci, 2008. **17**(7): p. 1175-87.

36. Pelham, H.R. and J.E. Rothman, *The debate about transport in the Golgi--two sides of the same coin?* Cell, 2000. **102**(6): p. 713-9.

37. Matsuura-Tokita, K., et al., *Live imaging of yeast Golgi cisternal maturation.* Nature, 2006. **441**(7096): p. 1007-1010.

38. Dunphy, W.G. and J.E. Rothman, *Compartmental organization of the Golgi stack.* Cell, 1985. **42**(1): p. 13-21.

39. Glick, B.S., T. Elstron, and G. Oster, *A cisternal maturation mechansim can explain the asymmetry of the Gogli stack.* FEBS Lett., 1997. **414**: p. 177-81.

40. Duden, R., *ER-to-Golgi transport: COP I and COP II function (Review).* Mol Membr Biol, 2003. **20**(3): p. 197-207.

41. Bonifacino, J.S. and J.H. Hurley, *Retromer.* Curr Opin Cell Biol, 2008. **20**(4): p. 427-36.

42. Tu, L. and D.K. Banfield, *Localization of Golgi-resident glycosyltransferases.* Cell Mol Life Sci, 2009.

43. Colley, K.J., *Golgi localization of glycosyltransferases: more questions than answers.* Glycobiology, 1997. **7**(1): p. 1-13.

44. Spiro, R.G., *Protein glycosylation: nature, distribution, enzymatic formation, and disease implications of glycopeptide bonds.* Glycobiology, 2002. **12**(4): p. 43R-56.

45. Tu, L., et al., *Signal-mediated dynamic retention of glycosyltransferases in the Golgi.* Science, 2008. **321**(5887): p. 404-7.

46. Uemura, S., et al., *The cytoplasmic tail of GM3 synthase defines its subcellular localization, stability, and in vivo activity.* Mol Biol Cell, 2009. **20**(13): p. 3088-100.

47. Lemmon, M.A., *Membrane recognition by phospholipid-binding domains.* Nat Rev Mol Cell Biol, 2008. **9**(2): p. 99-111.

48. Di Paolo, G. and P. De Camilli, *Phosphoinositides in cell regulation and membrane dynamics.* Nature, 2006. **443**(7112): p. 651-7.

49. Nicot, A.S. and J. Laporte, *Endosomal phosphoinositides and human diseases.* Traffic, 2008. **9**(8): p. 1240-9.

50. Toker, A. and L.C. Cantley, *Signalling through the lipid products of phosphoinositide-3-OH kinase.* Nature, 1997. **387**(6634): p. 673-6.

51. Blagoveshchenskaya, A., et al., *Integration of Golgi trafficking and growth factor signaling by the lipid phosphatase SAC1.* J Cell Biol, 2008. **180**(4): p. 803-12.

52. Faulhammer, F., et al., *Growth control of Golgi phosphoinositides by reciprocal localization of sac1 lipid phosphatase and pik1 4-kinase.* Traffic, 2007. **8**(11): p. 1554-67.

53. Audhya, A., M. Foti, and S.D. Emr, *Distinct roles for the yeast phosphatidylinositol 4-kinases, Stt4p and Pik1p, in secretion, cell growth, and organelle membrane dynamics.* Mol Biol Cell, 2000. **11**(8): p. 2673-89.

54. Hama, H., et al., *Direct involvement of phosphatidylinositol 4-phosphate in secretion in the yeast Saccharomyces cerevisiae.* J Biol Chem, 1999. **274**(48): p. 34294-300.

55. Demmel, L., et al., *The clathrin adaptor Gga2p is a phosphatidylinositol 4-phosphate effector at the Golgi exit.* Mol Biol Cell, 2008. **19**(5): p. 1991-2002.

56. D'Angelo, G., et al., *The multiple roles of PtdIns(4)P -- not just the precursor of PtdIns(4,5)P2.* J Cell Sci, 2008. **121**(Pt 12): p. 1955-63.

57. Huh, W.K., et al., *Global analysis of protein localization in budding yeast.* Nature, 2003. **425**(6959): p. 686-91.

58. Bell, A.W., et al., *Proteomics characterization of abundant Golgi membrane proteins.* J Biol Chem, 2001. **276**(7): p. 5152-65.

59. Bonangelino, C.J., E.M. Chavez, and J.S. Bonifacino, *Genomic screen for vacuolar protein sorting genes in Saccharomyces cerevisiae.* Mol Biol Cell, 2002. **13**(7): p. 2486-501.

60. Tong, A.H., et al., *Global mapping of the yeast genetic interaction network.* Science, 2004. **303**(5659): p. 808-13.

61. Dippold, H.C., et al., *GOLPH3 bridges phosphatidylinositol-4- phosphate and actomyosin to stretch and shape the Golgi to promote budding.* Cell, 2009. **139**(2): p. 337-51.

62. Scott, K.L., et al., *GOLPH3 modulates mTOR signalling and rapamycin sensitivity in cancer.* Nature, 2009. **459**(7250): p. 1085-90.

63. Lussier, M., et al., *The Ktr1p, Ktr3p, and Kre2p/Mnt1p mannosyltransferases participate in the elaboration of yeast O- and N-linked carbohydrate chains.* J. Biol. Chem., 1997. **272**(24): p. 15527-31.

