universität wien

DISSERTATION

Titel der Dissertation

Biochemical and structural characterization of preformed EGF receptor dimers

Verfasser

Mag. rer. nat. Jesús Fernández Rodríguez

angestrebter akademischer Grad

Doktor der Naturwissenschaften (Dr.rer.nat.)

Wien, 2011

Studienkennzahl lt. Studienblatt:	A 091 490
Studienrichtung lt. Studienblatt:	Dr.-Studium der Naturwissenschaften Molekulare Biologie
Betreuer:	Dr. Thomas Marlovits

September, 2011

Para mi madre, porque si he llegado

hasta aquí ha sido gracias a su ayuda.

Contents

Contents i

List of Figures v

List of Tables ix

1 Zusammenfassung 1

2 Abstract 5

3 Introduction 9

3.1 Signalling by receptor tyrosine kinases 9

3.2 The ErbB family of receptors 10

3.3 Structure and mode of action of the EGF receptor 19

3.3.1 Ectodomain organization 19

3.3.2 Kinase domain structure and activation 21

3.4 Ligand binding features of the EGF receptor 27

3.4.1 Curvilinear Scatchard plots of EGF receptors 27

3.4.2 Preformed, unliganded dimers and oligomers of EGF receptor . 30

3.5 Coiled-coil design 31

4 Results **35**

4.1 Characterization of the E3/K3 interaction in fusion proteins 35

4.1.1 Influence of peptide linkers and concentration in the solution behaviour of the K3 and E3 chimeric proteins 36

4.1.2 E3/K3-driven heterodimerization of the fusion proteins and distance estimation 38

4.1.3 K3 and E3 peptides are biologically inert when expressed *in vivo* 40

4.1.4 The E3/K3 interaction is stable under different buffer conditions 43

4.2 Purification of unliganded EGFR dimers for structural analysis 46

4.2.1 Full length EGFR unliganded dimers are degraded during purification 47

4.2.2 EGFRΔ998 can be purified in a dimeric form in the absence of ligand 48

4.2.3 Purification of other ErbB family dimers: EGFR/ErbB2 54

4.3 Structural analysis of the unliganded EGFRΔ998 dimer 56

4.3.1 Negative staining and image acquisition of purified samples . . . 56

4.3.2 3D model of the EGFRΔ998 E3/K3 dimer reveals ectodomain asymmetry prior to ligand binding 58

4.3.3 Binding of EGF induces large structural re-arrangements in EGFRΔ998 E3/K3 dimers 70

4.4 Biochemical characterization of dimeric EGFRΔ998 73

4.4.1 Unliganded EGFRΔ998 dimers are able to phosphorylate and trigger downstream signalling upon ligand binding 73

4.4.2 Tyrosine kinase activity of purified unliganded EGFRΔ998 dimers is independent of ligand addition 78

4.4.3 EGFRΔ998 dimers bind ligand with high and low affinities . . . 81

5 Discussion **85**

5.1 The E3/K3 coiled coil as a universal dimerization and purification tool and its applications for signal transduction studies 86

5.2 Phosphorylation and activity of EGFR preformed dimers *in vivo* and *in vitro* 88

5.3 Ectodomain asymmetry of preformed EGFR dimers and curvilinear Scatchard plots 89

5.4 Biological implications of ectodomain asymmetry in EGFR preformed dimers 95

6 Materials and Methods **99**

6.1 Materials 99

6.1.1 Constructs used in the characterization of the E3/K3 interaction 99

6.1.2 Constructs used for the purification and characterization of the ErbB receptor pairs 100

6.2 Methods 101

6.2.1 Cloning of the Venus/ECFP E3/K3 constructs 101

6.2.2 Expression, purification and TEV cleavage of E3/K3–tagged ECFP /Venus 102

6.2.3 Gel filtration standard curve 103

6.2.4 Fluorescence measurements, FRET efficiency and distance estimation 104

6.2.5 Isothermal Titration Calorimetry 105

6.2.6 Mammalian cell culture 105

6.2.7 Cloning of EGFR and ErbB2 E3/K3 constructs 106

6.2.8 Preliminary purification tests of EGFR constructs 107

6.2.9 Generation of CHO-K1 TetOn stable cell lines 108
6.2.10 Large scale expression and purification of EGFR constructs from CHO-K1 TetOn cells 108
6.2.11 Fluorescence microscopy of *in vivo* ligand binding 109
6.2.12 Kinase assays of purified EGFR dimers 110
6.2.13 Activation and signalling of EGFR constructs *in vivo* 110
6.2.14 Ligand saturation assays of purified EGFR dimers 111
6.2.15 Negative staining of purified EGFR dimers 112
6.2.16 Image acquisition and processing 113
6.2.17 3D model generation and refinement 113

References **115**

7 Acknowledgements **139**

8 Curriculum vitae **141**

List of Figures

1 General architecture of ErbB receptors 11
2 Conformations of tethered and extended ErbB receptors 15
3 Cellular pathways activated by ErbB receptors 18
4 Depiction of unliganded and liganded EGFR molecules 20
5 Conformations of the inactive and active kinase of the EGF receptor . . 23
6 Comparison between Src/CDK activation and asymmetric EGFR kinase dimers 24
7 Activation of the EGF receptor kinase domains 25
8 Activation pathway of EGF receptors 28
9 Ligand binding model for *Drosophila* EGF receptors 29
10 Helical view representation of a dimeric coiled-coil interaction 32
11 Depiction of the E3/K3 interaction 33
12 Effect of concentration in the association state of the different purified constructs 37
13 Gel filtration profile the of the tandem affinity purified heterodimer . . 39
14 Heterodimerization of SVenus-TEV-K3-pG-His and ECFP-TEV-E3-Strp 41
15 Analysis of samples after affinity purification and gel filtration 42
16 Stability of the fusion E3/K3 heterodimer under different buffer conditions 44

17 Gel filtration profile of the tandem affinity purified heterodimer at pH 5 45
18 Comparison of EGFRΔ998 and EGFR full-length homodimer purifications 48
19 Tandem affinity purification of EGFRΔ998 dimers 49
20 Gel filtration analysis of purified EGFRΔ998 dimers 51
21 Analysis of the second peak after gel filtration 52
22 Different profiles obtained in the purification of EGFRΔ998 dimers . . 53
23 Influence of polyglycine linkers and geldanamycin in the purification of truncated EGFR/ErbB2 heterodimers 55
24 Negative staining of purified EGFRΔ998 dimers 57
25 Comparison of negatively stained dimeric EGFRΔ998 and integrin . . . 58
26 First 50 ML2D classes sorted by number of members per class 59
27 Selected classes for initial 3D reconstruction 60
28 Initial Euler angle assignment and 3D reconstruction 61
29 First 3D model of a EGFRΔ998 E3/K3 preformed dimer 62
30 Final 3D reconstruction filtered to 25 $\mathring{A}$ 64
31 Euler angle coverage and resolution of the final 3D reconstruction . . . 65
32 Contoured ectodomains of the final 3D reconstruction 66
33 Crystal docking of EGFR ectodomains (1NQL) into the ectodomain electron density maps 68
34 ML3D results of the final 3D model 69
35 Contoured ectodomains of the two ML3D maps 70
36 First 50 ML2D classes of liganded EGFRΔ998 E3/K3 71
37 Top 70 classes of liganded EGFR E3/K3 dimers 72
38 EGFRΔ998 E3/K3 are present at the plasma membrane and internalized upon ligand binding 74

39 EGFRΔ998 E3/K3 dimers are able to phosphorylate and trigger downstream signalling upon ligand binding 75
40 Phosphorylation analysis of EGFRΔ998 E3/K3 and full-length EGFR reveals higher basal signals in CHO TetOn cells 77
41 *In vitro* kinase activity of purified EGFRΔ998 dimers 80
42 Ligand saturation analysis of purified EGFRΔ998 E3/K3 dimers 82
43 Comparison between two-site binding and one-site binding with Hill coefficient 91
44 Proposed activation pathway for the EGF receptor 96
45 Standard curve for TRF 112

List of Tables

3.1 Phosphotyrosine profile of ErbB receptors 13
3.2 Ligand specificity of ErbB receptors 14

6.1 Ligand saturation curve data 112

1

Zusammenfassung

Rezeptor-Tyrosinkinasen (RTK) sind eine Gruppe von Proteinen, die in die Plasmamembran eingebettet und für die Zell-Signalisierung verantwortlich sind. Im Menschen umfassen RTK 59 Gene,die in 20 Familien eingeteilt werden. Dazu gehört die Familie der Epidermalen-Wachstumsfaktor-Rezeptoren (ErbB oder HER), die wiederum aus vier Mitgliedern besteht: EGFR/ErbB1/HER1, ErbB2/HER2/Neu, ErbB3/HER3 und ErbB4/HER4. Diese binden unterschiedliche Liganden und bilden eine Reihe von Homo- und Heterodimeren um verschiedene zelluläre Antworten hervorzurufen. ErbB-Rezeptoren sind verantwortlich dafür, Signalisierung durch Liganden-induzierte Dimerisierung auszulösen und sind in Signalnetzwerke involviert, die mit Zellwachstum und Zelldifferenzierung in Verbindung gebracht werden. Viele Mutationen der ErbB-Rezeptoren führen zu Zelltransformation und Krebs .

Obwohl vieles über die Struktur und Aktivierung der ErbB-Rezeptoren bekannt ist, fehlt ein komplettes Bild der molekularen Mechanismen, die zu einer Umwandlung extrazellulärer in intrazelluläre Signale führen (der Aktivierung der Kinase-Domänen). Ein Teil der ErbB-Rezeptoren in der Plasmamembran können bei Abwesenheit von Liganden die Form von vorgeformten, nicht-kovalent verbundenen Dimeren annehmen.

Im Fall von EGFR bindet der Rezeptor *in vivo* Liganden mit zwei Affinitäten. Bei einer Aufreinigung des Rezeptors bleibt allerdings nur die Komponente mit niedriger Affinität übrig. Es wird vermutet, dass die hohe Affinitätsbindung und die vorgeformten Dimere in einem Zusammenhang stehen und eine Art negative Kooperativität aufweisen.

Aus diesem Grund entwickelte ich eine Methode, die es erlaubt, die strukturelle sowie biochemische Charakterisierung der vorgeformten EGFR-Homodimere durchzuführen. Die Schwierigkeit lag darin, ein Werkzeug einzuführen, das die Reinigung der vorgeformten Dimere ermöglicht, ohne auf die Zusetzung kovalenter Modifikationen zurückgreifen zu müssen, welche die inhärente Flexibilität und Struktur der Rezeptoren verhindern. In dieser Arbeit wird eine Methode vorgestellt, die es erlaubt, die EGF-Rezeptoren bei Abwesenheit der Liganden in einem dimerisierten Zustand während des Reinigungsprozesses zu erhalten. Dieses Werkzeug besteht aus einem *de novo* konstruierten heterodimerischen "coiled-coil" (Doppelwenden), das E3/K3 genannt wird. Dieses Paar wurde bereits beschrieben, jedoch noch nicht zu einem Protein fusioniert. Aus diesem Grund testete ich das Verhalten des E3/K3 Paares mit Hilfe der fluoreszierenden Proteinen Venus und ECFP. Durch die Verwendung von Affinitätsreinigungs- und Gelfiltrationstechniken sowie von FRET zeige ich, dass das E3/K3 Paar sehr gut für die Heterodimerisierung zweier Proteine geeignet ist. Es ist außerordentlich spezifisch und die Interaktion hält dem Reinigungsprozess stand.

Anschließend wurde das E3/K3 Dimerisierungsmodul in EGF-Rezeptoren transplantiert und in CHO-Zellen exprimiert. Die Ergebnisse zeigen, dass EGFR auch bei Abwesenheit des Liganden in dimerischer Form gereinigt werden und weiterhin *in vitro* Liganden binden kann, ohne Tyrosinkinase-Aktivität zu verlieren. Außerdem ist die Probe extrem rein und es kann für strukturelle Studien mit Single-Particle-Negative-Staining benutzt werden. Ich stelle eine dreidimensionale Rekonstruktion eines vorgeformten EGFR Homodimers vor wobei (soweit bekannt) zum ersten Mal eine in der

vollen Länge RTK gezeigt werden kann. Überraschenderweise zeigt der vorgeformte EGFR Dimer Ektodomänenasymmetrie. Eine der Ektodomänen scheint eine "tethered" (angebundene) Konfiguration aufzuweisen, während sich die andere in einem freieren Zustand befindet. Da die Anwesenheit struktureller Heterogenität in den Ektodomänen auf die Möglichkeit zweier verschiedener Bindungsstellen deutet, führte ich Ligandenbindung-Assays mit den gereinigten EGFR Dimeren durch. Das Scatchard-Diagramm zeigt eine kurviliniare, konkave Form, die charakteristisch für *in vivo* EGFR-Ligandenbindung ist. Es ist daher wahrscheinlich, dass die Ektodomänenasymmetrie direkt in Heterogenität der Bindungsstellen übersetzt wird, was wiederum die zwei Affinitäten generiert, die *in vivo* zu sehen sind. Das ist das erste Mal, dass ein solches Scatchard-Diagramm mit gereinigten EGF-Rezeptoren vorgestellt wird. In dieser Arbeit zeige ich eine unerwartete Architektur des vorgeformten EGFR Dimers. Anstatt symmetrische Dimer zu formen, gehen die Ektodomänen vor der Ligandenbindung in asymmetrischer Weise zusammen. Wir hoffen, dass sich diese Entdeckung als nützlich für die zukünftige Entwicklung von Medikamenten bei der Krebsbehandlung zeigt sowie Aufschluss über die allgemeinen Mechanismen der Signaltransduktion –nicht nur von ErbB, sonder auch von anderen Rezeptoren – geben kann.

2

Abstract

Receptor tyrosine kinases (RTKs) are a family of plasma-membrane embedded proteins that mediate cellular signalling. In humans, RTKs comprise 59 genes divided into 20 families. One of these is the family of the epidermal growth factor receptors (ErbBs or HERs), composed of four members: EGFR/ErbB1/HER1, ErbB2/HER2/Neu, ErbB3/HER3 and ErbB4/HER4. They can bind different ligands and form a wide array of homo- and heterodimers, thus generating different cellular responses. ErbB receptors are thought to trigger signalling by ligand-mediated dimerization, and are involved in pathways related to cellular growth and differentiation. Many mutations of ErbB receptors are known, and these typically lead to cell transformation and cancer.

Even though much is known about the structure and activation of ErbB receptors, we are still missing a complete picture of the molecular mechanisms that lead to the conversion of an extracellular signal into an intracellular event (the activation of the kinase domains). For example, ErbB receptors are known to exist as preformed, non-covalently linked dimers at the plasma membrane in the absence of ligand, but this association dissapears when the receptors are purified. EGFR is also known to bind ligand with two apparent affinities, but only *in vivo*. Once the receptor is purified,

only the low affinity component remains. The high affinity binding and the preformed dimers are thought to be related. These could present negative cooperativity that would explain the presence of two affinity sites. For this reason, I wanted to develop a method that allows both the structural and biochemical characterization of preformed EGFR homodimers. The challenge was to introduce a tool that allows the purification of a preformed dimer without the necessity of adding any covalent modifications, that would hinder the inherent flexibility and structure of the receptor.

In this work, I have introduced a heterodimerization tool that maintains the EGF receptors in a dimerized state in the absence of ligand throughout the purification procedure. This tool consists of a *de novo* designed heterodimeric coiled-coil, termed E3/K3. This pair has been previously described, but never fused to a protein. For this reason, I have tested the behaviour of the E3/K3 pair when fused to two model proteins, Venus and ECFP. By using affinity tag purification, gel filtration, isothermal titration calorimetry (ITC) as well as FRET, I have showed that the E3/K3 pair is very well suited for the heterodimerization of two proteins. It is extremely specific, and the interaction is able to withstand the purification procedure.

The E3/K3 dimerization module was then transplanted to EGF receptors, and expressed in CHO cells. The results show that EGFR can be purified in the absence of ligand in a dimeric form, and still bind ligand and possess kinase activity. Besides, the sample is extremely clean and can be readily used for structural studies using single-particle negative staining. I present a three-dimensional reconstruction of a preformed EGFR homodimer, which, to our knowledge, is the first one of a near-full length RTK. Strikingly, the EGFR preformed dimer presents ectodomain asymmetry. One of the ectodomains seems to exist in a tethered configuration while the other appears to be in a more extended state.

Since the presence of a structural heterogeneity in the ectodomains points to the

possibility of two different binding sites, I performed ligand binding assays with the purified EGFR dimers. The Scatchard plot shows a curvilinear, concave-up shape, characteristic of ligand binding to EGFR *in vivo*. Thus, it is very probable that the ectodomain asymmetry directly translates into a heterogeneity of binding sites, hence generating the two affinities seen *in vivo*. This is the first time that such a Scatchard plot is seen with purified EGF receptors.

With this work, I have shown that the architecture of preformed EGFR dimers differs from what was expected. Instead of forming symmetric dimers, the ectodomains associate in an asymmetric fashion prior to ligand binding. We hope that this discovery will prove to be useful for the future development of anticancer drugs, as well as shed light on the general mechanisms of signal transduction, not only of ErbBs, but also of other receptors.

3

Introduction

3.1 Signalling by receptor tyrosine kinases

Cells have different ways of transferring signals from the extracellular to the intracellular spaces. One of them involves the use of receptor tyrosine kinases (RTKs). All of these proteins share a general architecture, with ligand-binding ectodomains, a single-pass transmembrane helix, and an intracellular region comprising a kinase domain and a regulatory C-terminal tail [1]. RTKs are evolutionary well conserved, from *C. elegans* to humans [1, 2]. In the latter, the genome has 59 genes that encode 20 families of RTKs [3]. They regulate key processes in the cells, related to survival, migration, growth and differentiation, and cell cycle control [4]. Since they control very important cellular pathways, it seems obvious that an aberrant signalling by RTKs can lead to different pathologies, including cancer and inflammation, among others [1]. The general mechanism of RTK activation involves ligand-mediated dimerization of the extracellular domains and a subsequent activation of the intracellular kinase domains. However, there are different ways of achieving activation. Some RTKs exist as covalently bound, inactive dimers in the plasma membrane, such as the insulin receptors [5]. In some other

cases, like epidermal growth factor receptors (ErbBs), a fraction of them can exist as preformed dimers in the absence of ligand, although their function is still not clear [6]. Ephrin receptors (Eph), on the other hand, need to form higher-order oligomers to get activated [7]. Ligand binding can also differ among RTKs. The most common case is the cross-linking of the ectodomains by a bivalent ligand, like in the Flt1 vascular endothelial growth factor receptor (VEGF) [8]. For some other receptors, like ErbBs, the dimerization interface is entirely mediated by receptor contacts [6]. In any case, and in normal conditions, the ligand binding event is necessary in all RTKs in order to generate a cellular response. This will be started by the docking of several proteins to phosphorylated residues in the intracellular kinase and C-terminal domains of the RTKs. In this work, we have focused on the structural determination of a preformed epidermal growth factor receptor dimer. We hope that this work will shed light on the molecular mechanisms that link an extracellular binding event to intracellular kinase activation.

3.2 The ErbB family of receptors

The epidermal growth factor receptor (EGFR) is one of the most studied protein tyrosine kinases, since its first sequencing and detailed description in 1984 [9]. Since then, other three members of this family have been described in mammals, and collectively this family of receptors are termed ErbB from their sequence similarity to the viral erythroblastoma protein [10] or HER receptors (from Human Epidermal growth factor Receptors). They have homologs in many other organisms such as *Caenorhabditis elegans* [11] and *Drosophila melanogaster* [12]. In the case of mammals, the four members of the ErbB family are ErbB1/EGFR/HER1, ErbB2/HER2, ErbB3/HER3 and ErbB4/HER4, and have 60-75% sequence identity [13].

All members of the ErbB family are single-pass transmembrane receptors and share a common architecture: a glycosylated ectodomain with four subdomains (∼620 aa), a single-pass transmembrane domain, a juxtamembrane domain (∼45 aa), an intracellular tyrosine kinase domain (∼270 aa) and a regulatory C-terminal tail (∼230 aa) [6, 14, 15] (Figure 1).

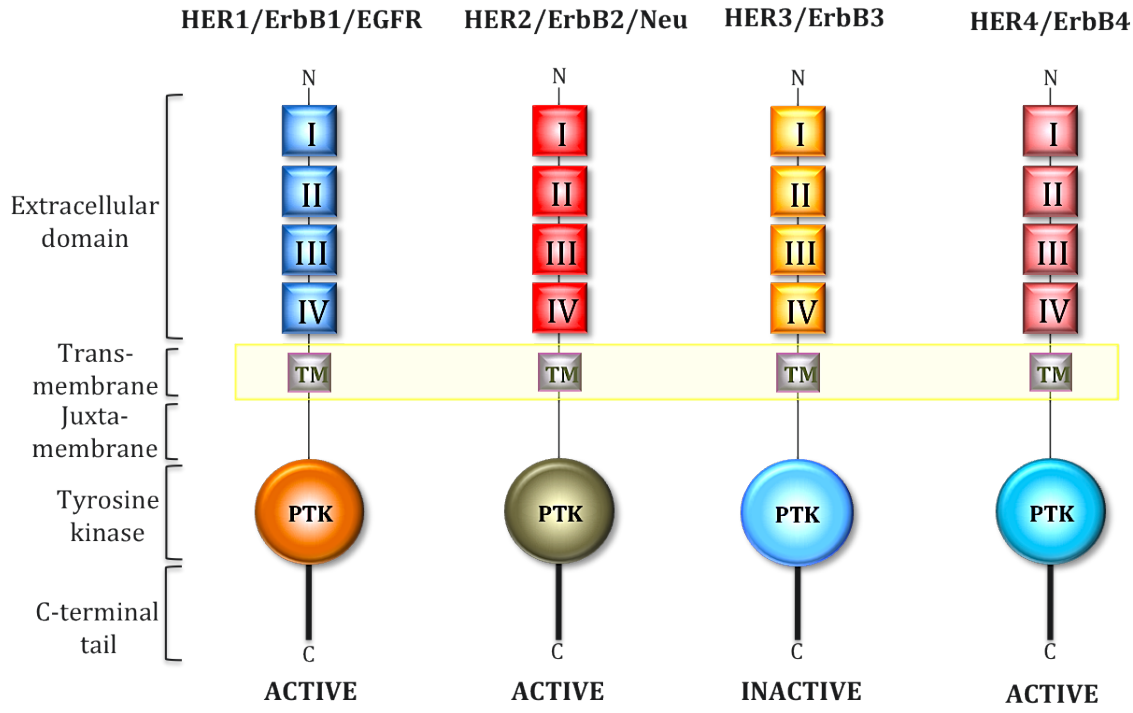


Figure 1: General architecture of ErbB receptors. ErbB receptors have a molecular weight of about 180 kDa and are composed of an extracellular, ligand binding domain, a transmembrane domain, a juxtamembrane domain, a protein tyrosine kinase domain and a regulatory C-terminal tail.

ErbB receptors have a molecular weight of about 180 kDa and are heavily glycosylated [15]. Up to 30 kDa of the total molecular weight is represented by oligosaccharide moieties. These sugars have proved to be crucial for ErbB action, and are especially well characterized for EGFR [16–21]. The EGFR immature protein is 1210 aminoacids

long, and will be cleaved upon plasma membrane insertion to generate the mature, glycosylated form consisting of 1186 residues [9]. This mechanism is common to the other ErbB members.

The ectodomain of the ErbB receptors can be subdivided into four subdomains, classically called L1 (leucine-rich), CR1 (cystein-rich), L2 and CR2, or alternatively, subdomains I to IV. Domains I and III are responsible for ligand binding, and domain II contains a loop that allows dimerization of the ectodomains [1, 22, 23]. Domain IV is thought to be involved in the maintenance of the undimerized configuration [24–26]. In normal conditions, ErbB receptors become activated after ligand-mediated dimerization. The ligand-mediated dimerization and activation of ErbB receptors causes the phosphorylation of several residues in the intracellular domains, including tyrosines, threonines and serines (Table 3.1, adapted from [15, 27]). These phosphorylated sites, in turn, serve as docking points for several intracellular proteins, and hence mediate different signaling pathways [2, 28–31]. Although several residues are known to be transphosphorylated by the kinase domains themselves, some others, especially *in vivo*, can be modified by cellular proteins [32–34] like Src and Jak-2 [35, 36].

EGFR family members can form homo- and heterodimers [37, 38]. They are able to dimerize after two molecules of ligand have bound the receptors, hence generating a 2:2 complex [14, 32, 39–45]. Hence, the ligands do not contribute to the dimerization interface, as in other RTKs [1]. Every ErbB receptor, except ErbB2, is able to bind different ligands (Table 3.2, taken from [15]), and if more than one ErbB receptor is expressed in the cell, a wide array of homo- and heterodimers are formed [6]. All ErbB receptor ligands are soluble and small (about 6-10 kDa) [15]. They are synthesised as large, transmembrane precursors that will be cleaved to yield a mature growth factor by ADAM metalloproteases at the plasma membrane [46].

EGFR		**ErbB2**		**ErbB3**		**ErbB4**	
Phospho-Y	*Domain*	*Phospho-Y*	*Domain*	*Phospho-Y*	*Domain*	*Phospho-Y*	*Domain*
703	Kinase	663	JM	661	JM	708	Kinase
740	Kinase	713	Kinase	770	Kinase	782	Kinase
777	Kinase	750	Kinase	804	Kinase	808	Kinase
789	Kinase	759	Kinase	805	Kinase	850	Kinase
803	Kinase	781	Kinase	849	Kinase	883	Kinase
845	Kinase	813	Kinase	869	Kinase	896	Kinase
867	Kinase	854	Kinase	878	Kinase	925	Kinase
876	Kinase	884	Kinase	893	Kinase	959	C-tail
891	Kinase	901	Kinase	922	Kinase	997	C-tail
920	Kinase	930	Kinase	956	C-tail	1010	C-tail
954	C-tail	983	C-tail	1035	C-tail	1031	C-tail
974	C-tail	1001	C-tail	1113	C-tail	1041	C-tail
992	C-tail	1090	C-tail	1140	C-tail	1056	C-tail
1045	C-tail	1105	C-tail	1178	C-tail	1103	C-tail
1068	C-tail	1117	C-tail	1180	C-tail	1125	C-tail
1086	C-tail	1174	C-tail	1203	C-tail	1137	C-tail
1101	C-tail	1199	C-tail	1205	C-tail	1163	C-tail
1114	C-tail	1200	C-tail	1241	C-tail	1177	C-tail
1148	C-tail	1226	C-tail	1243	C-tail	1183	C-tail
1173	C-tail			1257	C-tail	1196	C-tail
				1270	C-tail	1217	C-tail
				1288	C-tail	1233	C-tail
				1309	C-tail	1237	C-tail
						1241	C-tail
						1243	C-tail
						1259	C-tail
						1276	C-tail

Table 3.1: Phosphotyrosine profile of ErbB receptors. Listed are residues that get phosphorylated upon activation. Residues in red denote tyrosines that are identical in all receptors or whose surrounding sequence is identical to other ErbB receptors. JM, juxtamembrane domain. Adapted from [27].

Ligand	Receptor			
	EGFR	*ErbB2*	*ErbB3*	*ErbB4*
EGF	+	-	-	-
TGF-α	+	-	-	-
HB-EGF	+	-	-	+
Amphiregulin	+	-	-	-
Betacellulin	+	-	-	+
Epigen	+	-	-	-
Epiregulin	+	-	-	+
Neuregulin-1	-	-	+	+
Neuregulin-2	-	-	+	+
Neuregulin-3	-	-	-	+
Neuregulin-4	-	-	-	+

Table 3.2: Ligand specificity of ErbB receptors (taken from [15]).

In the last years, extensive structural analyses have shown how the ErbB ectodomains and kinase domains are organized. Crystal structures have revealed that when unbound, all ErbB receptors, except ErbB2, have a tethered, closed conformation [24–26, 47–49]. Upon ligand binding, the conformation changes to an extended one and allows ligand-mediated dimerization [23, 50, 51] (Figure 2). This event is thought to promote the association of the intracellular kinase domains that results in their activation via an asymmetric dimer [52]. The mechanism will be discussed extensively in section 3.3.2.

It seems clear that EGFR and ErbB4 are able to form homodimers upon ligand binding [53], and that ErbB2 is the preferred partner for heterodimerization [48, 54–56]. On the other hand, ErbB3 and ErbB2 are not able to homodimerize [40, 57]. ErbB3 is able to bind ligand and adopt an extended conformation in solution but without forming homodimers, in contrast to EGFR soluble ectodomains [58]. The only ErbB family member without a known ligand is ErbB2 [32, 40, 48]. ErbB2 adopts a constitutively extended configuration [47, 48] that makes domains I and III to be very close to each other, so that there is no space for a ligand molecule to bind. For

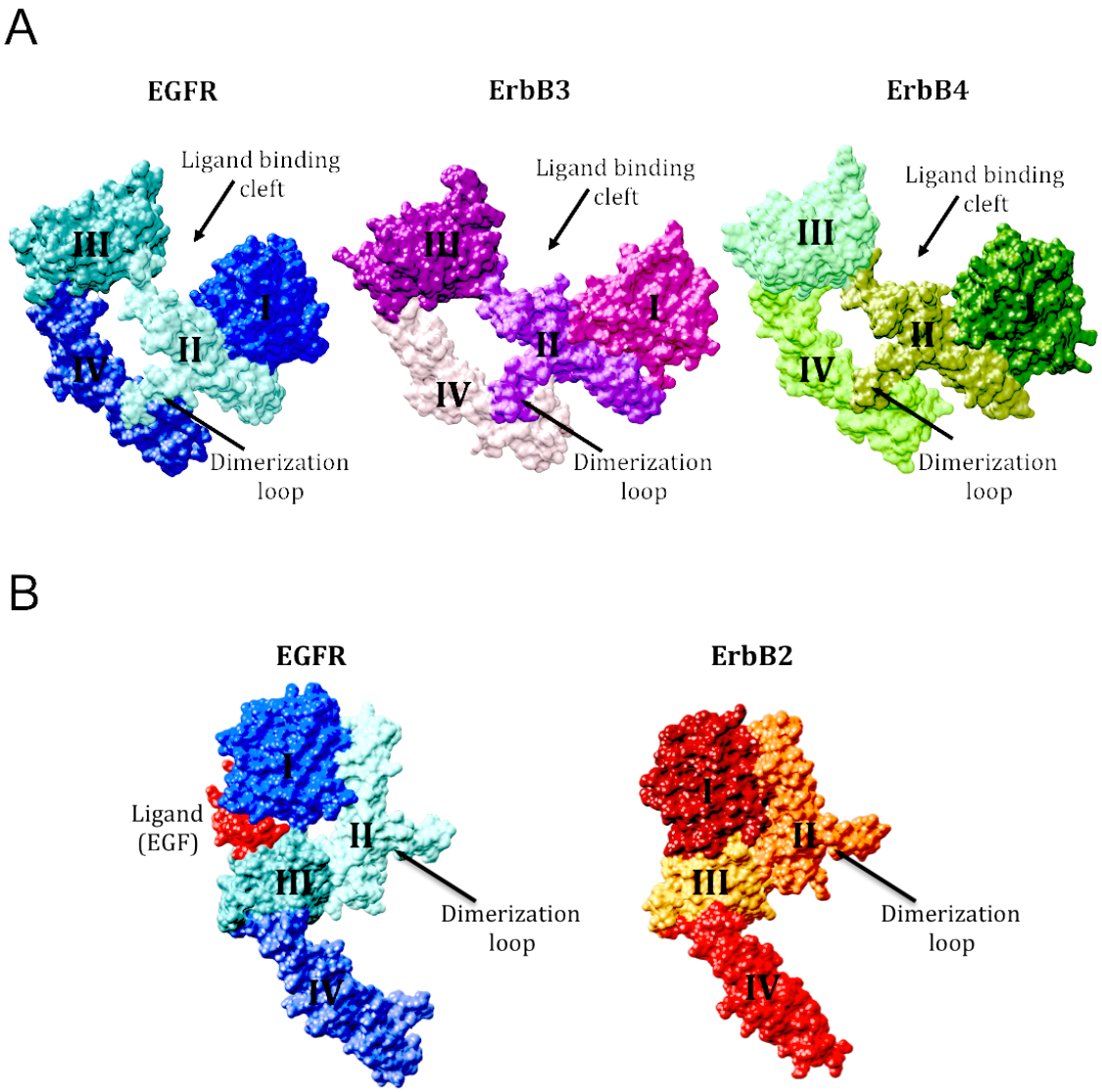


Figure 2: Conformations of tethered and extended ErbB receptors. A) Crystal structures showing the tethered, unbound conformations of EGFR, ErbB3 and ErbB4 (PDB codes 1NQL, 1M6B and 2AHX, respectively). Subdomains I to IV have been depicted in different colours. B) Crystal structures showing the extended, dimerization-prone configurations of EGFR bound to its ligand EGF and ErbB2 (PDB codes, 3NJP and 1N8Z, respectively). Note that the distance between domains I and III of ErbB2 is too small as for a ligand to bind.

this reason, it is thought to be the preferred heterodimerization partner, since it is poised to interact with other ErbB receptors. This is also confirmed by the fact that overexpression of ErbB2, but not EGFR, induces cell transformation, as is the case for a class of breast cancers [40, 59, 60]. Surprisingly, the *Drosophila* unbound EGFR adopts the same extended conformation as the human ErbB2, and so a different regulatory mechanism is involved in its autoinhibition [61].

ErbB3 is a kinase-defective receptor [40, 62, 63], and although it has recently been described to possess a measurable kinase activity, in comparison to other receptors it is very low [64]. For this reason, ErbB3 needs to hetero-associate with another ErbB member in order to generate an active dimer [40, 57]. ErbB2/ErbB3 heterodimers are the most potent signaling pair of all possible ErbB receptor dimers, due to its decreased downregulation [40, 65–67].

Several ligands can bind the same receptor. All ligands are structurally related [15, 32], and the fact that they can trigger different signaling pathways could be explained by how ligand binding affects the angle that subdomain II forms with the rest of the ectodomain after this event. Slightly different angles would promote homo- or heteroassociation, and could have an effect in the way the intracellular kinase domains associate [6]. The dimerization and activation of the kinase domain of ErbB receptors has also been extensively studied. The current model involves an allosteric mechanism in which the kinases form an active asymmetric dimer after ligand binding [13, 52, 68–72]. Since the C-terminal domains are different in each ErbB receptor, homo- and heterodimerization events will cause the phosphorylation of different residues and hence generate different signaling cascades [32]. The potency and length of the signals are tightly regulated, mainly by receptor degradation through internalization [66, 73, 74] and recycling to the plasma membrane [74]. Since ErbB receptors can be phosphorylated by different molecules *in vivo*, the array of cellular cascades, and hence responses,

that can be activated is very large. The main pathways that are activated by ErbB receptors are the RAS/ERK pathway, the PI3K/AKT pathway, the JAK/STAT pathway and the PLC/PKC pathway. These are involved in proliferation responses, such as anti-apoptosis, cell migration, survival and differentiation [2, 28–31]. A scheme depicting the pivotal members of these pathways can be found in Figure 3.

Apart from generating a cellular response, ErbB receptors have been reported to reach the nucleus and probably have some transcriptional activity [75, 76]. Since ErbB receptors control many pathways related to cell growth and apoptosis inhibition, their aberrant signalling can lead to cell transformation and cancer. They are involved in glioma, breast, head and neck, bladder, ovary and lung cancers, as well as in renal carcinomas [28, 77]. There are several mutations that are responsible for the anomalous behaviour ErbB receptors. EGFR variant III (EGFRvIII) lacks most of the extracellular domain, cannot bind ligand and exists in a constitutively active state at the plasma membrane [10]. There are other extracellular mutations of EGFR involved in cancer progression [78]. The kinase domain can also be mutated in different residues, or have insertions and deletions that make it constitutively active [79]. ErbB2 is overexpressed in several types of cancers and can be mutated in its kinase domain [40, 59, 60], and can also be associated with an overexpression of ErbB3 [30]. Mutations in ErbB4 are associated with melanoma [31]. For these reasons, ErbB family members are an active field of research for anticancer drugs. These include antibodies that bind extracellular domains of both the wild-type and the mutated receptors (like cetuximab/Erbitux^{TM} and Zalutumumab in the case of EGFR and trastuzumab/Herceptin^{TM} in the case of ErbB2 [80]), and small molecule tyrosine kinase inhibitors (TKIs). These block the ATP binding site of the kinase domain. Several TKIs are approved for cancer therapy, such as erlotinib/Tarceva^{TM} and gefitinib/Iressa^{TM} [79].

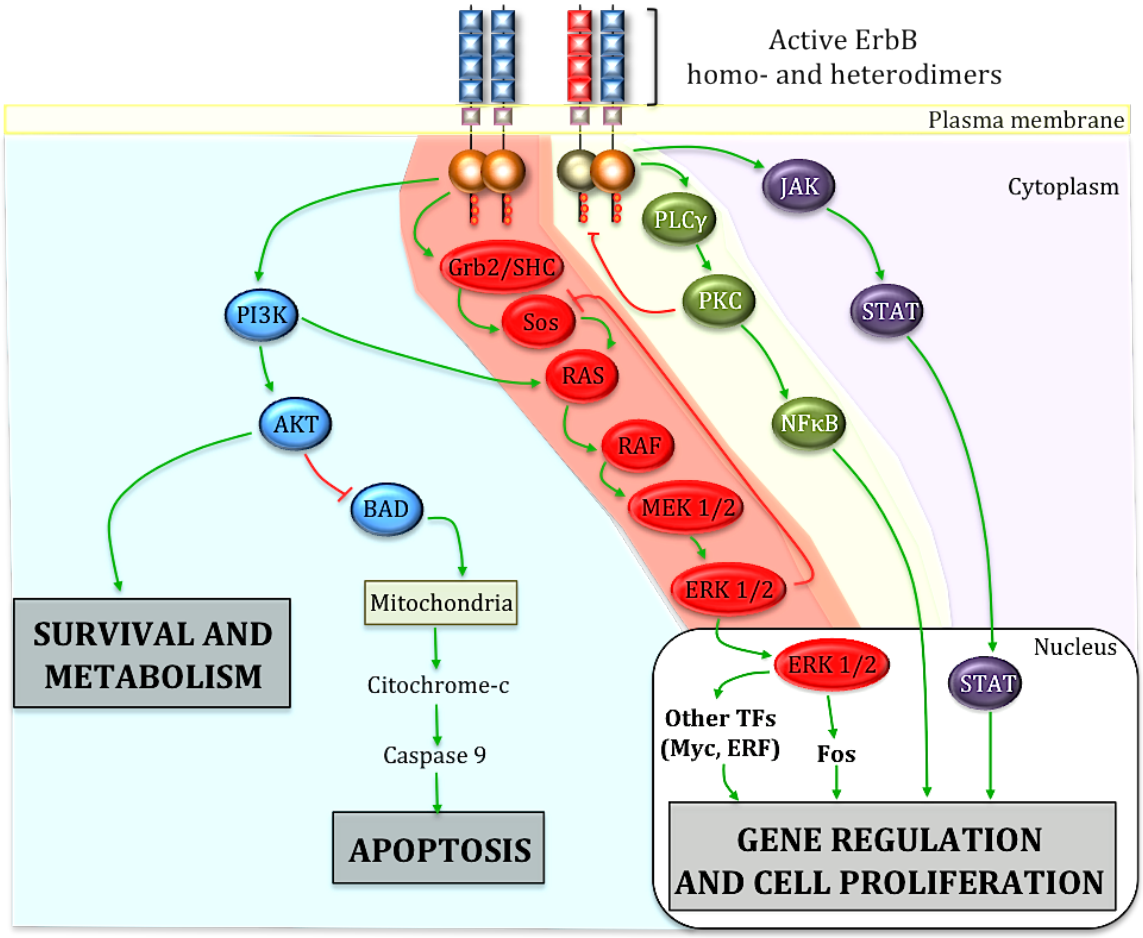


Figure 3: Cellular pathways activated by ErbB receptors. Phosphorylated ErbB receptors can activate four different pathways. The PI3K/AKT pathway (in blue): PI3K binds phosphorylated receptors via its SH2 domain. It induces the generation of phosphatidylinositol triphosphate (PIP_3), that serves as a docking site for AKT. AKT acts on many other substrates, such as BAD, which is downregulated, thus inhibiting the apoptosis response. The RAS/ERK pathway (in red): phosphorylated receptors bind Grb2 and SHC adaptors that in turn recruit SOS, a GDP/GTP exchanger. GTP is then loaded on RAS, and a kinase cascade is generated that results in the phosphorylation of ERK and its translocation to the nucleus. The PLC/PKC pathway (in green): phospholipase Cγ protein kinase (PLCγ) binds phosphorylated receptors, producing diacylglycerol (DAG) and inositol triphosphate (IP_3). These generate Ca^{2+} release, which in turn activates the protein kinase C family (PKC). The JAK/STAT pathway (in purple): Janus kinase (JAK) and STAT (Signal Transducer and Activator of Transcription) are able to bind phosphorylated receptors. JAK then phosphorylates STAT, which is translocated to the nucleus. Adapted from [28, 31].