64. Corbacho, I., I. Olivero, and L.M. Hernandez, *A genome-wide screen for Saccharomyces cerevisiae nonessential genes involved in mannosyl phosphate transfer to mannoprotein-linked oligosaccharides.* Fungal Genet. Biol., 2005. **42**(9): p. 773-90.

65. Lussier, M., A.M. Sdicu, and H. Bussey, *The KTR and MNN1 mannosyltransferase families of Saccharomyces cerevisiae.* Biochim Biophys Acta, 1999. **1426**(2): p. 323-34.

66. Schoenmakers, T.J., et al., *CHELATOR: an improved method for computing metal ion concentrations in physiological solutions.* Biotechniques, 1992. **12**(6): p. 870-4, 876-9.

67. Cole, C., J.D. Barber, and G.J. Barton, *The Jpred 3 secondary structure prediction server.* Nucleic Acids Res, 2008. **36**(Web Server issue): p. W197-201.

68. Ouali, M. and R.D. King, *Cascaded multiple classifiers for secondary structure prediction.* Protein Sci, 2000. **9**(6): p. 1162-76.

69. Snyder, C.M., et al., *GMx33 associates with the trans-Golgi matrix in a dynamic manner and sorts within tubules exiting the Golgi.* Mol. Biol. Cell, 2006. **17**(1): p. 511-24.

70. Wu, C.C., et al., *GMx33: a novel family of trans-Golgi proteins identified by proteomics.* Traffic, 2000. **1**(12): p. 963-75.

71. Holm, L., et al., *A database of protein structure families with common folding motifs.* Prot. Sci., 1992. **1**: p. 1691-1698.

72. Pearl, F., et al., *The CATH Domain Structure Database and related resources Gene3D and DHS provide comprehensive domain family information for genome analysis.* Nucleic Acids Res., 2005. **33**(Database issue): p. D247-51.

73. Murzin, A.G., et al., *SCOP: a structural classification of proteins database for the investigation of sequences and structures.* J. Mol. Biol., 1995. **247**(4): p. 536-40.

74. Bennett-Lovsey, R.M., et al., *Exploring the extremes of sequence/structure space with ensemble fold recognition in the program Phyre.* Proteins, 2007.

75. Aravind, L., et al., *The many faces of the helix-turn-helix domain: transcription regulation and beyond.* FEMS Microbiol. Rev., 2005. **29**(2): p. 231-262.

76. Gajiwala, K.S. and S.K. Burley, *Winged helix proteins.* Curr. Opin. Struct. Biol., 2000. **10**(1): p. 110-6.

77. Hierro, A., et al., *Structure of the ESCRT-II endosomal trafficking complex.* Nature, 2004. **431**(7005): p. 221-5.

78. Teo, H., et al., *ESCRT-II, an endosome-associated complex required for protein sorting: crystal structure and interactions with ESCRT-III and membranes.* Dev. Cell, 2004. **7**(4): p. 559-569.

79. Haering, C.H., et al., *Structure and stability of cohesin's Smc1-kleisin interaction.* Mol. Cell, 2004. **15**(6): p. 951-64.

80. Lo Conte, L., C. Chothia, and J. Janin, *The atomic structure of protein-protein recognition sites.* J. Mol. Biol., 1999. **285**(5): p. 2177-98.

81. Karathanassis, D., et al., *Binding of the PX domain of p47(phox) to phosphatidylinositol 3,4-bisphosphate and phosphatidic acid is masked by an intramolecular interaction.* Embo J, 2002. **21**(19): p. 5057-68.

82. Mao, Y., et al., *Crystal structure of the VHS and FYVE tandem domains of Hrs, a protein involved in membrane trafficking and signal transduction.* Cell, 2000. **100**(4): p. 447-56.

83. Lawrence, M.C. and P.M. Colman, *Shape complementarity at protein-/protein interfaces.* J. Mol. Biol., 1993. **234**(4): p. 946-950.

84. Ferguson, K.M., et al., *Structure of the high affinity complex of inositol trisphosphate with a phospholipase C pleckstrin homology domain.* Cell, 1995. **83**(6): p. 1037-46.

85. Bravo, J., et al., *The crystal structure of the PX domain from p40(phox) bound to phosphatidylinositol 3-phosphate.* Mol Cell, 2001. **8**(4): p. 829-39.

86. Zhou, C.Z., et al., *Crystal structure of the yeast Phox homology (PX) domain protein Grd19p complexed to phosphatidylinositol-3-phosphate.* J Biol Chem, 2003. **278**(50): p. 50371-6.

87. Dumas, J.J., et al., *Multivalent endosome targeting by homodimeric EEA1.* Mol Cell, 2001. **8**(5): p. 947-58.

88. Hendricks, K.B., et al., *Yeast homologue of neuronal frequenin is a regulator of phosphatidylinositol-4-OH kinase.* Nat Cell Biol, 1999. **1**(4): p. 234-41.

89. Konrad, G., et al., *Retention of the yeast Sac1p phosphatase in the endoplasmic reticulum causes distinct changes in cellular phosphoinositide levels and stimulates microsomal ATP transport.* J Biol Chem, 2002. **277**(12): p. 10547-54.