3.3 Structure and mode of action of the EGF receptor

3.3.1 Ectodomain organization

The ectodomain of EGF receptor has been crystallized in several forms, and these experiments have given valuable insights into the ectodomain organization and its consequences for EGFR activation. The crystal structure of unliganded EGFR [26] reveal a tethered configuration, in which domains II and IV are closely interacting via a tethering loop (Figure 4). Domains I and III are too far away as for the ligand to bind them simultaneously. This architecture has also been seen for the ectodomains of ErbB3 and ErbB4 [24, 25]. In the case of ligand bound ectodomains, two structures lacking the majority of domain IV were solved [23, 50] in which the dimerization loop of domain II, that was involved in maintaining the tethered conformation in the unliganded monomer, is now interacting with a second EGFR ectodomain. A more recent crystal structure shows the same architecture but this time including domain IV, which also seems to be involved in the interaction between two ectodomains but to a lesser extent than the dimerization loop of domain II (Figure 4) [51]. In either structure, two molecules of ligand are needed to stabilize the dimeric ectodomains (2:2 interaction), which makes EGFR a somewhat special receptor, in which the dimer interface is not stabilized by the ligand. Domains I and III are now close enough as to allow ligand to bind them simultaneously. The crystal structure of the liganded ectodomains shows a symmetric architecture with an overall symmetric heart-shape. This ectodomain arrangement has also been seen in electron-microscopy studies [81, 82].

From these studies, the mechanism of EGFR activation seemed to be easy: the ligand would allow the untethering of domain II, poising the dimerization loop for in-

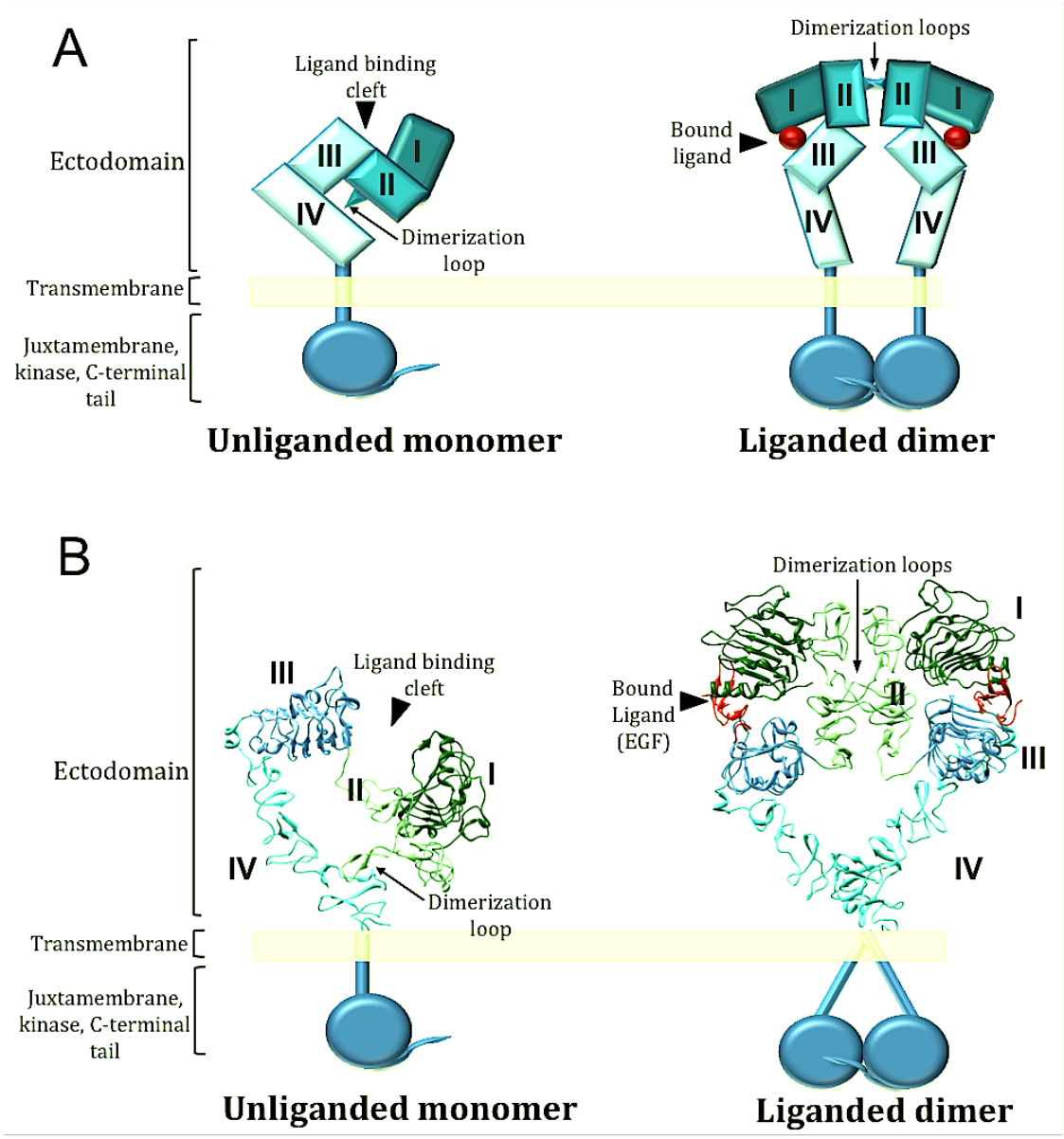


Figure 4: Depiction of unliganded and liganded EGFR molecules. Cartoon showing an unliganded, inactive monomer (left) and an active, liganded dimer (right). Ligand is depicted as a red sphere. B) Left, crystal structure of an inactive, unliganded ectodomain monomer (1NQL) coloured according to the subdomain boundaries (domains I to IV). Right, crystal structure of a liganded ectodomain dimer (3NJP) coloured following the same scheme as for the unbound structure. Ligand (EGF) is coloured in red.

teraction with other receptor. However, mutation studies of the residues involved in the tethering interaction failed to demonstrate this mechanism. In studies where soluble ectodomains were used, disrupting the tethering arm increased the affinity towards the ligand [83], which supports the idea of domains I and III coming in closer proximity. Yet another prediction of this model is that, if the tethering interaction is disrupted, the ectodomains should associate via their dimerization loops in the absence of ligand. However, mutational studies have failed to detect any ligand-independent dimerization of this kind [50, 84]. These effects are also observed in *in vivo* experiments. If the disruption of the tethering arm caused an increased rate of ligand-independent dimer formation and activation, this should be reflected in an increase of phosphorylation of the kinase domains. The mutational studies have shown that this is not the case [85, 86], although disrupting the tether increased the apparent binding affinity of EGF receptors [85, 87]. In [87] it is suggested that low affinity receptors represent tethered and extended configurations in a dynamic equilibrium. These and other studies on the disruption of the tethering interactions show that there must be other control mechanisms that maintain EGF receptor in its inactive state [6, 58] and that control its affinity to ligands.

3.3.2 Kinase domain structure and activation

After ligand-mediated ectodomain dimerization, the signal is transmitted along the EGFR molecule and results in the tyrosine kinase activation and phosphorylation of several tyrosine residues. These, in turn, will serve as recruitment points for different adaptor proteins containing PTB and SH2 domains, among others [88]. The tyrosine kinase domain of the EGF receptor can be divided into two subdomains, called the N- and the C-lobes. Between these two domains the so-called activation loop can be found (Figure 5, adapted from [89]). In the N-lobe, another important regulatory element is

found, the αC helix (Figure 5). A conserved glutamate residue is found in this helix, and in the active conformation it helps in the coordination of an ATP molecule with the N-lobe [89]. This organization is common to many other tyrosine kinases [89], in which the phosphorylation of a conserved tyrosine residue in the activation loop is necessary for the transition from the inactive to the inactive conformations. However, in the case of the EGF receptor, phosphorylation of this residue (Y845) does not seem to be involved in the kinase domain activation [35, 90].

In fact, the first crystal structure reported for the EGF receptor kinase [70] showed that it adopts an active conformation with Y845 unphosphorylated. It became clear then that there must be an alternative mechanism that prevents the EGFR kinase to be constitutively active. Other structures of the EGFR kinase domain in an inactive state [71, 91] were solved. They showed an organization that closely resembles the inactive state of Src tyrosine kinases and CDK2 (cyclin-dependent kinase 2) [92, 93]. Src kinases are found in an inactive state when bound to their SH2 and SH3 domains, and CDK2 is inactive when not bound to cyclins (Figure 6, adapted from [89]). The key features of this conformation are the position of the αC helix, which is facing away from the N-lobe and preventing the correct positioning of the conserved glutamate residue needed for the active conformation [89]. The activation loop in the inactive conformation also forms a short helix that helps in the stabilization of the αC helix away from the N-lobe. This is known as the CDK/Src-like conformation. In 2006, another crystal structure showed that the EGFR kinase in the active conformation forms an asymmetric, head-to-tail dimer [52]. The C-lobe of the activator kinase interacts with the N-lobe of the receiver kinase (Figure 6, adapted from [89]). This is also the mode of action for the allosteric regulation of ErbB4 and ErbB2 kinase domains [13, 68]. Mutations that disrupt the interaction interface, or binding to a cellular inhibitor of the EGFR kinase (the Mig6 peptide) prevent the kinases from adopting the active conformation [52, 91].

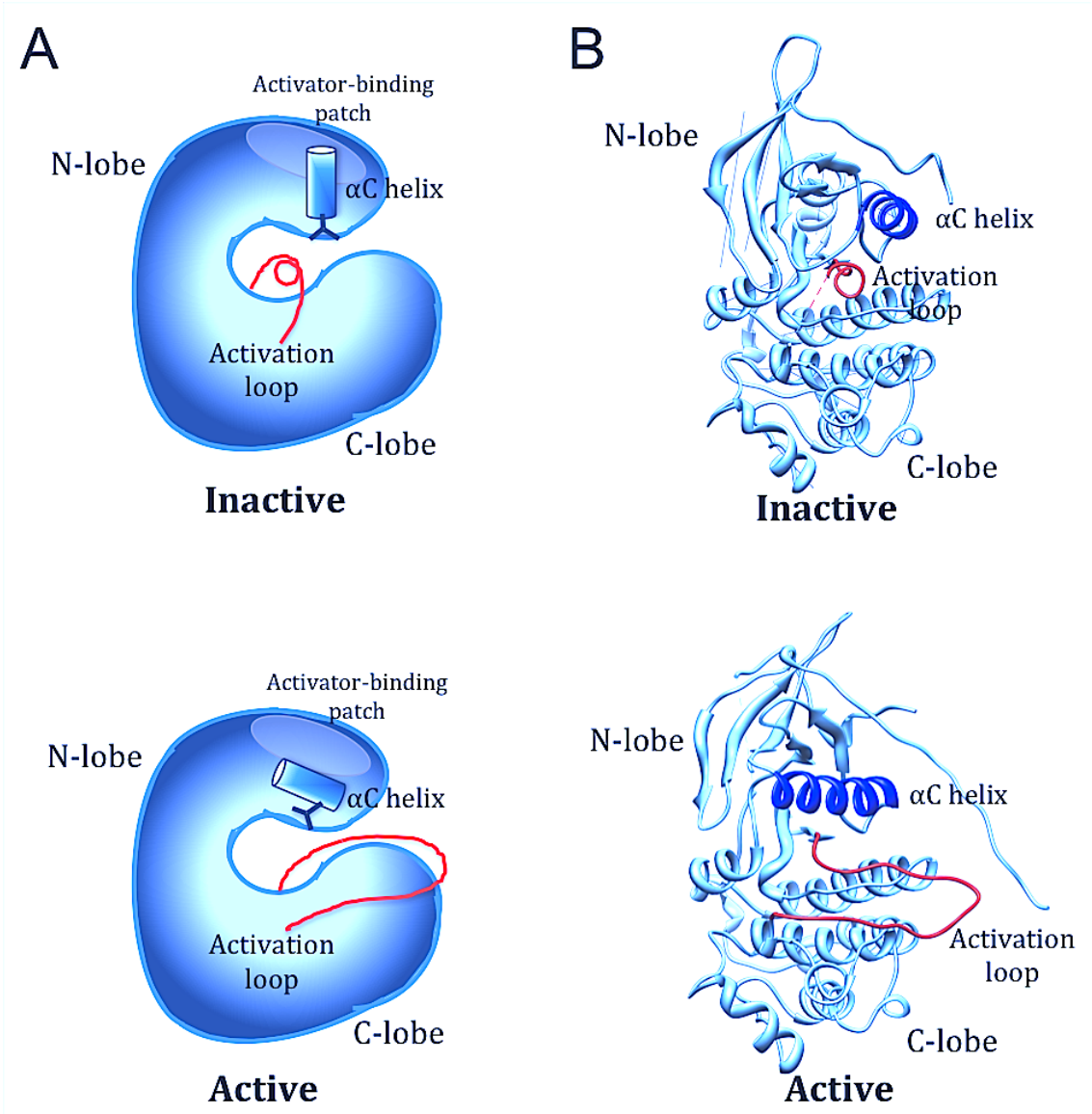


Figure 5: Conformations of the inactive and active kinase of the EGF receptor A) Cartoon showing the configuration of the αC helix, dimerization loop and activator-binding patch in the inactive (top panel) and active conformations (bottom panel). B) Crystal structures showing the same elements as in A) (PDB IDs, 3GT8 for the inactive and 2GS6 for the active conformations).

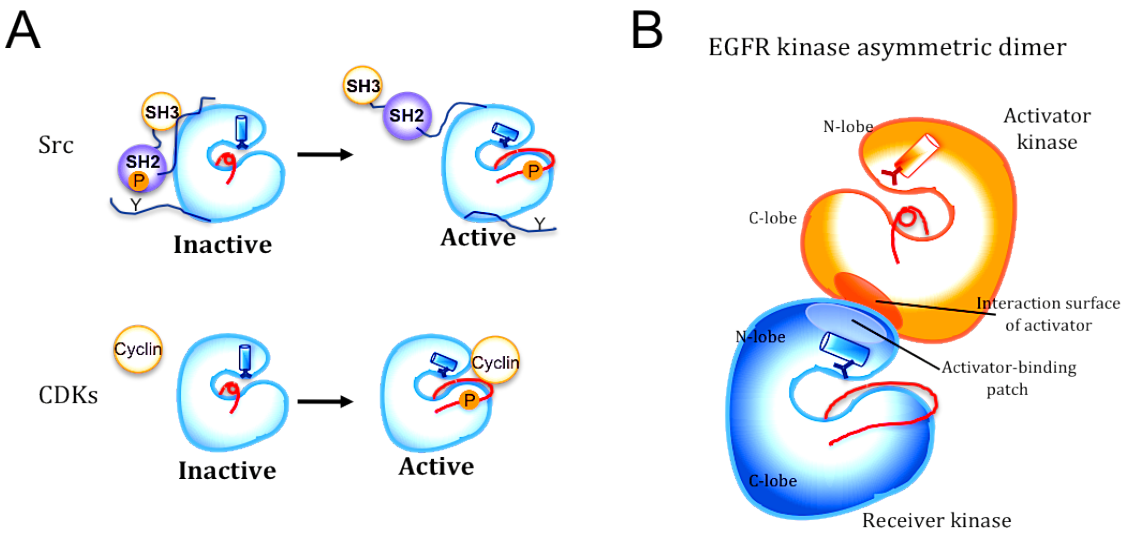


Figure 6: Comparison between Src/CDK activation and asymmetric EGFR kinase dimersA) Src kinases are active when not bound to SH2 and SH3 domains, while CDKs are only active after a cyclin has bound. B) EGFR asymmetric dimer, in which the activator kinase interacts with the receiver via the activator-binding patch. The receiver kinase adopts the same configuration as the one seen in Src kinases and CDKs. Adapted from [89].

However, the affinity of CDK2 and Src kinases towards their substrates is very high compared to EGFR kinases. Besides, isolated EGFR-family kinases have a very low basal activity and only increase their catalytic activity when the local concentration is raised, such as by attachment to lipid vesicles [52, 94]. The dimer interface seems to be very weak, and must be stabilized in some way. In fact, recent studies have shown that the juxtamembrane domain is enough to generate an asymmetric dimer in vitro [89]. It is well known that the juxtamembrane portion of the EGF receptor is involved in its proper activation [95, 96], and recent structural studies, both in EGFR and ErbB4 kinases, have shown that the C-terminal part of the juxtamembrane domain is directly involved in the stabilization of the asymmetric dimer [63, 97, 98].

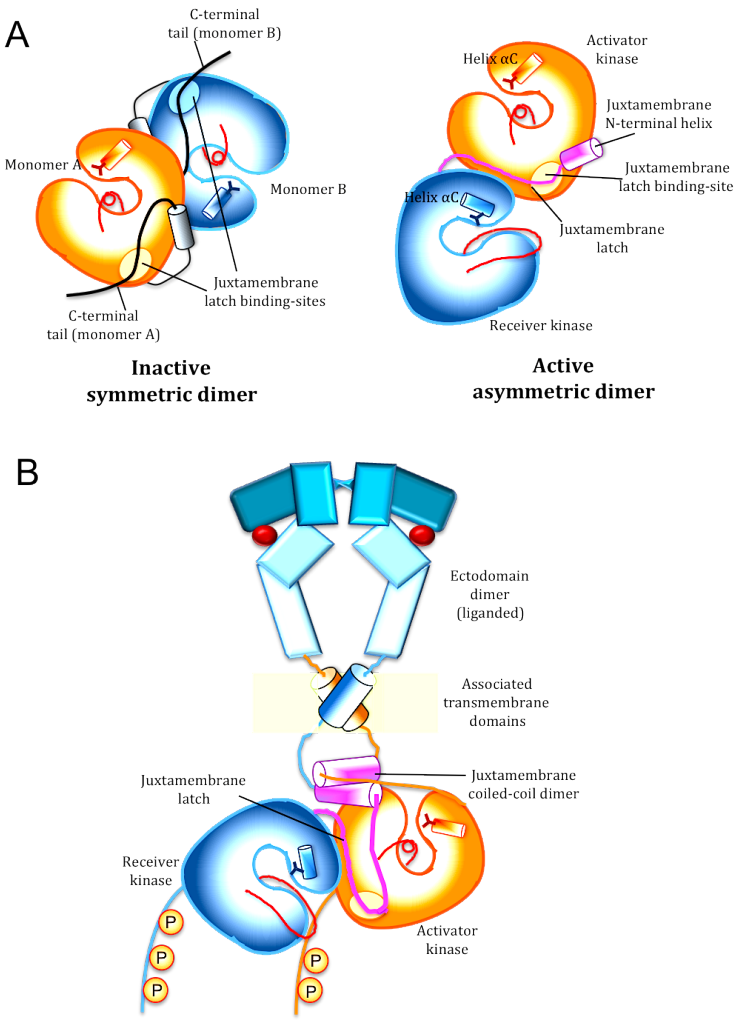


Figure 7: Activation of the EGF receptor kinase domains A) Depiction of an inactive, symmetric dimer (left panel). The kinase domains interact via their N-lobes (PDB 3GT8). Right panel, depiction of an active, asymmetric EGF receptor kinase dimer including the full juxtamembrane domain. The activator kinase interacts with the receiver via the C-lobe and N-lobe, respectively. The juxtamembrane domain of the receiver kinase wraps around the activator kinase, and the N-terminal part forms a helix (PDB 3GOP). B) A model of EGF receptor activation, showing the entire molecule. After ligand dependent dimerization, the juxtamembrane domains of both kinases form a coiled coil that brings them to an asymmetric, active configuration, allowing their transphosphorylation. (Domains not to scale). Adapted from [89].

In fact, the C-terminal tail of the activator kinase hides the juxtamembrane segment in its inactive conformation [63, 97, 98]. The current activation model proposes that the two juxtamembrane domains of the EGFR kinase dimer form a coiled-coil that is stabilized by the dimerization of the transmembrane helices (Figure 7, adapted from [89]). Although it seems clear that the transmembrane helices are able to interact [99–103], although the interaction interface does not seem to be very well defined [51]. The association of the juxtamembrane domains would in turn stabilize the asymmetric active conformation as opposed to the symmetric inactive one . However, in a recent structural study of near full-length activated EGFR homodimers, the kinase domains seem to be present in three different states: a putatively asymmetric, active dimer, a symmetric, inactive dimer, and symmetric monomers [82]. Mutations and binding to known tyrosine kinase inhibitors seem to change the equilibrium among these three states. The fact that a single conformation of the ectodomains (dimerized) can be coupled to several kinase configurations reveals a greater complexity than previously established, and suggests another set of regulatory mechanisms.

3.4 Ligand binding features of the EGF receptor

3.4.1 Curvilinear Scatchard plots of EGF receptors

Almost 30 years ago, ligand binding experiments of the EGF receptor *in vivo* revealed a concave-up Scatchard plot [104, 105], characteristic of two-site binding or a negative cooperativity event, and have been a subject of research ever since [85, 86, 95, 106–109]. These experiments reveal two affinities of the EGF receptor for its ligand, a high-affinity state with K_d = 0.01-0.1 nM, accounting for about 10% of the receptors, and a low affinity with K_d = 1-10 nM, accounting for about 90% [110]. The affinity ratio is thus about 100. Several models have been proposed in order to explain this behaviour, which include the assumption that monomeric EGF receptors are responsible for the low-affinity binding and that dimeric receptors represent the high-affinity state [45, 111], heterogeneity in the spatial distribution of EGF receptors [112], invoking an unknown intracellular component that would modulate the affinity of the EGF receptor toward its ligand [87, 113] and assuming that the extended and the tethered configurations of the extracellular domains represent the high- and low-affinity states, respectively [23, 26, 47, 50]. However, a series of recent experiments have shown that negative cooperativity between two EGF receptors is responsible for the apparent two affinity sites [95, 107, 108, 114, 115]. In this model, the unliganded EGF receptors are able to form inactive dimers (Figure 8 adapted from [107]).

The binding of a first ligand molecule (high affinity) would generate a structural change on the second site and reduce its affinity. This model, supported by structural data, has been shown for the *Drosophila* EGF receptor [116]. Negative cooperativity is also known to modulate the affinity of the insulin receptor [5, 117, 118], and the regulatory mechanisms of both insulin and EGF receptors are known to be strongly related [5, 119]. In this way, the extracellular domains of the *Drosophila* EGF receptor

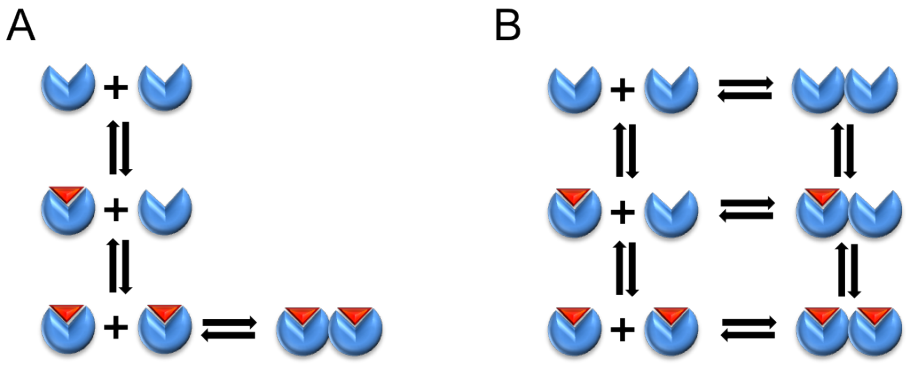


Figure 8: Activation pathway of EGF receptors. Receptors are depicted by indented circles and the ligand by a red triangle. A) Classical activation model, in which unliganded receptors are in a monomeric state. Ligand binds to its receptor in a sequential fashion. Once both sites are occupied, the receptors dimerize. B) Proposed model of EGF receptor activation in an aggregating system. Receptors are now able to dimerize in the absence of ligand without being activated. The binding of the ligand to the first site generates a negative cooperativity event that reduces the affinity of the second site (adapted from [107]).

have a symmetric organization when unbound. They form an asymmetric structure after the first ligand binding event that allows negative cooperativity to occur (Figure 9, adapted from [116]). However, in the case of the human EGF receptor, a structure of a preformed, unliganded receptor is still unknown, as are the precise regulatory mechanisms that allow ligand binding heterogeneity to take place. It is important to note, however, that the high-affinity state of EGF receptors is lost upon solubilisation [110]. A recent study shows that EGF receptors reconstituted into proteoliposomes present only low affinity binding [120]. In the same way, when vesicles containing EGF receptors are isolated from cells, these are able to bind ligand only with low affinity, even though in the cells they presented both affinities [121].

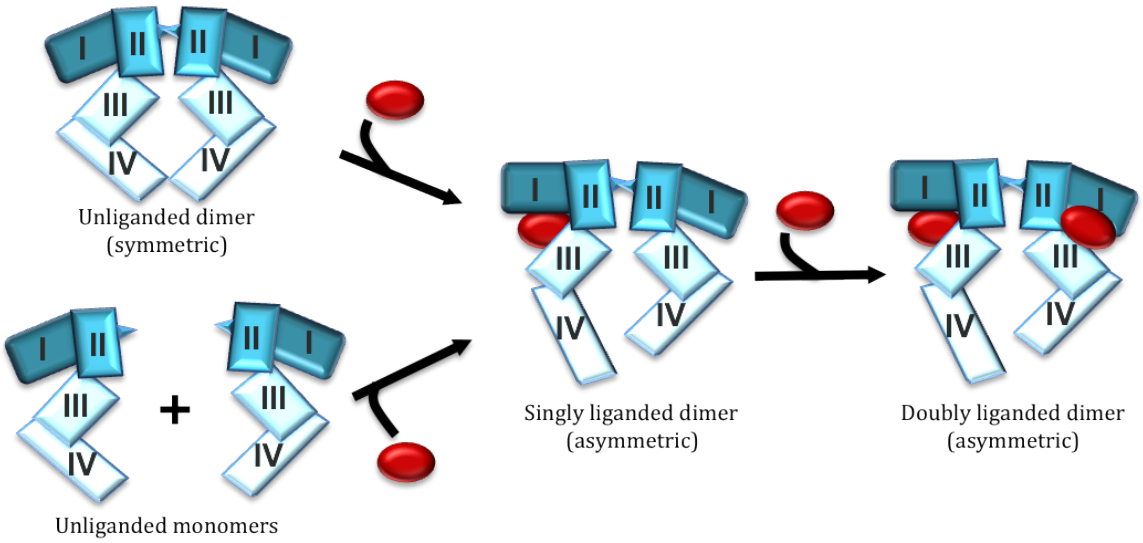


Figure 9: Ligand binding model for *Drosophila* EGF receptors. Initially, unbound receptors exist either as symmetric dimers (top left) or extended monomers (bottom left). Binding of SpitzEGF (red ellipses) to one of the sites generates a second site with less affinity in an asymmetric dimer (middle). The second binding event generates a doubly liganded dimer that shows a general ectodomain asymmetry, as opposed to the human liganded EGF receptor. Adapted from [116].

Taken together, these results suggest that the two-site affinity binding in EGF receptors is not only due to a simple negative cooperativity event caused by receptor association, but also to an external cellular agent involved in maintaining or regulating the high-affinity state, as was hypothesized before [87, 113].

3.4.2 Preformed, unliganded dimers and oligomers of EGF receptor

If an allosteric mechanism were responsible for the two affinity sites seen in the EGF receptor, we would expect that these are able to form dimers in the absence of ligand. In fact, this observation has been made several times and with a very diverse array of techniques in living cells [44, 76, 111, 122–131] and in purified receptors [120]. In some of these studies, higher order oligomers are also detected. In this sense, EGF receptors are similar to insulin receptors, with a main difference: insulin receptors form constitutive, covalently bound dimers in the absence of ligand [5, 117, 118], while ligand-independent dimerization of EGF receptors occurs spontaneously and in a dynamic fashion. In order to understand the activation mechanism of the EGF receptor, a characterization of the preformed, dimeric state is needed, since it seems to play a very important role in its activation.

For these reasons, I decided to centre my research in developing a method that allows the purification of unliganded EGF receptor dimers, and that can be used as a general system to purify and characterize not only EGF receptor homodimers, but also any other dimeric receptors or proteins, including other heterodimeric members of the EGF receptor family. For this goal, a tool that is simple yet stable enough to withstand a whole purification procedure is needed. The idea is to use of a dimerization module that allows the proteins to stay in close contact during the purification process. There are several dimerization strategies exploited by proteins, but one of the simplest and most specific is the coiled-coil interaction. Besides, coiled-coils are one of the only protein structures that can be rationally designed in order to modify their affinity and behaviour, a very useful feature for the purposes of this investigation.

3.5 Coiled-coil design

Proteins can associate using a wide range of mechanisms. One of them includes the use of coiled-coils, and some analyses suggest that up to 10% of eukaryotic proteins contain coiled-coil sequences [132]. Since coiled coils have a very characteristic primary sequence, it is relatively easy to predict the areas of a protein that will adopt this conformation. Coiled-coils have a basic unit that is repeated along their sequence, called a heptad [133, 134], represented by the letters *abcdefg*. The residues occupying the positions *a* and *d* have a hydrophobic character, and permit the association of different coiled-coil strands following the knobs-into-holes model proposed by Crick in 1953 [135]. All the residues in positions *a* and *d* associate, creating the so-called hydrophobic core. The members of the coiled coil adopt an alpha helical structure and wrap around each other, creating a higher-order supercoil. The residues of the heptad in positions *e* and *g* are usually charged, and contribute to the specificity of the coiled-coil association. The residues in the rest of positions are normally polar, and are exposed to the solvent [133, 136, 137]. A scheme depicting the general architecture of a coiled-coil, in helical wheel representation, can be seen in Figure 10.

Coiled-coils can homo- or heteroassociate, forming dimers, trimers or higher-order oligomers, in parallel or antiparallel configurations. These structures are an active subject of research, and have proved to be very good candidates for *de novo* protein design [138–145]. Synthetic coiled-coils offer the possibility of working with rationally designed protein structures that can be tested *in vitro*, and minimize the possible interactions with cellular proteins. Properties like the association constant, specificity and oligomerization state can also be controlled to some extent [141]. The number of heptad repeats can also be modified, and so very powerful yet small dimerization tags can be generated. In this sense, the use of such a tag would prove to be very valuable

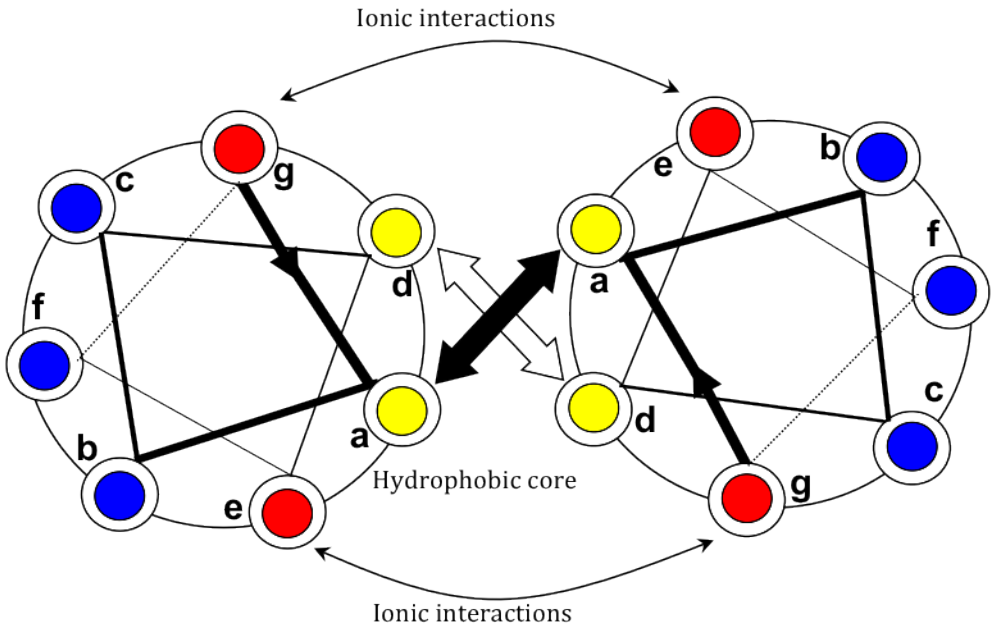


Figure 10: Helical view representation of a dimeric coiled-coil interaction. The scheme shows the first heptad repeat *abcdefg*, that can be repeated afterwards. Hydrophobic residues are shown in yellow, ionic ones in red and polar aminoacids in blue. The coiled-coiled is stabilized by a continuous hydrophobic core and the specificity is given by the interactions between the residues *e* and *g*.

in order to characterize preformed EGFR dimers, since these are known not to survive the purification procedure, as has been commented on before. In this work, we have chosen one of these *de novo* designed coiled-coil heterodimerization tags, both because it is only 21 aminoacids long and because it seems to behave as a strict heterodimer in solution [141, 146]. Its structure was solved by NMR ([146], PDB code 1U0I) and can be seen in Figure 11 B.

The combination of a small tag, which would prevent any steric hindrance effects with the tagged protein, and a high specificity for its heterodimerization partner make this synthetic coiled-coil pair a very good option for our purposes. This pair is termed the E3/K3 tag, and consists of two separate three-heptad repeats. E3 contains the

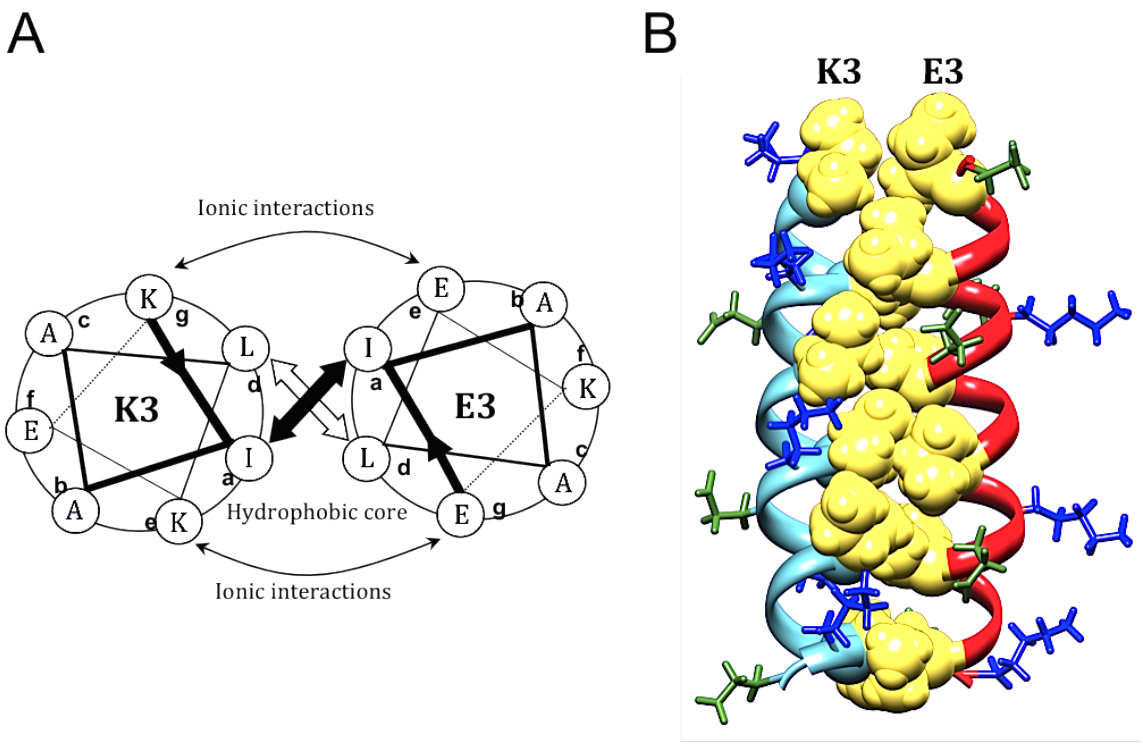


Figure 11: Depiction of the E3/K3 interaction. A) Helical wheel representation of the E3/K3 coiled-coil . B) NMR solution structure of the E3/K3 coiled-coil (PDB code 1U0I). The backbone of the K3 peptide is shown in green, the backbone of the E3 in red. Leucines and isoleucines are depicted by yellow spheres, glutamic acids in green and lysines in blue.

sequence $(EIAALEK)_3$ and K3 is composed of three repeats of KIAALKE. A helical wheel representation can be seen in Figure 11 A. In the first part of my work, I have described the interaction when tagging soluble model proteins. In the second part, I successfully purified EGFR preformed dimers with these two tags.

4

Results

4.1 Characterization of the E3/K3 interaction in fusion proteins

Before starting with the purification procedure of EGFR, I performed different experiments in order to elucidate how the heterodimeric coiled-coil pair E3/K3 behaves when fused to a protein. As a simple yet informative method, the pair Venus/Enhanced Cyan Fluorescent Protein (ECFP) was chosen and tagged with either E3 or K3 preceded by a TEV cleavage site and followed by a StrepTagII and a His–tag, respectively. Several experiments regarding the purity and stability of the E3/K3 system were performed. In particular, I carried out FRET experiments in order to test the dimerization ability under different conditions. Once the synthetic system was validated and described, I continued with the purification and characterization of a more complex system as EGF receptors are.