90. Strahl, T. and J. Thorner, *Synthesis and function of membrane phosphoinositides in budding yeast, Saccharomyces cerevisiae.* Biochim Biophys Acta, 2007. **1771**(3): p. 353-404.

91. Narayan, K. and M.A. Lemmon, *Determining selectivity of phosphoinositide-binding domains.* Methods, 2006. **39**(2): p. 122-33.

92. Yu, J.W. and M.A. Lemmon, *All phox homology (PX) domains from Saccharomyces cerevisiae specifically recognize phosphatidylinositol 3-phosphate.* J Biol Chem, 2001. **276**(47): p. 44179-84.

93. Levine, T.P. and S. Munro, *Targeting of Golgi-specific pleckstrin homology domains involves both PtdIns 4-kinase-dependent and -independent components.* Curr Biol, 2002. **12**(9): p. 695-704.

94. Altschul, S.F., et al., *Basic local alignment search tool.* J Mol Biol, 1990. **215**(3): p. 403-10.

95. Cavalier-Smith, T., *Protist phylogeny and the high-level classification of Protozoa.* European Journal of Protistology, 2003. **39**(4): p. 338-348.

96. Cronan, J.E., *Bacterial membrane lipids: where do we stand?* Annu Rev Microbiol, 2003. **57**: p. 203-24.

97. Goldberg, J., *Decoding of Sorting Signals by Coatomer through a GTPase Switch in the COPI Coat Complex.* 2000. **100**(6): p. 671-679.

98. Myers, J.K. and C.N. Pace, *Hydrogen bonding stabilizes globular proteins.* Biophys J, 1996. **71**(4): p. 2033-9.

99. Sheu, S.Y., et al., *Energetics of hydrogen bonds in peptides.* Proc Natl Acad Sci U S A, 2003. **100**(22): p. 12683-7.

100. Scott, K.L. and L. Chin, *Signaling From the Golgi: Mechanisms and Models for Golgi Phosphoprotein 3-Mediated Oncogenesis.* Clin Cancer Res.

101. Xie, M.W., et al., *Insights into TOR function and rapamycin response: chemical genomic profiling by using a high-density cell array method.* Proc Natl Acad Sci U S A, 2005. **102**(20): p. 7215-20.

102. Carpenter, G., *Receptors for Epidermal Growth Factor and Other Polypeptide Mitogens.* Annual Review of Biochemistry, 1987. **56**(1): p. 881-914.

103. Klein, S. and A. Levitzki, *Targeting the EGFR and the PKB pathway in cancer.* Curr Opin Cell Biol, 2009. **21**(2): p. 185-93.

104. Lin, C.R., et al., *Expression Cloning of Human EGF Receptor Complementary DNA: Gene Amplification and Three Related Messenger RNA Products in A431 Cells.* Science, 1984. **224**(4651): p. 843-848.

105. Lemmon, M.A., *Ligand-induced ErbB receptor dimerization.* Exp Cell Res, 2009. **315**(4): p. 638-48.

106. Yarden, Y. and M.X. Sliwkowski, *Untangling the ErbB signalling network.* Nat Rev Mol Cell Biol, 2001. **2**(2): p. 127-37.

107. Schlessinger, J. and M.A. Lemmon, *SH2 and PTB domains in tyrosine kinase signaling.* Sci STKE, 2003. **2003**(191): p. RE12.

108. Zeng, F., A.B. Singh, and R.C. Harris, *The role of the EGF family of ligands and receptors in renal development, physiology and pathophysiology.* Exp Cell Res, 2009. **315**(4): p. 602-10.

109. Morgan, S. and J.R. Grandis, *ErbB receptors in the biology and pathology of the aerodigestive tract.* Exp Cell Res, 2009. **315**(4): p. 572-82.

110. Hynes, N.E. and G. MacDonald, *ErbB receptors and signaling pathways in cancer.* Curr Opin Cell Biol, 2009. **21**(2): p. 177-84.

111. Voldborg, B.R., et al., *Epidermal growth factor receptor (EGFR) and EGFR mutations, function and possible role in clinical trials.* Annals of Oncology, 1997. **8**(12): p. 1197-1206.

112. Wilson, K.J., et al., *Functional selectivity of EGF family peptide growth factors: implications for cancer.* Pharmacol Ther, 2009. **122**(1): p. 1-8.

113. Ross, J.S. and J.A. Fletcher, *The HER-2/neu Oncogene in Breast Cancer: Prognostic Factor, Predictive Factor, and Target for Therapy.* Oncologist, 1998. **3**(4): p. 237-252.

114. Nicholson, R.I., J.M. Gee, and M.E. Harper, *EGFR and cancer prognosis.* Eur J Cancer, 2001. **37 Suppl 4**: p. S9-15.

115. Hynes, N.E. and H.A. Lane, *ERBB receptors and cancer: the complexity of targeted inhibitors.* Nat Rev Cancer, 2005. **5**(5): p. 341-54.

116. Somers, W., et al., *The X-ray structure of a growth hormone-prolactin receptor complex.* Nature, 1994. **372**(6505): p. 478-81.

117. Philo, J.S., et al., *Dimerization of the extracellular domain of the erythropoietin (EPO) receptor by EPO: one high-affinity and one low-affinity interaction.* Biochemistry, 1996. **35**(5): p. 1681-91.