4.1.1 Influence of peptide linkers and concentration in the solution behaviour of the K3 and E3 chimeric proteins

It has been shown that the K3 and E3 peptides form a stable heterodimeric coiled coil interaction, but the free monomers have a different behaviour. A fraction of free K3 peptides are present as homodimers in solution [147]. For this reason, I have used the fluorescent variants Venus and ECFP (enhanced cyan fluorescent protein) to test the behaviour of the E3/K3 pair when fused to a protein. Both chimeras (Venus-K3 and ECFP-E3) were cloned into a pET-21 plasmid for *E. coli* expression, but the Venus construct was preceded by a small S-tag at the N-terminus in order to increase its size and differentiate it from the E3-tagged ECFP in SDS-PAGE analysis. In this study I show that both the concentration and the neighbouring residues seem to affect the behaviour in solution of K3. The endogenous sequence of the pET-21a linker preceding the hexahistidine tag (KLAATLE) strongly resembles the heptad repeat of the K3 peptide (KIAALKE). The number of heptad repeats is crucial in the solution behaviour of synthetic coiled coil sequences [141, 142]. The purified SVenus-K3-KLAATLE-his behaves as a tetramer at higher concentrations, and the minimum molecular weight species seen is a dimer (Figure 12 A). However, when the KLAATLE linker is exchanged with a polyglycine linker (GGSG, denoted pG), the homodimerization of the fusion protein at lower concentrations is disrupted (Figure 12 B) although this effect is still seen as the concentration rises, just as with the KLAATLE construct. This points to an intrinsic tendency of the K3 peptide to form homodimers, as shown in previous studies [147]. We further investigated the role of the C-terminal linker in the solution behaviour of K3 fusion proteins. We reasoned that if the polyglycine linker is mutated to prolines (PPSG), an aminoacid usually found at the end of helices and in loops, the limits of the heptad would be better defined, and hence the homodimerization of the

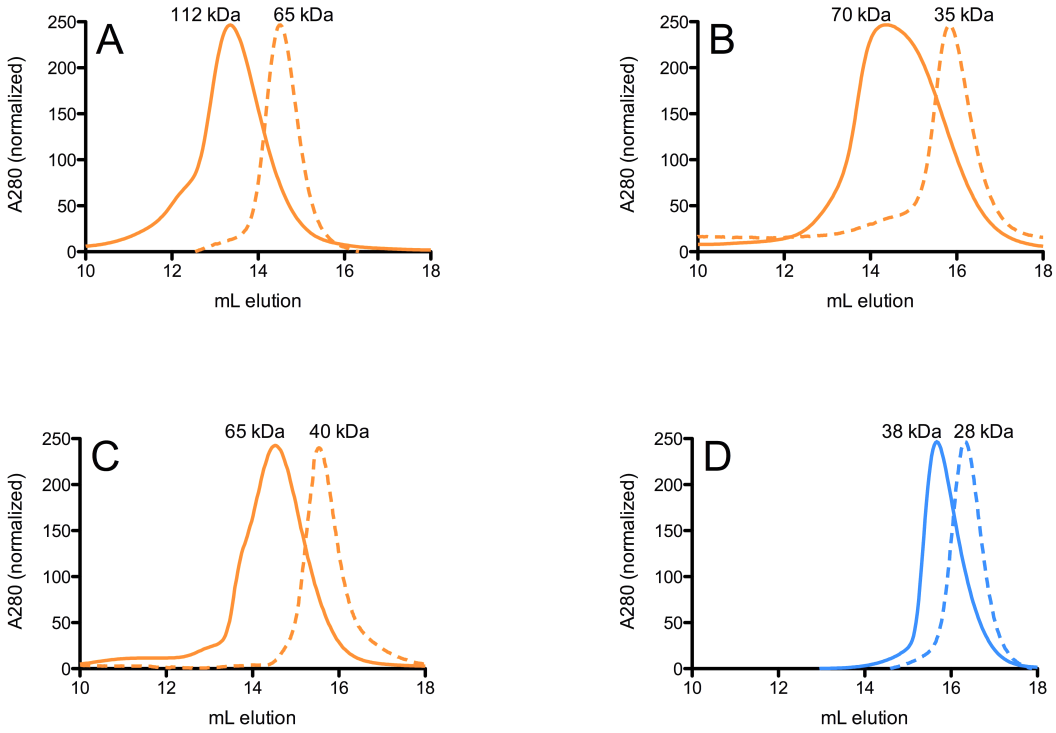


Figure 12: Effect of concentration in the association state of the different purified constructs. All samples were brought to the desired concentration in a final volume of 250 μL and loaded onto a Superdex200 column. Theoretical molecular weights: SVenus-K3-KLAATLE-His, 33.1 kDa, SVenus-TEV-K3-pG-His 33.7 kDa, SVenus-TEV-K3-PPG-His 33.8 kDa, ECFP-TEV-E3-Strp 32 kDa. The calculated molecular weights in kDa are noted above each elution peak. A) SVenus-K3-KLAATLE-his gel filtration profiles of 4 mg/mL (continuous line) and 0.2 mg/mL (dashed line) loaded sample. At higher concentrations, this construct elutes as a high order oligomer (possibly a tetramer, continuous line). At lower concentrations, the protein forms dimers (dashed line). B) SVenus-TEV-K3-pG-his gel filtration profiles of 4 mg/mL mg (continuous line) and 0.2 mg/mL (dashed line) loaded sample. At higher concentrations, this construct elutes as a dimer (continuous line) but the equilibrium is shifted towards a monomeric form with lower concentrations (dashed line). C) SVenus-TEV-K3-PPSG-his gel filtration profiles of 4 mg/mL mg (continuous line) and 0.2 mg/mL (dashed line) loaded sample. The behaviour is similar to the polyglycine linker construct in B. D) ECFP-TEV-E3-Strp gel filtration profiles of 4 mg/mL (dashed line) and 0.2 mg/mL (continuous line) loaded sample. There is a light tendency towards homodimerization, but the effect is very low.

chimeric proteins would be disrupted. However, this construct behaved exactly as the polyglycine fusion variant (Figure 12 C), which implies again an intrinsic homodimerization tendency of the K3 peptide. SDS-PAGE analysis shows unequivocally that the change in molecular weight is not due to any contamination or association with other bacterial proteins, but intrinsic to the fusion proteins (Figure 15). In the case of ECFP- TEV-E3-Strp, the concentration-dependent homodimerization effect is very low (Figure 12 D). Even though the fluorescent protein variants used here have a light tendency to homodimerize at higher concentrations [148], the fact that the ECFP fusion protein behaves as a monomer in solution, but the Venus fusions do not, rules out the possibility of a Venus- or ECFP-driven homodimerization. Besides, the protein concentration in the gel filtration assays is clearly below the concentrations needed for a Venus or ECFP homodimerizing effect. However, it is important to point out that K3 homodimerization does not hinder the formation of the E3/K3 heterodimeric coiled coil, possibly because of the higher stability of the heterodimer versus the homodimer. This is proven by the ability of SVenus-TEV-K3-pG-his to recruit all of the ECFP-TEV-E3-Strp, both in gel filtration assays (Figure 13), ITC measurements (Figure 14 B) and fluorescence assays (Figure 14 C).

4.1.2 E3/K3-driven heterodimerization of the fusion proteins and distance estimation

The tandem affinity purified species behaves strictly as a heterodimer in solution (Figure 13 A) at the concentrations tested. Figure 13 B compares the behaviour of the monomeric fusion proteins versus the heterodimer. The analysis with SDS-PAGE and Coomassie staining (Figure 15 D) together with the gel filtration profile and ITC measurements (Figure 14 B) points to a 1:1 interaction of the E3/K3 fusion proteins.

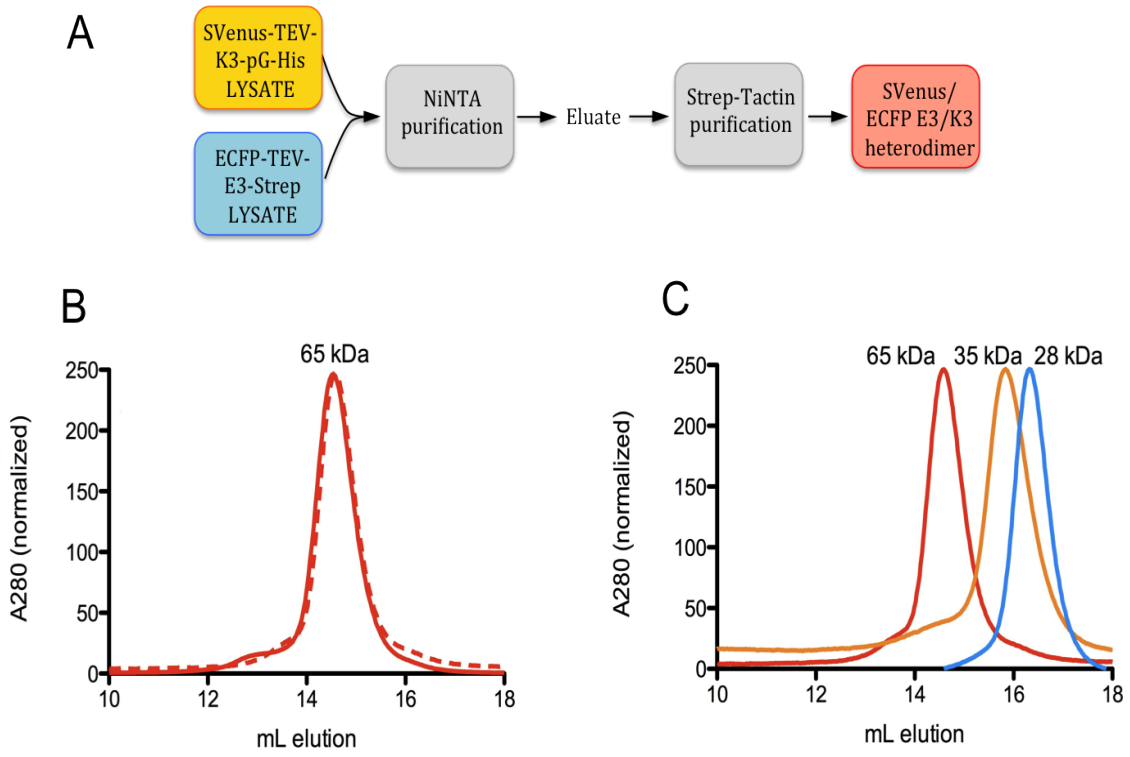


Figure 13: Gel filtration profile of the tandem affinity purified heterodimer. A) SVenus-TEV-K3-pG-his and ECFP-TEV-E3-Strp cell lysates were mixed and subjected to tandem affinity purification (StrepTactin after NiNTA). The behaviour of the eluted species at different concentrations was tested by gel filtration. All the samples were brought to the desired concentration in a final volume of 250 μL. The theoretical molecular weight of the heterodimer is 66 kDa. The calculated molecular weights in kDa are noted above each elution peak. B) 4 mg/mL (continuous line) and 0.2 mg/mL (dashed line) were loaded in a Superdex200 column. This species elutes as a heterodimer regardless of its concentration. C) Comparison of the monomeric SVenus-TEV-K3-pG-his and ECFP-TEV-E3-Strp (orange line and blue line, respectively) and the heterodimer (red line), 0.2 mg/mL each.

The free E3/K3 coiled coil has been shown to have a K_d in the range of 70 nM and interact in a one to one fashion [141, 146, 147]. ITC measurements of the K3 and E3 fusion proteins (SVenus-TEV-K3-pG-His and ECFP- TEV-E3-Strp) give a K_d of

89.3±5.2 nM, (Figure 14 B) which is in good agreement with previous dissociation constant determinations. In order to test the heterodimerization ability of the free tagged monomers, purified SVenus-TEV-K3-pG-his and ECFP-TEV-E3-Strp were mixed at equimolar concentrations (2 μM each) in 200 μL lysis buffer. The fluorescence levels in the different channels (Venus, FRET and ECFP) were measured and compared to those of a 2 μM solution of the tandem affinity purified heterodimer (Figure 14 C), and the FRET efficiency and distance estimation calculated (Figure 14 D).

The results show that the fluorescence levels of the mixed monomers and the purified heterodimer are the same, confirming that virtually all the monomeric molecules are able to heterodimerize. The FRET efficiency (E) and distance estimation (d) are also equal (E = 0.46±0.03 for the mixed samples and 0.45±0.03 for the tandem affinity purified heterodimer; d = 51.3±0.5 $\mathring{A}$ mixed and 51.6±0.5 $\mathring{A}$, tandem affinity purified).

4.1.3 K3 and E3 peptides are biologically inert when expressed *in vivo*

Even after a single affinity purification, the eluted samples are extremely clean (Figure 15), independently of the construct used. In the case of the heterodimer tandem affinity purification, the sample is >99% pure as judged by Coomassie staining (Figure 15 D).

The fusion proteins were purified from a whole bacterial lysate, and hence were exposed to all the proteins and peptides present in *E. coli*. The fact that the eluted samples are clean of any contaminants shows that the K3 and E3 peptides are inert in terms of heterologous association with other proteins. As demonstrated with the tandem affinity purification of the heterodimer, they only associate with their designed partner. These features are of extreme usefulness if the E3/K3 system is to be implemented as a general affinity purification method.

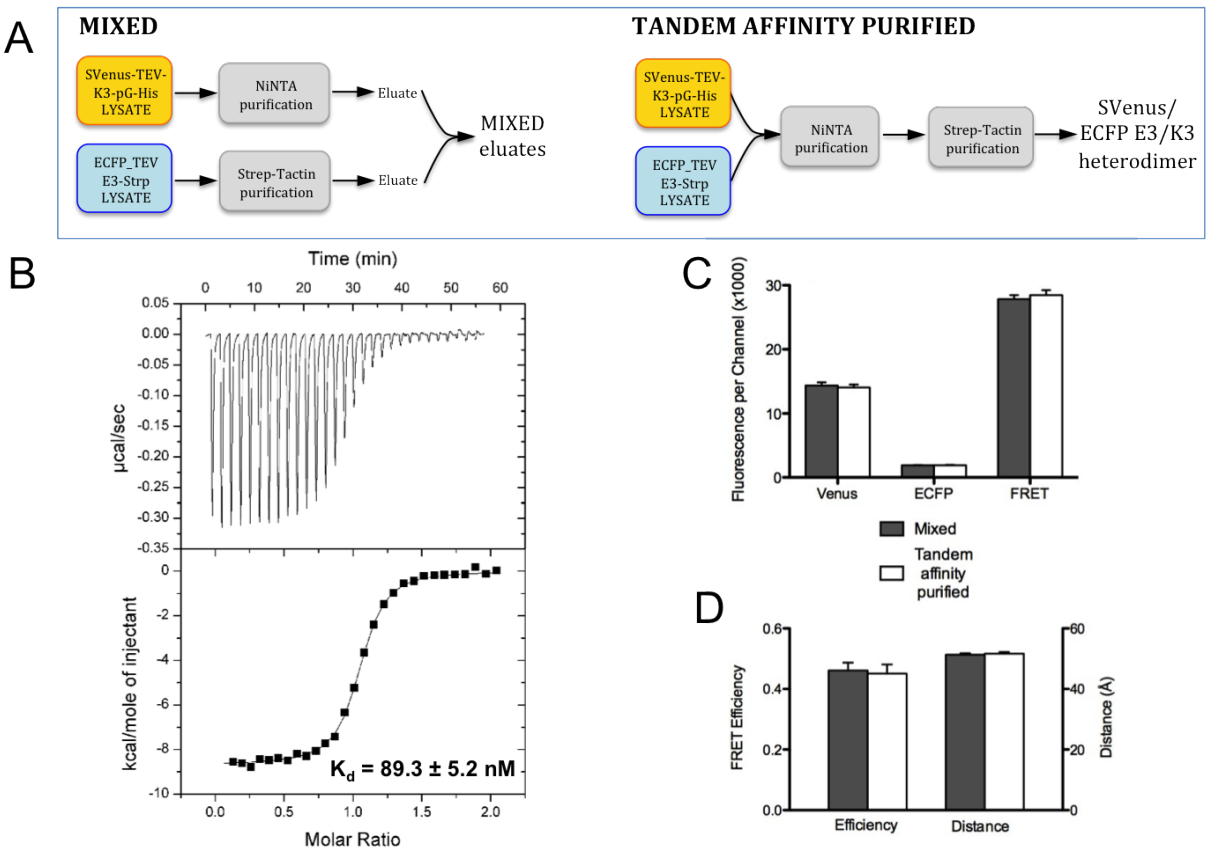


Figure 14: Heterodimerization of SVenus-TEV-K3-pG-His and ECFP-TEV-E3-Strp. A) Workflow for Mixed SVenus and ECFP eluates (left panel) and tandem affinity purified samples (right panel) B) Isothermal titration calorimetry of ECFP-TEV-E3-Strp binding to SVenus-TEV-K3- pG-His. Top figure, data from 30 injections of 100 μM ECFP-TEV-E3-Strp into the cell containing 10 μM of SVenus-TEV-K3-pG-His in lysis buffer at 25°C. Bottom figure, best Chi-square fit (continuous line) to the calorimetry data (squares), analysed with MicroCal Origin software. The calculated dissociation constant is K_d = 89.3±5.2 nM. B) Equimolar amounts of the affinity purified SVenus-TEV-K3-pG-his and ECFP-TEV-E3-Strp monomers were mixed (grey bars) and its fluorescence compared to the affinity purified heterodimer (white bars) per triplicate. Final concentration 2 μM of each species in 200 μL lysis buffer. Error bars show the standard error (σ/n). C) Comparison of efficiency and distance estimation for the mixed and tandem affinity purified samples. The similar levels of fluorescence in B) and FRET efficiency in C) demonstrate that virtually all free monomers are able to heterodimerize.

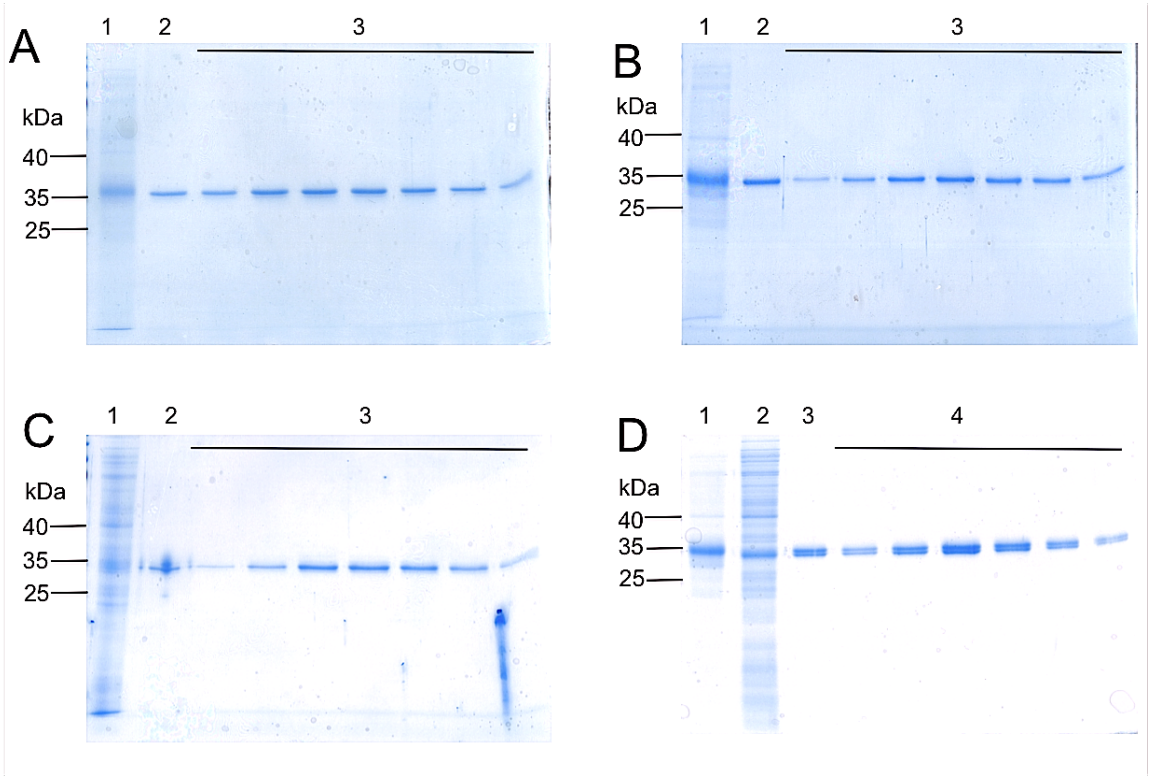


Figure 15: Analysis of samples after affinity purification and gel filtration. 12% SDS-PAGE analysis of purified samples from SVenus-KLAATLE-His (A), SVenus-TEV-K3- pG-His (B) and ECFP-TEV-E3-Strp (C), Coomassie stained, 0.5 μg per lane. Lane 1, cell lysate. Lane 2, eluate from affinity column (HisTrap or StrepTrap). Lane 3, peak from size exclusion chromatography after affinity purification, preceded and followed by the neighbouring fractions. The K3 and E3 peptides are inert in terms of heterologous binding to endogenous *E.coli* proteins when expressed *in vivo*. After a single affinity purification, all the samples are >95% pure as judged with Coomassie staining. D) 14% SDS-PAGE analysis of tandem affinity purified TEV constructs. Lane 1, SVenus-TEV-K3-pG-His cell lysate. Lane 2, ECFP-TEV-E3-Strp lysate. Lane 3, eluate after second affinity purification (StrepTrap). Lane 4 and neighbouring lanes, peak from size exclusion chromatography (below the black line). If a tandem affinity purification is performed, an extremely clean, monodisperse sample is obtained. The two monomers are clearly seen as two bands with different apparent molecular weight (lane 4 and neighbouring lanes)

4.1.4 The E3/K3 interaction is stable under different buffer conditions

In order to confirm that the fusion products still retain this stability, several conditions for heterodimerization were tested and the FRET efficiency of the sample measured. The estimated distance between SVenus-TEV-K3-pG-His and ECFP-TEV-E3-Strp in lysis buffer (20 mM Tris, 150 mM NaCl, 10 mM imidazole, pH 7.5) is 51.3±0.5 $\AA$ (Figure 16). Estimating the distance that a given linker provides between the E3/K3 tagged proteins can be crucial for their successful co-purification. If the proteins are bulky and the linker too short, the K3 and E3 peptides might fail in forming a stable coiled-coil interaction due to steric hindrance. Previous studies have shown that the coiled-coils with similar sequences to the E3/K3 coiled-coil are stable at high salt concentration [143]. Thus, if the E3/K3 system is to be introduced as a general tool for protein purification, it has to be resistant to commonly used reagents and protocols, such as detergent solubilisation and high salt washes, among others. In order to test this, the stability of the E3/K3 coiled-coil interaction in terms of FRET efficiency and distance estimation was measured in buffers containing high salt, detergent, glycerol and ethanol (Figure 16). The monomers were dialysed against the chosen buffer (except for Triton X-100, where the detergent was added to the desired final concentration) and mixed following the same protocol explained in 4.1.2. The results indicate that the E3/K3 coiled-coil is stable under a variety of buffer conditions, which makes it a very interesting tool for protein purification. The FRET efficiency test is a valid tool because the fluorescence of known amounts of the monomers in a buffer can be measured separately. This allows the effects of buffer-dependent fluorescence variation to be cancelled out. The small differences seen in the FRET efficiency could be due to the conformation of the C-terminal end, comprising the TEV cleavage site, the coiled-

coil region and the affinity tag. These could have a looser or tighter conformation depending on the buffer used, which would translate into a modest distance change between the monomers (Figure 16).

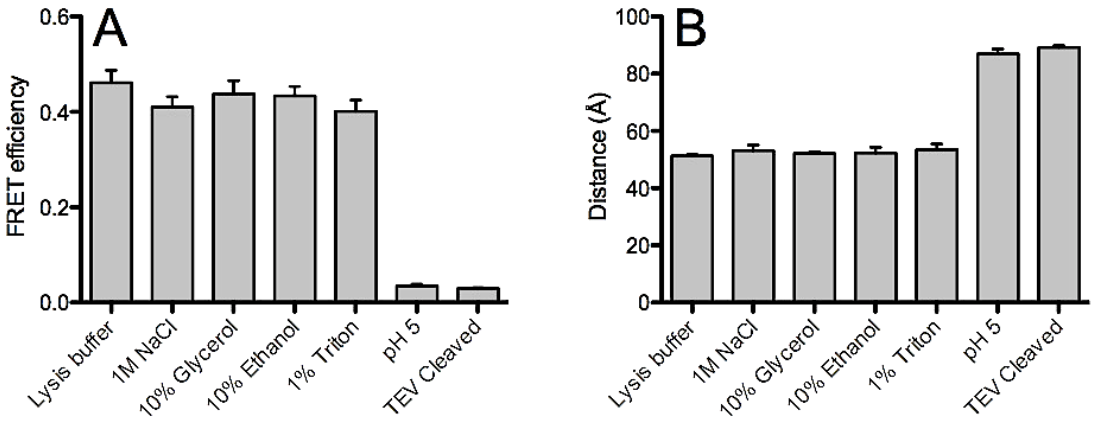


Figure 16: Stability of the fusion E3/K3 heterodimer under different buffer conditions. SVenus-TEV-K3-pG-His and ECFP-TEV-E3-Strp monomers were mixed at a final concentration of 2 μM each in different buffer conditions. The FRET efficiency, measured as the donor fluorescence decrease in the presence of acceptor, was calculated and this value used to estimate the distance between monomers. A) The FRET efficiency seems to fluctuate slightly at pH 7.5 depending on the buffer used (first 5 bars). This difference could be due to a differential packing of the linker and the coiled-coil in the different buffers, leading to more compact or stretched conformations. At pH 5, the FRET efficiency drops drastically, as compared to other buffer conditions at pH 7.5, showing a disruption of the E3/K3 interaction. This value is similar to the TEV cleaved mixed monomers, where no FRET occurs. B) Distance calculations show little difference among buffers at pH 7.5 (first 5 bars). At pH 5, the disruption of the coiled-coil interaction leads to a decrease in FRET efficiency, which in turn translates into an increase of the distance estimation. At pH5, the distance estimation is similar to that of the TEV cleaved monomers. Note that the estimated distance of the pH 5 and TEV cleaved samples (87.0$\pm$1.7 $\mathring{A}$ and 89$\pm$0.6 $\mathring{A}$, respectively) is very close to the effective limits for FRET (around 100 $\mathring{A}$).

Since the distance estimation is proportional to the sixth root of the inversed efficiency, a moderate change in efficiency produces almost no effect in the distance estimation. If the E3/K3 are cleaved by TEV treatment, the FRET signal is lost (Figure 16). A recent study has shown that the E3/K3 coiled-coil stability is pH-dependent [147]. In this study, the free E3/K3 dimer is disrupted at pH 5. E3 forms homotrimers and K3 stays mainly as a monomer with a small fraction of homodimers. We are able to reproduce and monitor this disruption at pH 5 as a sharp drop in the FRET efficiency of the mixed monomers (Figure 16). However, when this sample is subjected to size exclusion chromatography, a single-peak elution profile is seen (Figure 17).

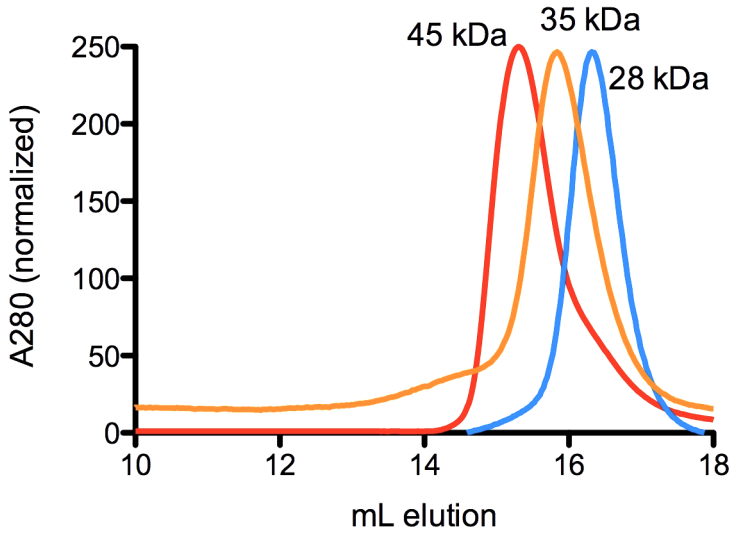


Figure 17: Gel filtration profile of the tandem affinity purified heterodimer at pH 5. All samples were brought to the desired concentration in a final volume of 250 μL and loaded onto a Superdex200 column. Red line, 0.2 mg/mL of the heterodimer at pH 5, as compared to the monomeric proteins (SVenus-Tev-K3-pG-His in orange and CFP-TEV-E3-Strp in blue).

The elution peak corresponds roughly to the size of the monomeric proteins. The fact that the peak is slightly above the size of monomeric SVenus-Tev-K3-pG-His could

mean that there is still a small percentage of dimers in the sample. The absence of the E3-homotrimers observed in previous studies is probably due to steric hindrance effects and charge effects of the linker, coiled-coil sequence and purification tag at lower pH, which would prevent homomeric interactions. As mentioned in [147], the pH-dependent disruption of the interaction can be very valuable in terms of drug delivery to eukaryotic cells, because endosomal pH lies within this range. It can also be devised as a method for elution from an affinity column if the tagged proteins are stable at pH 5.

4.2 Purification of unliganded EGFR dimers for structural analysis

Preformed EGFR dimers are known to occur *in vivo* and to be involved in high-affinity ligand binding (see sections 3.4.1 and 3.4.2). However, these dimers are only stable *in vivo*, and the purification procedure, if not using chemical cross-linkers, disrupts the interaction. For this reason, I have used the heterodimeric E3/K3 coiled-coil as a tool to allow two EGFR receptors to stay in close contact, yet being able to retain their flexibility. The chosen expression system are CHO-K1 TetOn cells. They don't have endogenous EGFR, are known to express high amounts of protein and the glycosylation pattern is very similar to that of human proteins. Besides, the doxycycline–induced system (pBI vector, containing a bidirectional promoter) allows us to control the timing and level of expression.

4.2.1 Full length EGFR unliganded dimers are degraded during purification

Several attempts to purify the full-length EGFR dimer have been performed, since its C-terminal tail is known to be involved in the regulation of the EGFR tyrosine kinase activity [52, 149–152]. As a fast method to assess the co-elution of an E3/K3 induced dimeric species after the second tandem affinity purification, a test purification using transiently transfected CHO TetOn cells and spin columns has been used. A 10-cm culture plate is transfected with the desired plasmid (either pBI containing both full-length EGFR-E3Strp and K3pGHis; or pBI containing EGFRΔ998–E3Strp and K3pGHis), solubilised, cleared and loaded onto one of the spin columns (either NiNTA or Strep-tactin) and a normal tandem affinity purification is performed. The analysis of the fractions in Figure 18 shows that, while the purification of a C-terminally truncated EGFR dimer is possible, the full-length dimers seem to be degraded in the process.

It is well known that the purification of Δ998 EGFR is feasible in different cell lines [52, 69, 81, 97, 153] be it the kinase domain only or including the ectodomains. However, it has been shown in the literature that EGFR tends to be degraded during purification, even in the presence of protease inhibitors ([153] and this study). It is true that in a recent study [120] a full length EGFR has been successfully purified by using insect cells as an expression system. It could well be that the source of the material has a real impact in the degradation of EGFR during purification, or that it is a matter of the purification conditions (affinity tags, etc.). In any case, this effect was not further analyzed and the characterization of the unliganded EGFR dimer was done by using a Δ998 truncated receptor.

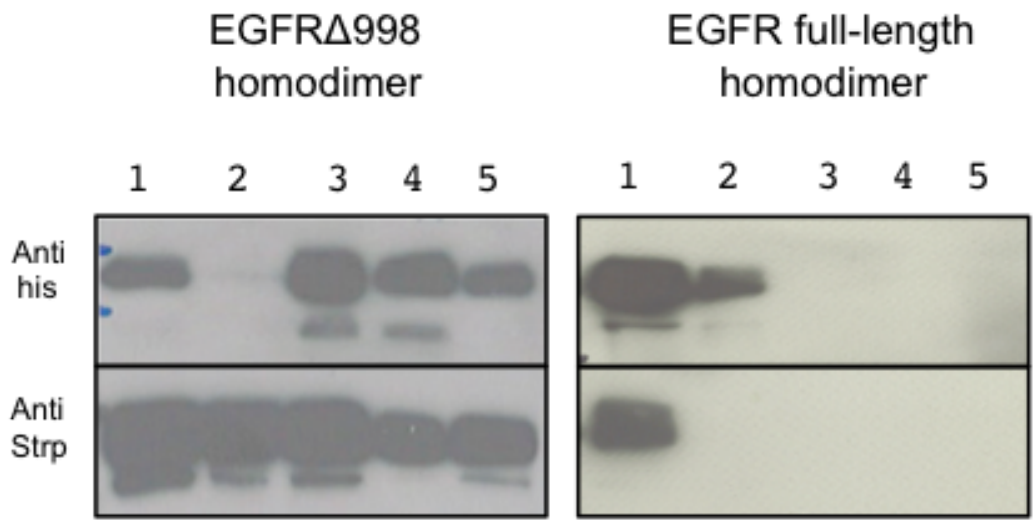


Figure 18: Comparison of EGFRΔ998 and EGFR full-length homodimer purifications. CHOTet-On cells (1x 10-cm dish) were transfected with a plasmid coding for EGFRΔ998 E3/K3 homodimer or EGFR full-length E3/K3 homodimer and induced o/n with 1μg/mL doxycycline. Soluble cell extracts were used in a tandem-affinity purification (NiNTA and Strep-tactin). Left panels, EGFRΔ998 homodimer purification in spin columns. 1: soluble cell extract. 2: NiNTA flowthrough (FT). 3: NiNTA elution. 4: Strp-tactin FT. 5: Strep-tactin elution. Right panels, EGFR full-length purification in spin columns. 1: soluble cell extract. 2: Strp-tactin FT. 3: Strp-tactin elution. 4: NiNTA FT. 5: NiNTA elution.

4.2.2 EGFRΔ998 can be purified in a dimeric form in the absence of ligand

Unliganded EGFRΔ998 dimers can effectively be purified from stable CHO-K1 TetOn lines co-expressing EGFRΔ998-K3-His and EGFRΔ998-E3-Strp. The purification scheme involves three steps: two affinity columns (in tandem, Strep-tactin followed by nickel affinity) and a gel filtration column (Figure 19). The order of the affinity purifications can be reversed, but it seems that many proteins from CHO-K1 TetOn lysates can bind to the nickel affinity column, including some contaminants that could degrade the

sample. Besides, EDTA cannot be added to the solubilisation buffer if a nickel affinity column is used as a first step, which implies that metalloproteases can still be active.

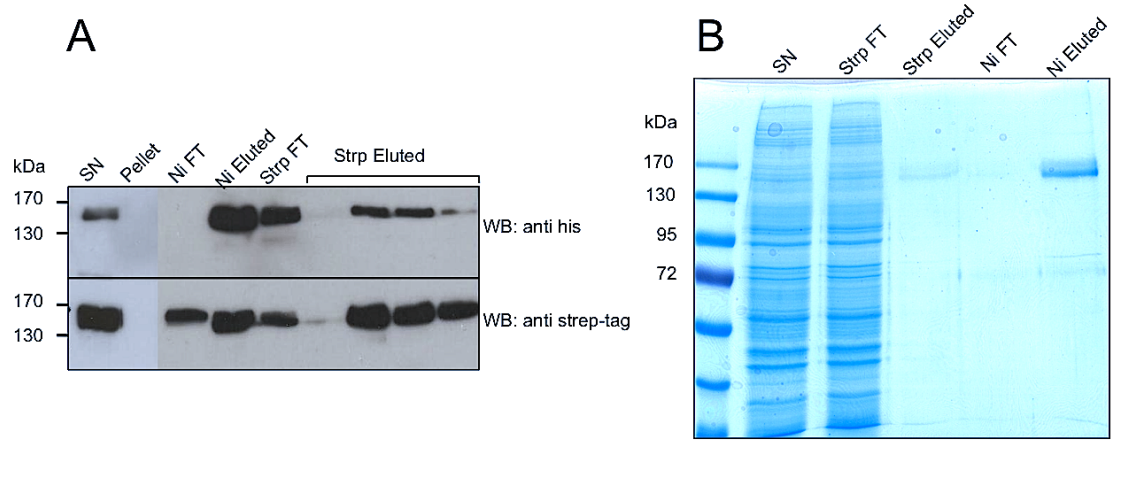


Figure 19: Tandem affinity purification of EGFRΔ998 dimers. A) Analysis of the solubilisation and co-elution of EGFRΔ998 E3/K3 stably expressed in CHO-K1 TetOn cells (10 x 175 cm^2 dishes). The K3 peptide is followed by a hexahistidine tag and the E3 peptide, by a StrepTagII. The order of purification was nickel affinity followed by Strep-tactin column. The expected molecular weight of the monomers is 118 kDa plus 20-30 kDa of glycosylation. SN, supernatant after clearing the lysate by centrifugation. Ni FT, nickel affinity flow-through fraction. Strp FT, Strep-tactin flow-through. B) Coomassie staining of fractions from 30 x 175 cm^2 dishes. In this case the order of purification was StrepTactin followed by NiNTA. The Strep Eluted lane does not correspond to the elution peak in order to maximize yield.

The typical yield from 30 x 175 cm^2 dishes after the second affinity purification is in the order of 0.1-0.2 mg/mL in a volume of approximately 750 μL, depending on the cell line that is being used. The gel filtration analysis shows that the eluted sample is mainly composed of a peak with an approximate molecular weight of 505 kDa (Figure 20 A). The glycosylated dimer has an expected molecular weight of about 320 kDa, plus the weight of a Triton X-100 micelle (90 kDa). The fact that the calculated molecular weight is higher than expected could indicate that one of the molecule's dimensions is

higher than the other (an oblate spheroid, for example). The SDS-PAGE and western-blot analyses don't show any other protein bound to the sample (Figure 20 B).

Next, I wanted to test if the unliganded sample really elutes as a dimer, and if it is able to bind ligand. For this, a tandem affinity purified sample was divided in two fractions and either loaded onto a Superdex200 column directly or treated with a fluorescent ligand (EGF-Rhodamine) plus ATP (Figure 20 C). After ligand addition, there is no change in the apparent molecular weight of the sample, and the fluorescence maximum correlates with the A280 peak. This result indicates that the original, non-treated sample elutes as a dimer, and that it is able to bind ligand in a normal way. These samples were analyzed with SDS-PAGE followed by silver staining. The gels show no other bound proteins, and confirm that both the unliganded and liganded samples behave in the same way after gel filtration (Figure 20 D). However, in both gel filtration profiles in Figure 20 A and C, a small shoulder is seen just after the main elution peak. This peak does not show as any contaminant in the gels (Figure 20 B and D), and it is not able to bind ligand, because after ligand addition its apparent molecular weight does not change (Figure 20 C). A western-blot analysis of this peak revealed a low molecular weight band that associates with the second peak (Figure 21). We concluded that this species is a degradation product of the receptors, although the addition to protease inhibitors to various steps in the purification didn't eliminate this peak, as has been shown in the literature [153].

It must be noted though, that the second peak is not always seen. In some purifications, the protein eluted as a monodisperse peak, while in the majority the small shoulder was detected. This behaviour seems to be related to the level of expression of the stable cell lines. The monodisperse and shoulder peaks were always seen when the receptors were purified from a cell line having a low expression level as compared to another cell line, where the expression levels were higher. In this cell line, the intensity

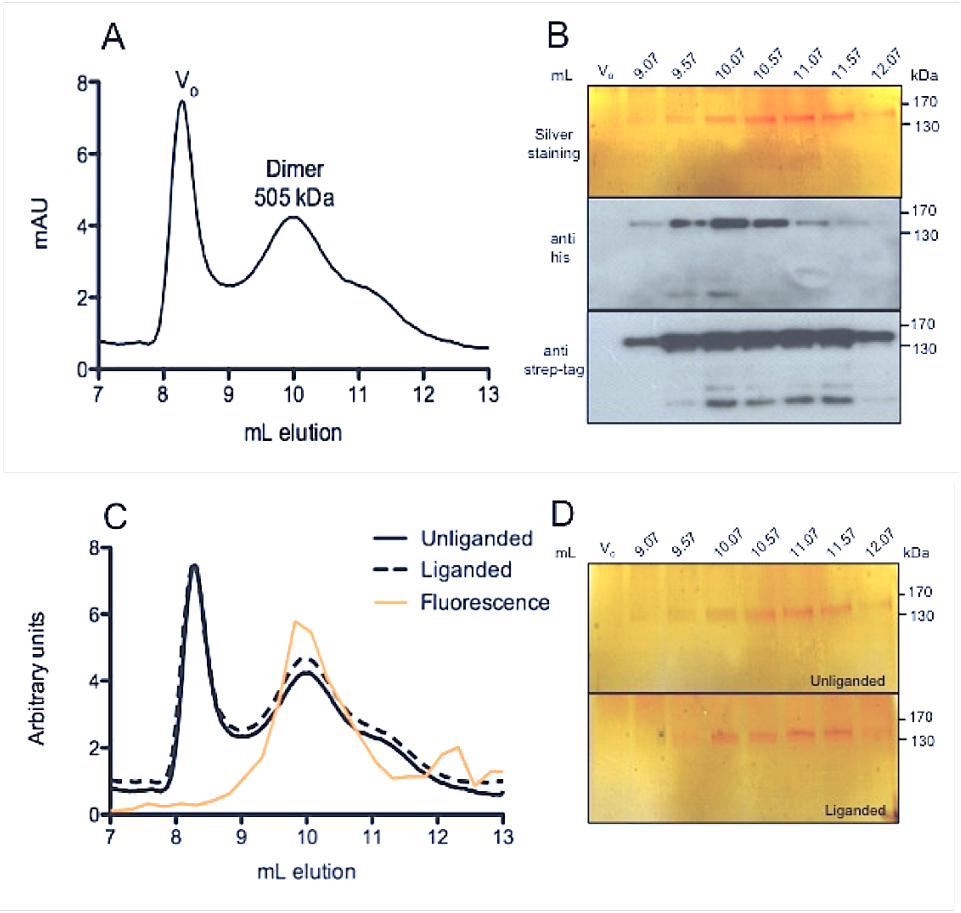


Figure 20: Gel filtration analysis of purified EGFRΔ998 dimers. A) Superdex200 profile of a tandem affinity purified EGFRΔ998 dimeric sample in DDM-150 buffer. No ligand was added. B) Silver staining of the eluted fractions from A) (top panel), anti-his tag (middle panel) and anti-strep tag (bottom panel). The elution volume of each sample is noted above the corresponding lane. C) The eluted fraction from the tandem affinity purification was divided in two and either loaded directly onto a Superdex200 column (unliganded, continuous line) or treated with 20 μM EGF-Rhodamine and 1 mM ATP prior to loading (liganded, dashed line). The fluorescence profile of this sample was measured and plotted (fluorescence, orange line) D) The silver staining of both gel filtration runs shows no change in the apparent molecular weight after ligand addition.