118. Schlessinger, J., *Cell signaling by receptor tyrosine kinases.* Cell, 2000. **103**(2): p. 211-25.

119. Wiesmann, C., et al., *Crystal structure at 1.7 A resolution of VEGF in complex with domain 2 of the Flt-1 receptor.* Cell, 1997. **91**(5): p. 695-704.

120. Bella, J., et al., *The leucine-rich repeat structure.* Cell Mol Life Sci, 2008. **65**(15): p. 2307-33.

121. Burgess, A.W., et al., *An open-and-shut case? Recent insights into the activation of EGF/ErbB receptors.* Mol Cell, 2003. **12**(3): p. 541-52.

122. Ozcan, F., et al., *On the nature of low- and high-affinity EGF receptors on living cells.* Proc Natl Acad Sci U S A, 2006. **103**(15): p. 5735-40.

123. Macdonald-Obermann, J.L. and L.J. Pike, *The intracellular juxtamembrane domain of the epidermal growth factor (EGF) receptor is responsible for the allosteric regulation of EGF binding.* J Biol Chem, 2009. **284**(20): p. 13570-6.

124. Zhou, M., et al., *Real-time measurements of kinetics of EGF binding to soluble EGF receptor monomers and dimers support the dimerization model for receptor activation.* Biochemistry, 1993. **32**(32): p. 8193-8.

125. Mayawala, K., D.G. Vlachos, and J.S. Edwards, *Heterogeneities in EGF receptor density at the cell surface can lead to concave up scatchard plot of EGF binding.* FEBS Lett, 2005. **579**(14): p. 3043-7.

126. Klein, P., et al., *A structure-based model for ligand binding and dimerization of EGF receptors.* Proc Natl Acad Sci U S A, 2004. **101**(4): p. 929-34.

127. Wofsy, C., et al., *Implications of epidermal growth factor (EGF) induced egf receptor aggregation.* Biophys J, 1992. **63**(1): p. 98-110.

128. Schmitz, K.R. and K.M. Ferguson, *Interaction of antibodies with ErbB receptor extracellular regions.* Exp Cell Res, 2009. **315**(4): p. 659-70.

129. Lax, I., et al., *Noncontiguous regions in the extracellular domain of EGF receptor define ligand-binding specificity.* Cell Regul, 1991. **2**(5): p. 337-45.

130. Lenferink, A.E., et al., *The linear C-terminal regions of epidermal growth factor (EGF) and transforming growth factor-alpha bind to different epitopes on the human EGF receptor.* Biochem J, 1998. **336 (Pt 1)**: p. 147-51.

131. Woof, J.M. and D.R. Burton, *Human antibody-Fc receptor interactions illuminated by crystal structures.* Nat Rev Immunol, 2004. **4**(2): p. 89-99.

132. Tao, M.H., R.I. Smith, and S.L. Morrison, *Structural features of human immunoglobulin G that determine isotype-specific differences in complement activation.* J Exp Med, 1993. **178**(2): p. 661-7.

133. Gill, G.N., et al., *Monoclonal anti-epidermal growth factor receptor antibodies which are inhibitors of epidermal growth factor binding and antagonists of epidermal growth factor binding and antagonists of epidermal growth factor-stimulated tyrosine protein kinase activity.* J Biol Chem, 1984. **259**(12): p. 7755-60.

134. Bonner, J.A., et al., *Radiotherapy plus cetuximab for squamous-cell carcinoma of the head and neck.* N Engl J Med, 2006. **354**(6): p. 567-78.

135. Kirkpatrick, P., J. Graham, and M. Muhsin, *Cetuximab.* Nat Rev Drug Discov, 2004. **3**(7): p. 549-550.

136. Murthy, U., et al., *Binding of an antagonistic monoclonal antibody to an intact and fragmented EGF-receptor polypeptide.* Arch Biochem Biophys, 1987. **252**(2): p. 549-60.

137. Seiden, M.V., et al., *A phase II trial of EMD72000 (matuzumab), a humanized anti-EGFR monoclonal antibody, in patients with platinum-resistant ovarian and primary peritoneal malignancies.* Gynecol Oncol, 2007. **104**(3): p. 727-31.

138. Waterfield, M.D., et al., *A monoclonal antibody to the human epidermal growth factor receptor.* J Cell Biochem, 1982. **20**(2): p. 149-61.

139. Defize, L.H., et al., *Signal transduction by epidermal growth factor occurs through the subclass of high affinity receptors.* J Cell Biol, 1989. **109**(5): p. 2495-507.

140. Winkler, M.E., et al., *Epidermal growth factor and transforming growth factor alpha bind differently to the epidermal growth factor receptor.* Biochemistry, 1989. **28**(15): p. 6373-8.

141. Gooi, H.C., et al., *The carbohydrate specificities of the monoclonal antibodies 29.1, 455 and 3C1B12 to the epidermal growth factor receptor of A431 cells.* Biosci Rep, 1985. **5**(1): p. 83-94.

142. Reuter, C.W., M.A. Morgan, and A. Eckardt, *Targeting EGF-receptor-signalling in squamous cell carcinomas of the head and neck.* Br J Cancer, 2007. **96**(3): p. 408-16.