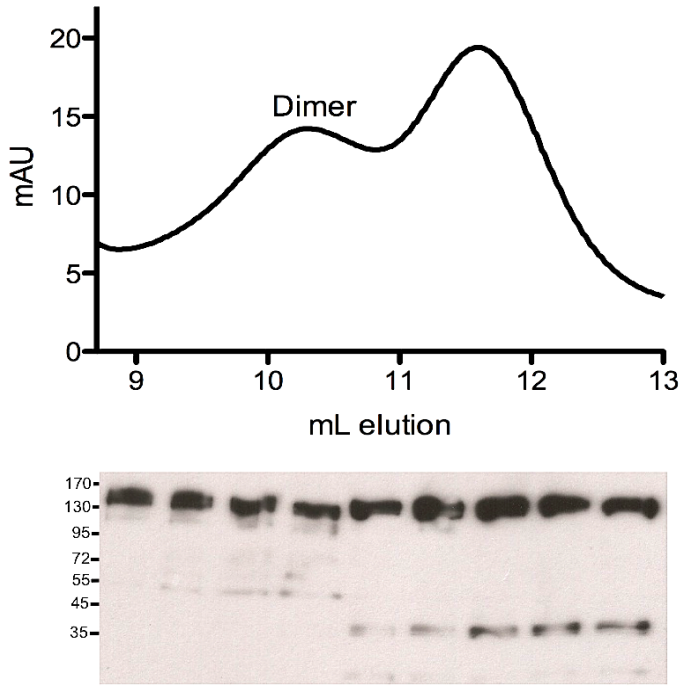
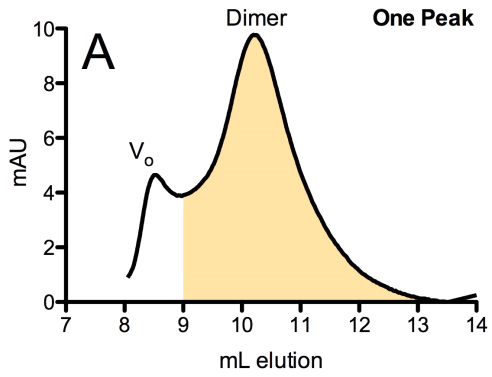
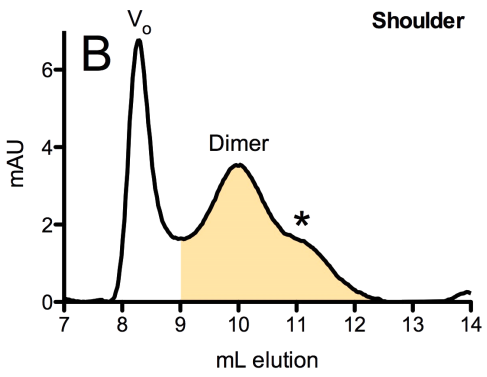


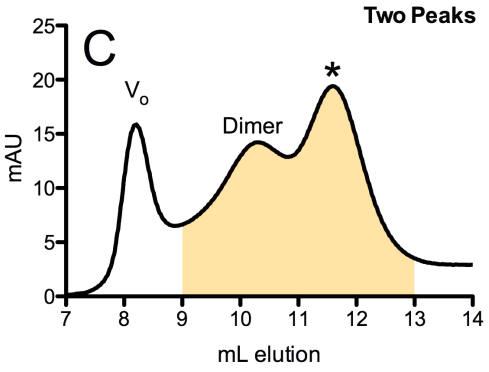
Figure 21: Analysis of the second peak after gel filtration. Top panel shows the two-peak profile obtained after gel filtration of NiNTA purified EGFRΔ998. Bottom panel, samples from the gel filtration were analyzed with WB (anti strep-tag) and show a degradation product associated with the second peak.

of the second peak was always in the same range as the first peak corresponding to the mature dimers (Figure 22).

As a way of estimating how much protein is being produced by each cell line, the area under the curve was calculated (Figure 22 D) and divided by the total surface of cells used for expression in each experiment. This is a valid method of comparison since all the cells were induced in the same conditions (80-90% confluence, 20-22 hours of induction). The cell line in Figure 22 C seems to express higher protein levels than the cell line in A and B. Thus, the intensity of the second peak seems to be related to the protein expression levels. In any case, all subsequent experiments, especially those







D

	One Peak	Shoulder	Two Peaks
Area under curve	16.6	6.0	47.1
Surface of cells, cm^2	14000	3500	15000
Ratio area/cm^2 (x1000)	**1.19**	**1.72**	**3.14**

Figure 22: Different profiles obtained in the purification of EGFRΔ998 dimers. A) Monodisperse, one peak profile obtained from a low-expression cell line. The peak corresponds to the total protein eluted from 80 x 175 cm^2 dishes. B) Shoulder peak profile obtained from the same cell line as in A). The peak corresponds to the total protein eluted from 20 x 175 cm^2 dishes. C) Two-peak sample profile obtained from a high-expression cell line. The peak corresponds to the total protein eluted from 30 x 500 cm^2 dishes. D) The area under the curve (shadowed in orange) was calculated using the trapezoid method, and divided by the total surface of cells used in each purification. With these to values, a ratio curve area/cm^2 cells was calculated.

related to structure determination, were done with fractions corresponding to the first peak (elution volume of 9.57 mL) to avoid any interference from the second peak.

4.2.3 Purification of other ErbB family dimers: EGFR/ErbB2

There are numerous studies proving that EGFR exists as preformed homodimers before ligand addition. Some other studies show that, in fact, all ErbB family members are able to form this kind of preformed dimers [44, 76, 111, 122–131]. The E3/K3 coiled-coil system is perfectly suited to study these interactions in the absence of ligand. For this reason, several attempts were performed in order to purify an unliganded EGFR/ErbB2 dimer. The constructs used consist of the C-terminally truncated versions of both receptors, because we have shown that full-length molecules are somehow degraded during the procedure. The initial attempt followed the same protocol as for the purification of an EGFRΔ998 homodimer: transient transfection of the cells, solubilisation and tandem affinity purification using spin columns (Figure 23, first panel). This experiment revealed that, although initially EGFR coelutes with ErbB2 in the first affinity purification (Strp-tactin, Figure 23, left panel, lane 3), there seems to be some process that degrades EGFR (lanes 4 and 5) and prevents ErbB2 from binding to the NiNTA column (lanes 4 and 5). We were really interested in knowing why ErbB2 was unable to bind to the column, even though it carries a his-tag.

In one study [154] ErbB2 was found to associate with Hsp90 via its kinase domain. This is probably related to the maintenance of ErbB2 in an inactive state by preventing association of the kinase domains in the absence of ligand. If Hsp90 associates with the kinase domain of ErbB2, it could prevent the E3/K3 interaction by steric hindrance. In order to remove Hsp90, the cells were treated with 10μM geldanamycin, a molecule known to bind Hsp90 and separate it from its cargo protein. In this case, EGFR seemed to escape degradation and some ErbB2 was now able to bind the NiNTA

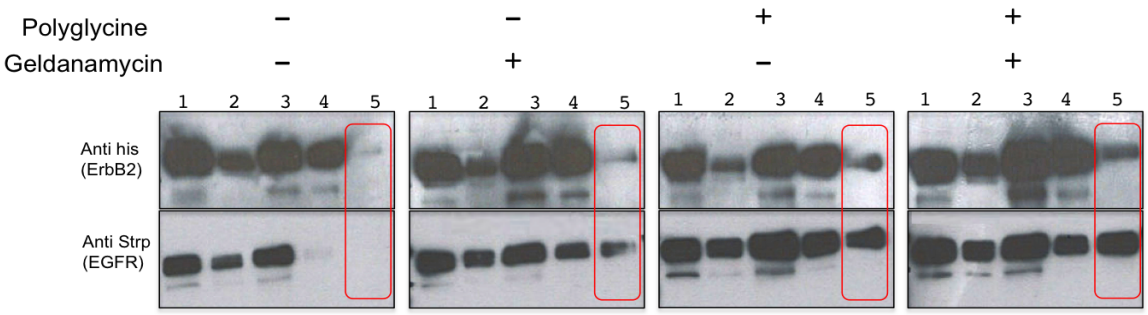


Figure 23: Influence of polyglycine linkers and geldanamycin in the purification of truncated EGFR/ErbB2 heterodimers. CHOTet-On cells (1x 10-cm dish per experiment) were transfected with a plasmid coding for EGFRΔ998/ErbB2Δ1007 E3/K3 heterodimer or a version of it including a 10-residue polyglycine arm just before the E3/K3 sequences, and induced o/n with 1μg/mL doxycycline. Soluble cell extracts were used in a tandem-affinity purification (NiNTA and Strep-tactin) in spin columns. 1: soluble cell extract. 2: Strp-tactin FT. 3: Strp-tactin elution. 4: NiNTA FT. 5: NiNTA elution.

column (Figure 23, second panel, lanes 3, 4 and 5). To further test the steric hindrance hypothesis, a different construct was used for the purification of EGFR/ErbB2 dimers. In this case, a longer polyglycine linker (10 aminoacids) was inserted between the end of the kinase domain and the beginning of the E3/K3 peptides. EGFR also seemed to escape degradation and the amount of eluted ErbB2 increased (Figure 23, third panel, lane 5). A combination of both approaches (longer polyglycine arm and geldanamycin treatment) seems to have an additive effect (Figure 23, last panel), although still the majority of ErbB2 does not bind the NiNTA column. Although we still don't completely understand why only some of the heterodimers co-elute, the purification of a EGFR/ErbB2 heterodimer seems feasible. However, the amount of starting material needed for a structural analysis would be higher. There seems to be a steric hindrance effect that would need further investigation.

4.3 Structural analysis of the unliganded EGFR△998 dimer

4.3.1 Negative staining and image acquisition of purified samples

Several attempts were performed before a suitable staining method was found. Both the ability of the stain to give a good contrast for image acquisition and the reduction of the background were thoroughly tested (Figure 24). PTA staining was too thick, and uranyl acetate gave positively stained samples. The best stains were uranyl formate and Nanovan. To completely remove background, the sample was washed 5X with water prior to staining.

The negatively stained samples seem to have a high heterogeneity. Yet the biochemical analysis has proved that the samples are clean and homogeneous. The apparent heterogeneity comes from the particle itself, and is due to a high degree of asymmetry plus the combination of different settling positions on the grid and possibly an inherent flexibility of the molecule. The maximum radius of the particles ranges from 22 to 25 nm, and is in very good agreement with the expected size from crystal structures. For other receptors, such as integrins, the negatively stained samples looked very similar to the micrographs of EGFRΔ998 (Figure 25) and the resolution achieved was 30 $\mathring{A}$ [155].

However, for image acquisition in a 300 kV Polara, Nanovan gave a very low contrast, and needed a very high cumulative electron dose for the particles to be seen. On the other hand, uranyl formate gives a very good contrast at low electron doses and hence was chosen as the stain for further analysis.

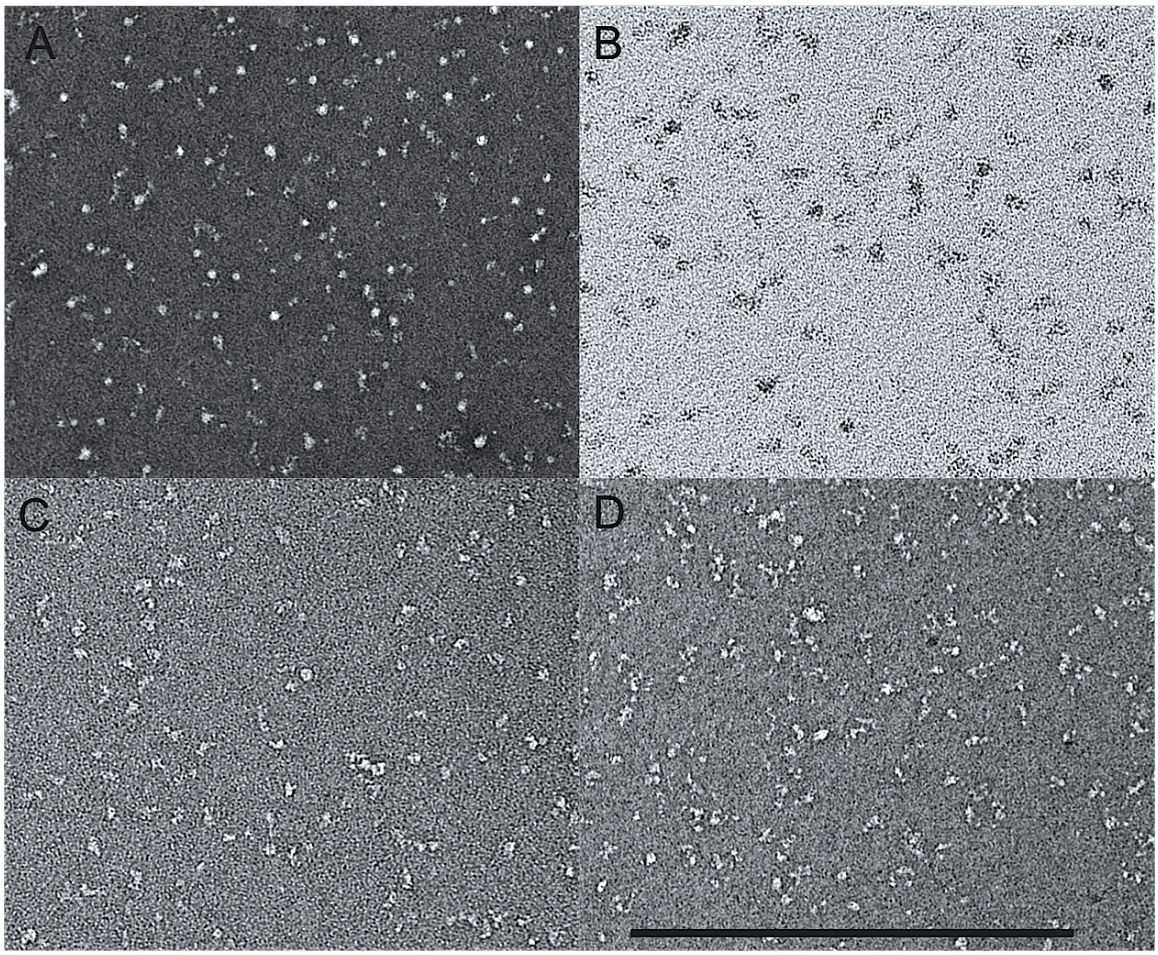


Figure 24: Negative staining of purified EGFRΔ998 dimers. A) Phosphotungstic acid staining (PTA). B) Uranyl acetate pH 5 staining. C) Nanovan pH 7 staining. D) Uranyl formate pH 4 staining. High tension 80 kV, magnification 71000x. Scale bar is 500 nm.

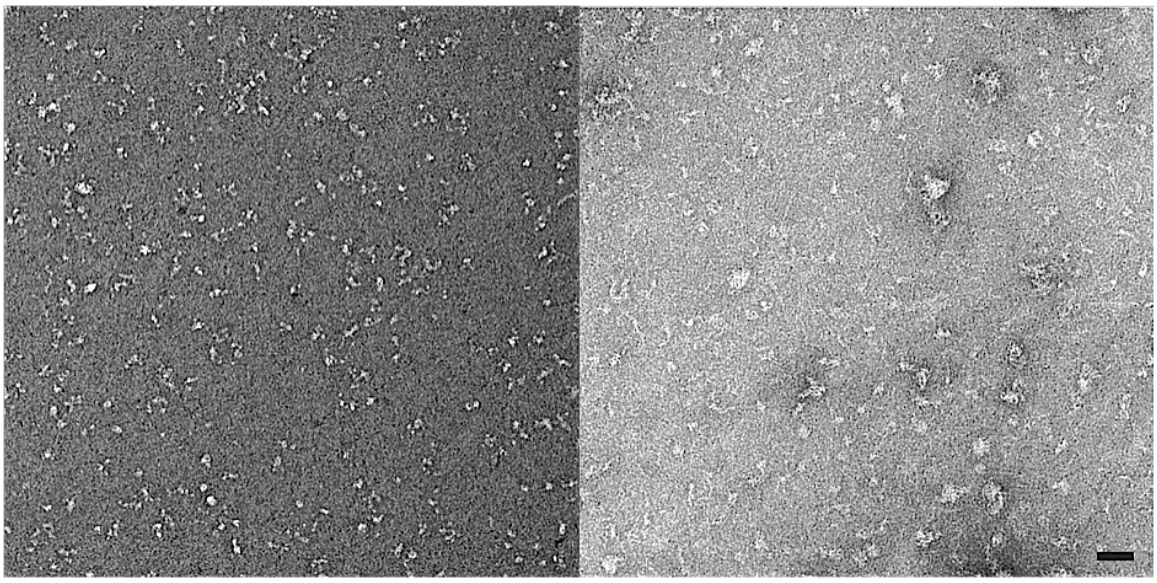


Figure 25: Comparison of negatively stained dimeric EGFRΔ998 and integrin. Left panel, EGFRΔ998 stained with uranyl formate. Right panel, integrin $\alpha V\beta 3$ complexes stained with uranyl acetate (taken from [155]). Scale bar is 100 $\mathring{A}$.

4.3.2 3D model of the EGFRΔ998 E3/K3 dimer reveals ectodomain asymmetry prior to ligand binding

Once a suitable negative staining method was found and image acquisition set up, single particles were manually picked from the micrographs and extracted. A total of about ∼12000 particles were then used in an unbiased, reference-free maximum-likelihood 2D round (ML2D) in Xmipp [156–158], and sorted into 50 different classes. Even though the micrographs showed structural heterogeneity, the initial ML2D analysis yielded some very well defined views that can be seen in Figure 26.

The fifty classes generated in the ML2D round were split into 150 classes that were taken as initial references for subsequent multi-reference analysis (MRA) and multi-statistical analysis (MSA). The results of the fourth refinement step are shown in Figure 27. This analysis showed that the particles are able to generate distinct

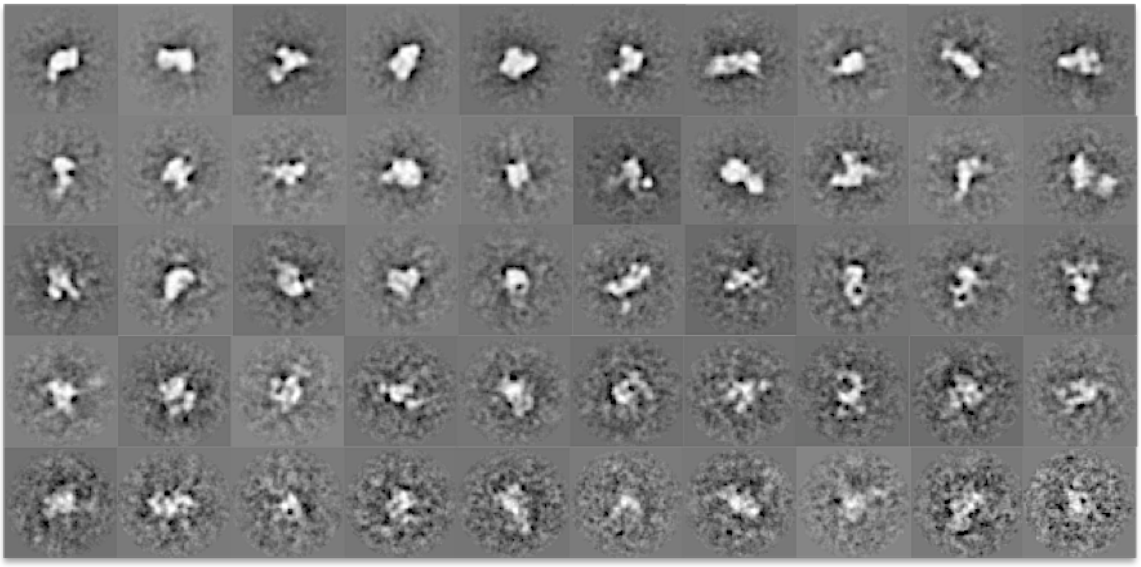


Figure 26: First 50 ML2D classes sorted by number of members per class (from top left to bottom right) of unliganded EGFRΔ998 dimers. $\sim$12000 particles were manually picked from the micrographs, extracted and used in an initial reference-free ML2D round. In spite of the apparent structural heterogeneity, the ML2D yielded 50 different classes, some of them having a very defined shape.

views of what seems to be a highly asymmetric structure, as most of the classes show completely different shapes. Initial attempts for reconstructing a 3D model involved a built-in algorithm, but yielded no results. For this reason, I decided to take another approach. Since X-ray crystallography and NMR have been used to solve the structures of the different domains of the EGFR molecule, I created an initial dimeric model of an inactive EGFRΔ998 molecule by placing all these parts together (ectodomains, PDB code 1NQL, transmembrane domains of a EGFR/ErbB2 heterodimer PBD code 2KS1, juxtamembrane plus kinase domains PDB code 3GOP). The model generated in this way was a symmetric homodimer, as can be seen in Figure 28 A and B. This model was then filtered to 25 $\AA$ resolution and used to generate a set of different Euler angles. These were then visually compared to the classes from the fourth MRA/MSA round

(Figure 27, red squares and Figure 28 C), and the Euler angles for these views assigned (supervised classification). Surprisingly, this method was able to generate an initial 3D reconstruction that looked similar to the modelled chimera but differed in other ways (Figure 29).