143. Talavera, A., et al., *Nimotuzumab, an antitumor antibody that targets the epidermal growth factor receptor, blocks ligand binding while permitting the active receptor conformation.* Cancer Res, 2009. **69**(14): p. 5851-9.

144. Fernandez, A., et al., *A new monoclonal antibody for detection of EGF-receptors in western blots and paraffin-embedded tissue sections.* J Cell Biochem, 1992. **49**(2): p. 157-65.

145. Giusti, R.M., et al., *FDA review of a panitumumab (Vectibix) clinical trial for first-line treatment of metastatic colorectal cancer.* Oncologist, 2009. **14**(3): p. 284-90.

146. Rivera, F., et al., *Current situation of zalutumumab.* Expert Opin Biol Ther, 2009. **9**(5): p. 667-74.

147. Li, S., P. Kussie, and K.M. Ferguson, *Structural basis for EGF receptor inhibition by the therapeutic antibody IMC-11F8.* Structure, 2008. **16**(2): p. 216-27.

148. Kuenen, B., et al., *A phase I pharmacologic study of necitumumab (IMC-11F8), a fully human IgG1 monoclonal antibody directed against EGFR in patients with advanced solid malignancies.* Clin Cancer Res, 2010. **16**(6): p. 1915-23.

149. Scott, A.M., et al., *A phase I clinical trial with monoclonal antibody ch806 targeting transitional state and mutant epidermal growth factor receptors.* Proc Natl Acad Sci U S A, 2007. **104**(10): p. 4071-6.

150. Liu, M., et al., *Identification and characterization of a fully human antibody directed against epidermal growth factor receptor for cancer therapy.* AACR Meeting Abstracts, 2004. **2004**(1): p. 163-c-.

151. Nahta, R. and F.J. Esteva, *Trastuzumab: triumphs and tribulations.* Oncogene, 2007. **26**(25): p. 3637-43.

152. Cai, Z., et al., *Differential binding patterns of monoclonal antibody 2C4 to the ErbB3-p185her2/neu and the EGFR-p185her2/neu complexes.* Oncogene, 2008. **27**(27): p. 3870-4.

153. Barnes, D.W., *Epidermal growth factor inhibits growth of A431 human epidermoid carcinoma in serum-free cell culture.* J Cell Biol, 1982. **93**(1): p. 1-4.

154. Sato, J.D., et al., *Biological effects in vitro of monoclonal antibodies to human epidermal growth factor receptors.* Mol Biol Med, 1983. **1**(5): p. 511-29.

155. Lu, D., et al., *Simultaneous blockade of both the epidermal growth factor receptor and the insulin-like growth factor receptor signaling pathways in cancer cells with a fully human recombinant bispecific antibody.* J Biol Chem, 2004. **279**(4): p. 2856-65.

156. Garrett, T.P., et al., *The crystal structure of a truncated ErbB2 ectodomain reveals an active conformation, poised to interact with other ErbB receptors.* Mol Cell, 2003. **11**(2): p. 495-505.

157. Fendly, B.M., et al., *Characterization of murine monoclonal antibodies reactive to either the human epidermal growth factor receptor or HER2/neu gene product.* Cancer Res, 1990. **50**(5): p. 1550-8.

158. Takai, N., et al., *2C4, a monoclonal antibody against HER2, disrupts the HER kinase signaling pathway and inhibits ovarian carcinoma cell growth.* Cancer, 2005. **104**(12): p. 2701-8.

159. Agus, D.B., et al., *Targeting ligand-activated ErbB2 signaling inhibits breast and prostate tumor growth.* Cancer Cell, 2002. **2**(2): p. 127-37.

160. Adams, C.W., et al., *Humanization of a recombinant monoclonal antibody to produce a therapeutic HER dimerization inhibitor, pertuzumab.* Cancer Immunol Immunother, 2006. **55**(6): p. 717-27.

161. Hudziak, R.M., et al., *p185HER2 monoclonal antibody has antiproliferative effects in vitro and sensitizes human breast tumor cells to tumor necrosis factor.* Mol Cell Biol, 1989. **9**(3): p. 1165-72.

162. Kumar, R., H.M. Shepard, and J. Mendelsohn, *Regulation of phosphorylation of the c-erbB-2/HER2 gene product by a monoclonal antibody and serum growth factor(s) in human mammary carcinoma cells.* Mol Cell Biol, 1991. **11**(2): p. 979-86.

163. Lewis, G.D., et al., *Differential responses of human tumor cell lines to anti-p185HER2 monoclonal antibodies.* Cancer Immunol Immunother, 1993. **37**(4): p. 255-63.

164. Park, J.W., et al., *Anti-p185HER2 monoclonal antibodies: biological properties and potential for immunotherapy.* Cancer Treat Res, 1992. **61**: p. 193-211.

165. Carter, P., et al., *Humanization of an anti-p185HER2 antibody for human cancer therapy.* Proc Natl Acad Sci U S A, 1992. **89**(10): p. 4285-9.

166. Wehrman, T.S., et al., *A system for quantifying dynamic protein interactions defines a role for Herceptin in modulating ErbB2 interactions.* Proc Natl Acad Sci U S A, 2006. **103**(50): p. 19063-8.

167. Molina, M.A., et al., *Trastuzumab (herceptin), a humanized anti-Her2 receptor monoclonal antibody, inhibits basal and activated Her2 ectodomain cleavage in breast cancer cells.* Cancer Res, 2001. **61**(12): p. 4744-9.

168. Wong, A.J., et al., *Structural alterations of the epidermal growth factor receptor gene in human gliomas.* Proc Natl Acad Sci U S A, 1992. **89**(7): p. 2965-9.

169. Yamazaki, H., et al., *Amplification of the structurally and functionally altered epidermal growth factor receptor gene (c-erbB) in human brain tumors.* Mol Cell Biol, 1988. **8**(4): p. 1816-20.

170. Heimberger, A.B., et al., *The natural history of EGFR and EGFRvIII in glioblastoma patients.* J Transl Med, 2005. **3**: p. 38.

171. Heimberger, A.B., et al., *Prognostic effect of epidermal growth factor receptor and EGFRvIII in glioblastoma multiforme patients.* Clin Cancer Res, 2005. **11**(4): p. 1462-6.

172. Sok, J.C., et al., *Mutant epidermal growth factor receptor (EGFRvIII) contributes to head and neck cancer growth and resistance to EGFR targeting.* Clin Cancer Res, 2006. **12**(17): p. 5064-73.

173. Johns, T.G., et al., *Novel monoclonal antibody specific for the de2-7 epidermal growth factor receptor (EGFR) that also recognizes the EGFR expressed in cells containing amplification of the EGFR gene.* Int J Cancer, 2002. **98**(3): p. 398-408.

174. Johns, T.G., et al., *Identification of the epitope for the epidermal growth factor receptor-specific monoclonal antibody 806 reveals that it preferentially recognizes an untethered form of the receptor.* J Biol Chem, 2004. **279**(29): p. 30375-84.

175. Rudnick, S.I. and G.P. Adams, *Affinity and avidity in antibody-based tumor targeting.* Cancer Biother Radiopharm, 2009. **24**(2): p. 155-61.

176. Adams, G.P., et al., *High affinity restricts the localization and tumor penetration of single-chain fv antibody molecules.* Cancer Res, 2001. **61**(12): p. 4750-5.

177. Davis, C.G., M.L. Gallo, and J.R. Corvalan, *Transgenic mice as a source of fully human antibodies for the treatment of cancer.* Cancer Metastasis Rev, 1999. **18**(4): p. 421-5.

178. Yang, X.D., et al., *Development of ABX-EGF, a fully human anti-EGF receptor monoclonal antibody, for cancer therapy.* Crit Rev Oncol Hematol, 2001. **38**(1): p. 17-23.

179. Seront, E., et al., *Successful long-term management of a patient with late-stage metastatic colorectal cancer treated with panitumumab.* Cancer Treat Rev, 2010. **36 Suppl 1**: p. S11-4.

180. Bleeker, W.K., et al., *Dual mode of action of a human anti-epidermal growth factor receptor monoclonal antibody for cancer therapy.* J Immunol, 2004. **173**(7): p. 4699-707.

181. Mateo, C., et al., *Humanization of a mouse monoclonal antibody that blocks the epidermal growth factor receptor: recovery of antagonistic activity.* Immunotechnology, 1997. **3**(1): p. 71-81.

182. Crombet-Ramos, T., et al., *Antiproliferative, antiangiogenic and proapoptotic activity of h-R3: A humanized anti-EGFR antibody.* Int J Cancer, 2002. **101**(6): p. 567-75.

183. Allan, D.G., *Nimotuzumab: evidence of clinical benefit without rash.* Oncologist, 2005. **10**(9): p. 760-1.

184. Ramakrishnan, M.S., et al., *Nimotuzumab, a promising therapeutic monoclonal for treatment of tumors of epithelial origin.* MAbs, 2009. **1**(1): p. 41-8.

185. Aboud-Pirak, E., et al., *Efficacy of antibodies to epidermal growth factor receptor against KB carcinoma in vitro and in nude mice.* J Natl Cancer Inst, 1988. **80**(20): p. 1605-11.

186. Hamers-Casterman, C., et al., *Naturally occurring antibodies devoid of light chains.* Nature, 1993. **363**(6428): p. 446-8.

187. Cortez-Retamozo, V., et al., *Efficient cancer therapy with a nanobody-based conjugate.* Cancer Res, 2004. **64**(8): p. 2853-7.

188. Dooley, H. and M.F. Flajnik, *Antibody repertoire development in cartilaginous fish.* Dev Comp Immunol, 2006. **30**(1-2): p. 43-56.

189. Nguyen, V.K., A. Desmyter, and S. Muyldermans, *Functional heavy-chain antibodies in Camelidae.* Adv Immunol, 2001. **79**: p. 261-96.

190. Nguyen, V.K., et al., *Loss of splice consensus signal is responsible for the removal of the entire C(H)1 domain of the functional camel IGG2A heavy-chain antibodies.* Mol Immunol, 1999. **36**(8): p. 515-24.

191. Holliger, P., T. Prospero, and G. Winter, *"Diabodies": small bivalent and bispecific antibody fragments.* Proc Natl Acad Sci U S A, 1993. **90**(14): p. 6444-8.

192. Labrijn, A.F., R.C. Aalberse, and J. Schuurman, *When binding is enough: nonactivating antibody formats.* Curr Opin Immunol, 2008. **20**(4): p. 479-85.

193. Bird, R.E., et al., *Single-chain antigen-binding proteins.* Science, 1988. **242**(4877): p. 423-426.

194. Ewert, S., et al., *Biophysical properties of camelid V(HH) domains compared to those of human V(H)3 domains.* Biochemistry, 2002. **41**(11): p. 3628-36.

195. Vincke, C., et al., *General strategy to humanize a camelid single-domain antibody and identification of a universal humanized nanobody scaffold.* J Biol Chem, 2009. **284**(5): p. 3273-84.

196. Omidfar, K., et al., *Production and characterization of a new antibody specific for the mutant EGF receptor, EGFRvIII, in Camelus bactrianus.* Tumour Biol, 2004. **25**(4): p. 179-87.

197. Omidfar, K., et al., *Production of a novel camel single-domain antibody specific for the type III mutant EGFR.* Tumour Biol, 2004. **25**(5-6): p. 296-305.

198. Omidfar, K., et al., *Expression of EGFRvIII in thyroid carcinoma: immunohistochemical study by camel antibodies.* Immunol Invest, 2009. **38**(2): p. 165-80.

199. Roovers, R.C., et al., *Efficient inhibition of EGFR signaling and of tumour growth by antagonistic anti-EFGR Nanobodies.* Cancer Immunol Immunother, 2007. **56**(3): p. 303-317.

200. Gottlin, E.B., et al., *Isolation of novel EGFR-specific VHH domains.* J Biomol Screen, 2009. **14**(1): p. 77-85.

201. Gainkam, L.O., et al., *Comparison of the biodistribution and tumor targeting of two 99mTc-labeled anti-EGFR nanobodies in mice, using pinhole SPECT/micro-CT.* J Nucl Med, 2008. **49**(5): p. 788-95.

202. Elleman, T.C., et al., *Identification of a determinant of epidermal growth factor receptor ligand-binding specificity using a truncated, high-affinity form of the ectodomain.* Biochemistry, 2001. **40**(30): p. 8930-9.

203. Roovers, R.C., G.A. van Dongen, and P.M. van Bergen en Henegouwen, *Nanobodies in therapeutic applications.* Curr Opin Mol Ther, 2007. **9**(4): p. 327-35.

204. Muyldermans, S., C. Cambillau, and L. Wyns, *Recognition of antigens by single-domain antibody fragments: the superfluous luxury of paired domains.* Trends Biochem Sci, 2001. **26**(4): p. 230-5.

205. Roovers, R.C., et al., *A dual specific anti-EGFR nanobody efficiently inhibits solid tumor growth.* in preparation, 2010.

206. Dong, J., et al., *A single-domain llama antibody potently inhibits the enzymatic activity of botulinum neurotoxin by binding to the non-catalytic alpha-exosite binding region.* J Mol Biol, 2010. **397**(4): p. 1106-18.

207. Decanniere, K., et al., *A single-domain antibody fragment in complex with RNase A: non-canonical loop structures and nanomolar affinity using two CDR loops.* Structure, 1999. **7**(4): p. 361-70.

208. Desmyter, A., et al., *Crystal structure of a camel single-domain VH antibody fragment in complex with lysozyme.* Nat Struct Biol, 1996. **3**(9): p. 803-11.

209. Spinelli, S., et al., *Lactococcal bacteriophage p2 receptor-binding protein structure suggests a common ancestor gene with bacterial and mammalian viruses.* Nat Struct Mol Biol, 2006. **13**(1): p. 85-9.

210. Muyldermans, S., et al., *Sequence and structure of VH domain from naturally occurring camel heavy chain immunoglobulins lacking light chains.* Protein Eng, 1994. **7**(9): p. 1129-35.

211. Stanfield, R.L., et al., *Antibody elbow angles are influenced by their light chain class.* J Mol Biol, 2006. **357**(5): p. 1566-74.

212. Dechant, M., et al., *Complement-dependent tumor cell lysis triggered by combinations of epidermal growth factor receptor antibodies.* Cancer Res, 2008. **68**(13): p. 4998-5003.

213. Dumoulin, M., et al., *A camelid antibody fragment inhibits the formation of amyloid fibrils by human lysozyme.* Nature, 2003. **424**(6950): p. 783-8.

214. Decanniere, K., et al., *Degenerate interfaces in antigen-antibody complexes.* J Mol Biol, 2001. **313**(3): p. 473-8.

215. De Genst, E., et al., *Strong in vivo maturation compensates for structurally restricted H3 loops in antibody repertoires.* J Biol Chem, 2005. **280**(14): p. 14114-21.

216. Saerens, D., et al., *Identification of a universal VHH framework to graft non-canonical antigen-binding loops of camel single-domain antibodies.* J Mol Biol, 2005. **352**(3): p. 597-607.

217. Koide, A., et al., *Exploring the capacity of minimalist protein interfaces: interface energetics and affinity maturation to picomolar KD of a single-domain antibody with a flat paratope.* J Mol Biol, 2007. **373**(4): p. 941-53.

218. Cortez-Retamozo, V., et al., *Efficient tumor targeting by single-domain antibody fragments of camels.* Int J Cancer, 2002. **98**(3): p. 456-62.

219. Davies, D.R. and H.C. Gerson, *Interactions of Protein Antigens with Antibodies.* Proceedings of the National Academy of Sciences of the United States of America, 1996. **93**(1): p. 7-12.

220. Akilesh, S., et al., *Podocytes use FcRn to clear IgG from the glomerular basement membrane.* Proc Natl Acad Sci U S A, 2008. **105**(3): p. 967-72.

221. Caliceti, P. and F.M. Veronese, *Pharmacokinetic and biodistribution properties of poly(ethylene glycol)-protein conjugates.* Advanced Drug Delivery Reviews, 2003. **55**(10): p. 1261-1277.

222. Luo, F., et al., *Correlation of pharmacokinetics with the antitumor activity of Cetuximab in nude mice bearing the GEO human colon carcinoma xenograft.* Cancer Chemotherapy and Pharmacology, 2005. **56**(5): p. 455-464.

223. Tan, A.R., et al., *Pharmacokinetics of Cetuximab After Administration of Escalating Single Dosing and Weekly Fixed Dosing in Patients with Solid Tumors.* Clinical Cancer Research, 2006. **12**(21): p. 6517-6522.

224. Vanhoefer, U., et al., *Phase I Study of the Humanized Antiepidermal Growth Factor Receptor Monoclonal Antibody EMD72000 in Patients With Advanced Solid Tumors That Express the Epidermal Growth Factor Receptor.* J Clin Oncol, 2004. **22**(1): p. 175-184.

225. Leyland-Jones, B., et al., *Pharmacokinetics, Safety, and Efficacy of Trastuzumab Administered Every Three Weeks in Combination With Paclitaxel.* J Clin Oncol, 2003. **21**(21): p. 3965-3971.

226. Chapman, A.P., *PEGylated antibodies and antibody fragments for improved therapy: a review.* Advanced Drug Delivery Reviews, 2002. **54**(4): p. 531-545.

227. Zuker, M., *Mfold web server for nucleic acid folding and hybridization prediction.* Nucleic Acids Res, 2003. **31**(13): p. 3406-15.

228. Kapust, R.B., et al., *Tobacco etch virus protease: mechanism of autolysis and rational design of stable mutants with wild-type catalytic proficiency.* Protein Eng., 2001. **14**(12): p. 993-1000.

229. Guerrero, S.A., et al., *Production of selenomethionine-labelled proteins using simplified culture conditions and generally applicable host/vector systems.* Appl. Microbiol. Biotechnol., 2001. **56**(5-6): p. 718-23.

230. Ferguson, K.M., et al., *Extracellular domains drive homo- but not hetero-dimerization of erbB receptors.* Embo J, 2000. **19**(17): p. 4632-43.

231. Beckett, D., E. Kovaleva, and P.J. Schatz, *A minimal peptide substrate in biotin holoenzyme synthetase-catalyzed biotinylation.* Protein Sci, 1999. **8**(4): p. 921-9.

232. Studier, F.W., *Protein production by auto-induction in high density shaking cultures.* Protein Expr Purif, 2005. **41**(1): p. 207-34.

233. Abbott, J. and D. Beckett, *Cooperative binding of the Escherichia coli repressor of biotin biosynthesis to the biotin operator sequence.* Biochemistry, 1993. **32**(37): p. 9649-56.

234. Evans, G. and R.F. Pettifer, *CHOOCH: a program for deriving anomalous-scattering factors from X-ray fluorescence spectra.* Journal of Applied Crystallography, 2001. **34**(1): p. 82-86.

235. Otwinowski, Z. and W. Minor, *Processing of X-ray Diffraction Data Collected in Oscillation Mode*, in *Methods in Enzymology*, C.W. Carter Jr. and R.M. Sweet, Editors. 1997, Academic Press: New York. p. 307-326.

236. Kraut, J., et al., *Chymotrypsinogen: a three-dimensional fourier synthesis at 5 angstrom resolution.* Proc Natl Acad Sci U S A, 1962. **48**: p. 1417-24.

237. Schneider, T.R. and G.M. Sheldrick, *Substructure solution with SHELXD.* Acta Crystallogr., 2002. **D58**(Pt 10 Pt 2): p. 1772-9.

238. Sheldrick, G.M., *Macromolecular phasing with SHELXE.* Z. Kristallogr., 2002. **217**: p. 644-650.

239. Sheldrick, G.M., et al., *Ab initio phasing*, in *International Tables for Crystallography*, M.G. Rossmann and E. Arnold, Editors. 2001, IUCr and Kluwer Academic Publishers: Dordrecht. p. 333-351.

240. Emsley, P. and K. Cowtan, *Coot: model-building tools for molecular graphics.* Acta Crystallogr., 2004. **D60**(Pt 12 Pt 1): p. 2126-2132.

241. McCoy, A.J., et al., *Phaser crystallographic software.* J Appl Crystallogr, 2007. **40**(Pt 4): p. 658-674.

242. Schuck, P., *Size-distribution analysis of macromolecules by sedimentation velocity ultracentrifugation and lamm equation modeling.* Biophys. J., 2000. **78**(3): p. 1606-19.

243. Garcia De La Torre, J., M.L. Huertas, and B. Carrasco, *Calculation of hydrodynamic properties of globular proteins from their atomic-level structure.* Biophys. J., 2000. **78**(2): p. 719-30.