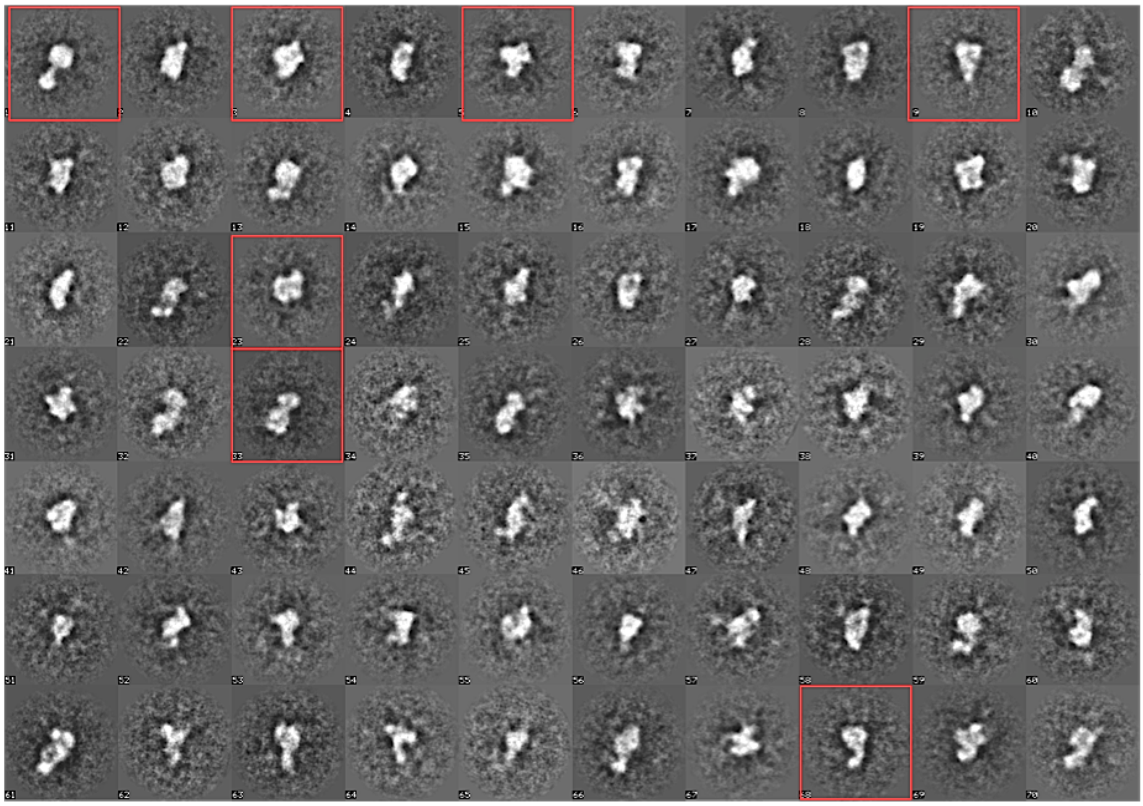


Figure 27: Selected classes for initial 3D reconstruction. The 50 initial classes obtained by free ML2D were used as references for subsequent rounds of MRA/MSA. After the fourth MRA/MSA iteration, 150 classes were generated and sorted by score. Only the top 70 are shown in this figure. Red squares denote the classes that were manually selected for the first 3D reconstruction.

Apart from the slight size change (the PDB structures don't contain all the sugar chains, and the stain depositing around the molecule creates a shell that can distort the size), the most prominent difference is the fact that, in the reconstruction, the

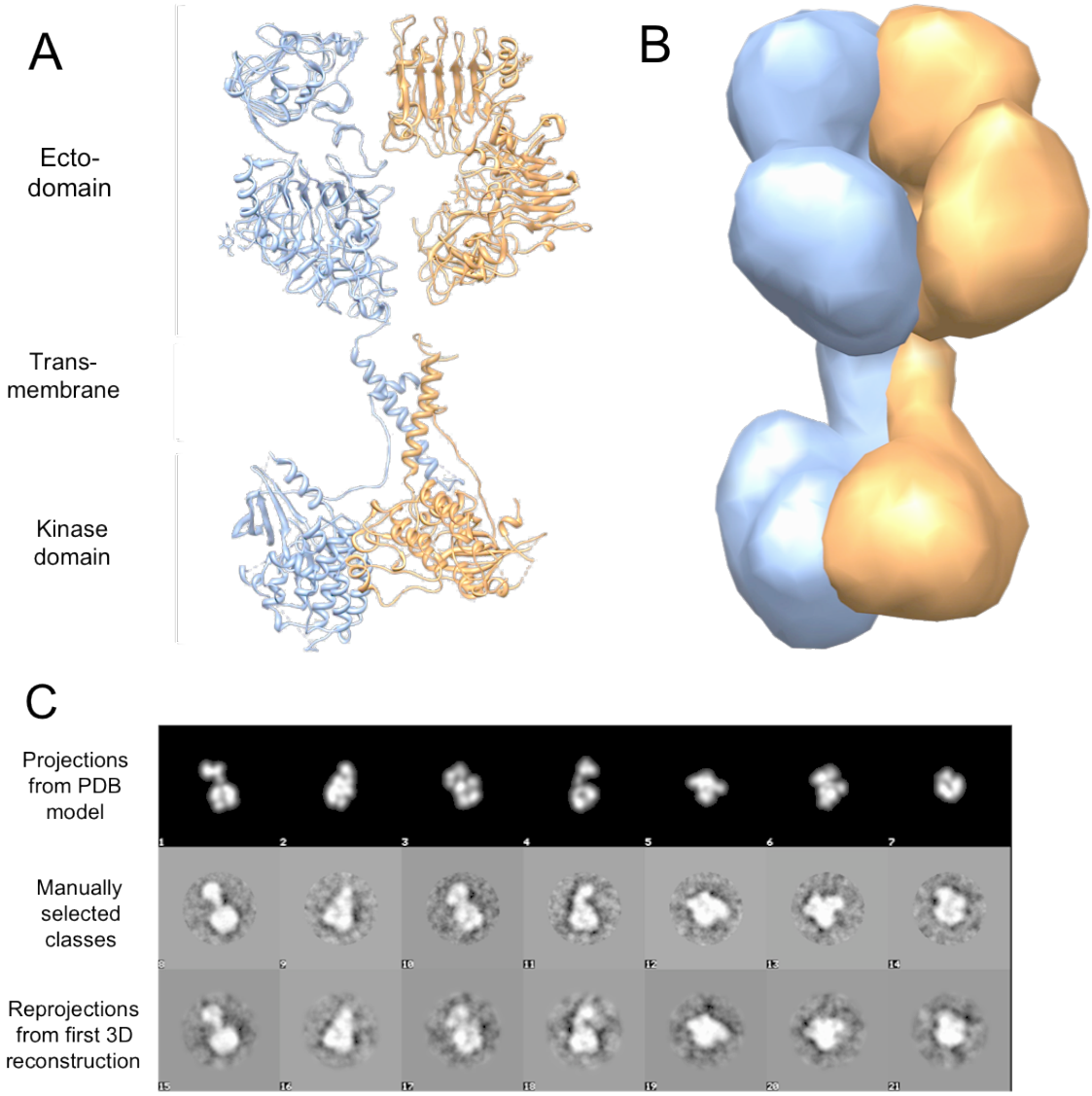


Figure 28: Initial Euler angle assignment and 3D reconstruction. A) Crystal structures for the EGFR tethered ectodomain (PDB code 1NQL), transmembrane (2KS1) and kinase domains (3GOP) were manually docked to resemble an inactive, symmetric homodimer. B) The dimeric structure showed in A) filtered to a resolution of 25$\mathring{A}$. C). The filtered initial 3D reference was used to obtain projections with known Euler angles (top panel) that were manually compared to the classes shown in Figure 27. Those with high visual correlation were selected for Euler angle assignment and initial 3D reconstruction (middle panel). Bottom panel shows reprojections of the initial 3D model after the reconstruction.

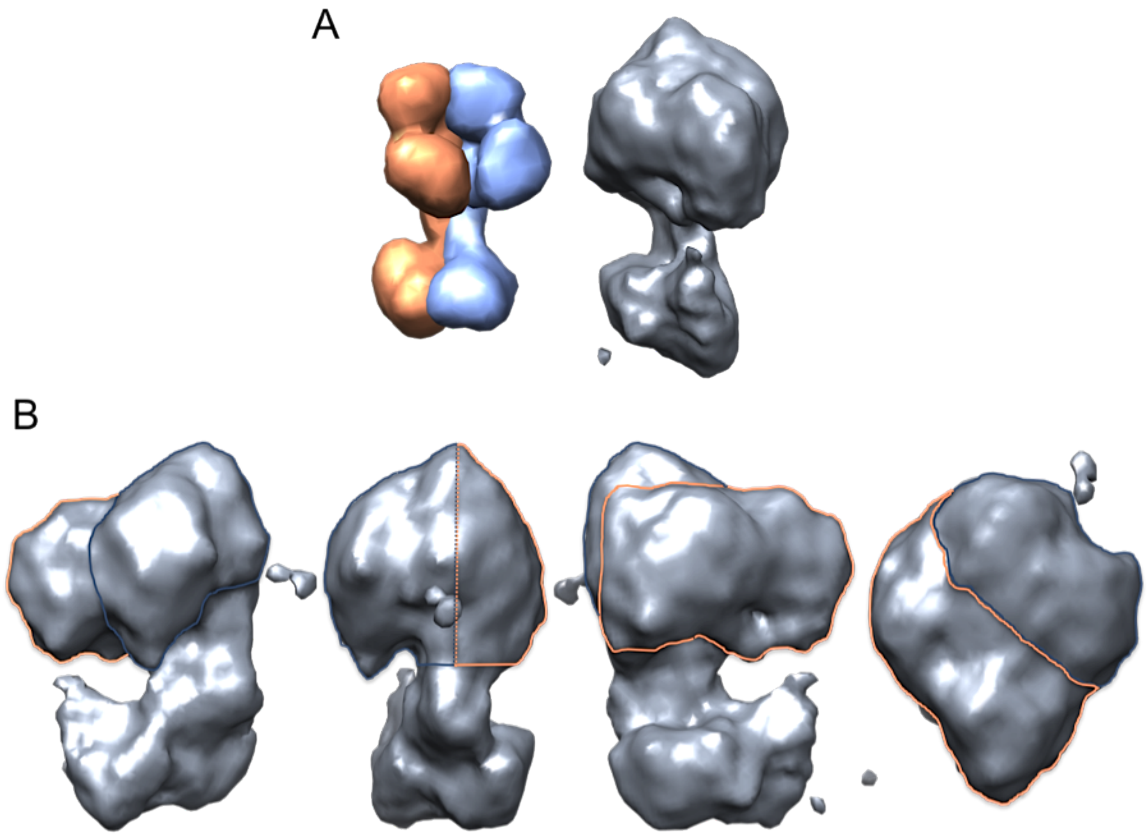


Figure 29: First 3D model of a EGFRΔ998 E3/K3 preformed dimer. The seven selected classes in Figure 28 were used for the first reconstruction. A) Comparison between the initial 3D model (left) and first 3D reconstruction of the EGFRΔ998 E3/ K3 dimer, showing a front view. B) Different views of the reconstruction showing the boundaries between the two ectodomains. From left to right, left, back, right and top views (relative to the back view).

ectodomains seem to interact in an asymmetric fashion. The boundaries between the two ectodomains are highlighted in Figure 29 B. Unfortunately, the resolution was not enough to show the way the kinase domains interact. In any case, and since we had an initial 3D model that could be used as a reference for subsequent analyses, the number of particles was increased (final number ∼24000) in order to refine the 3D reconstruction. The single particles were then matched to the references using an Euler

angle increment of 4°.

The increase in the number of particles and their subsequent analysis improved the resolution of the 3D reconstruction, as can be seen in Figure 30. The final resolution achieved is about 33 $\AA$, and the Euler angle plot shows that many different views were covered (Figure 31). The most striking feature of this model is that, as seen with the initial 3D in Figure 29, the ectodomains interact in a highly asymmetric way. The left ectodomain (with respect to the back view) closely resembles the tethered monomer seen in X-ray crystallography structures, although domains III and IV seem to form a 90° angle with domain II (Figure 30, back view, bottom right panel). In this respect, the junction between domain II and domain III appears to be quite flexible. The right ectodomain, on the other hand, seems to adopt an extended conformation (Figure 30, front and right views, top left and bottom right panels, respectively). This semi-extended configuration of the ectodomain has been seen in other studies [159], although as part of a monomeric receptor. In this publication, the extracellular domain seems to flap back and forth from the tethered to the semi-extended configuration.

The difference between the conformation of the ectodomains can clearly be appreciated in Figure 32, where the final 3D map was segmented and coloured according to density boundaries (built-in function in Chimera). In this way, domain IV of the left ectodomain (blue) wraps around and behind the right ectodomain (orange). The interaction interface between the ectodomains appears to span from domain I to domain IV of the left monomer (blue) and domain II to domain IV of the right monomer. The closest interactions seem to be between domains I-II of the left monomer and domains II-III of the right one, and both domains IV. However, the resolution does not allow us to show a more detailed view of what these interactions are.

With respect to the transmembrane, juxtamembrane and kinase domains, we found the same limitation. The resolution is enough to say that both kinase domains fit into

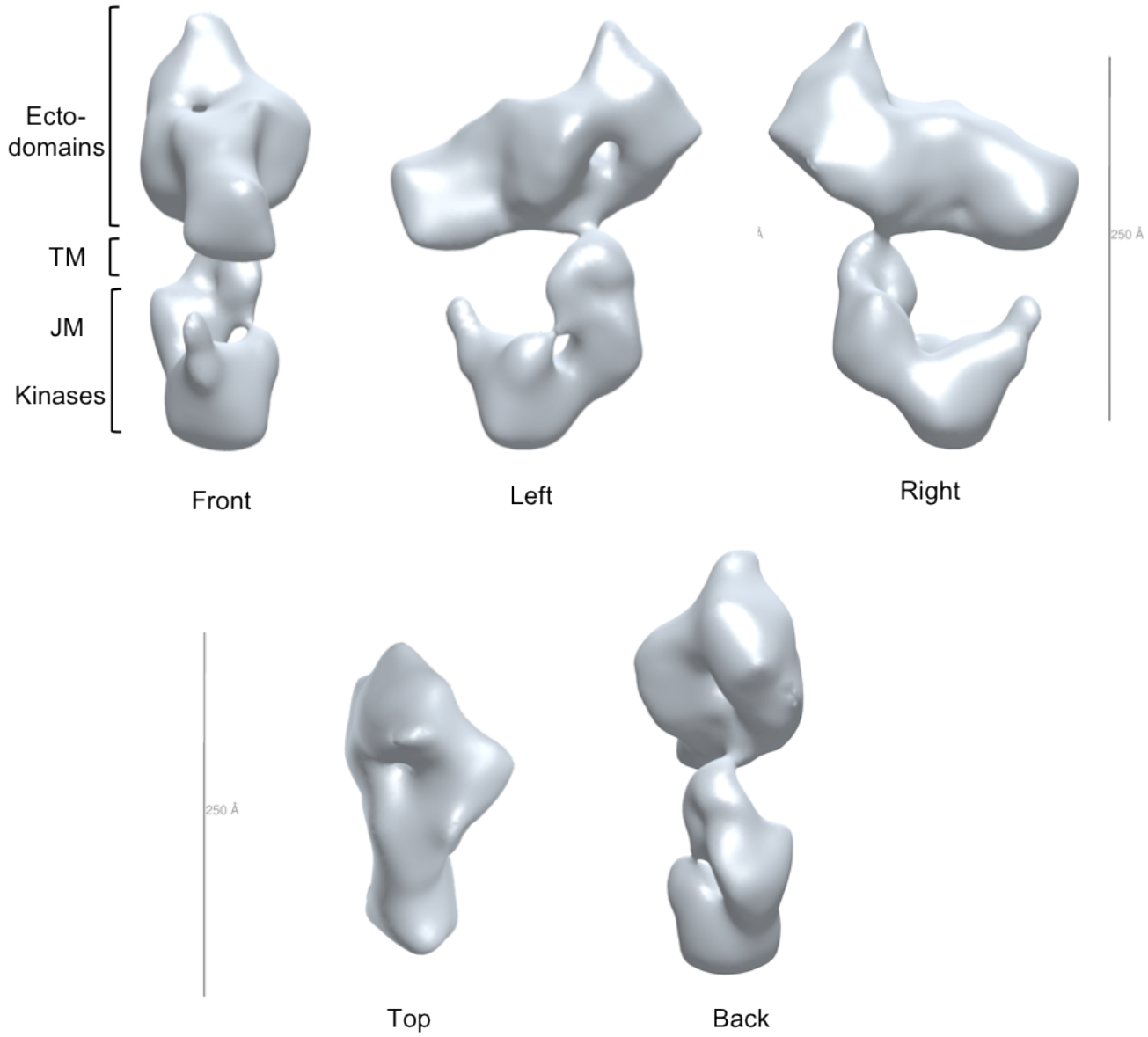


Figure 30: Final 3D reconstruction filtered to 25 $\AA$. After several rounds of automated single-particle projection matching in Imagic, a final 3D structure of the whole molecule was reconstructed and filtered to 25$\AA$. The model presents an unmistakable ectodomain asymmetry, as can be appreciated in the different views. The kinase domains are also clearly defined. Scale bar is 250 $\AA$.

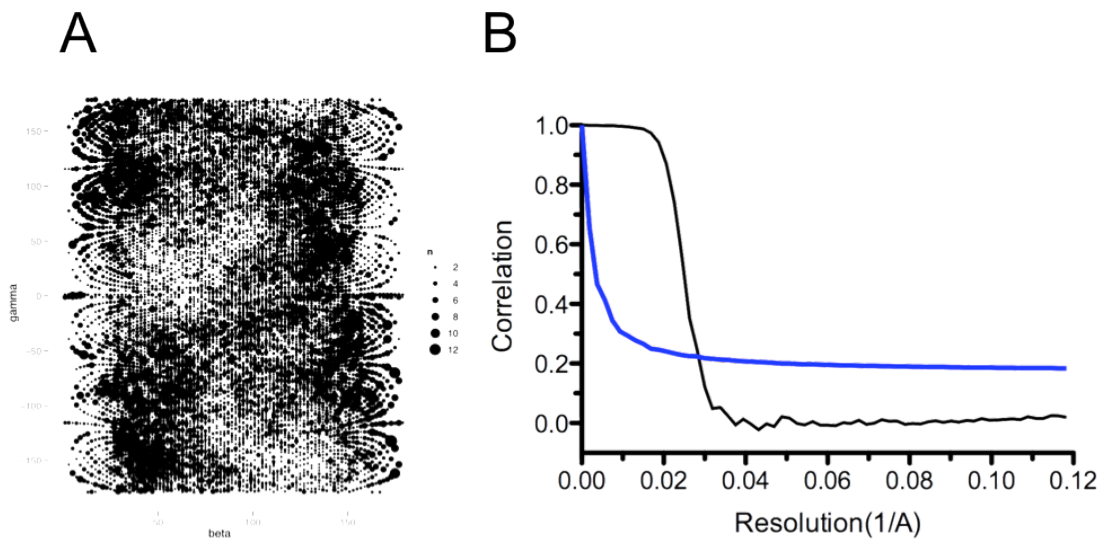


Figure 31: Euler angle coverage and resolution of the final 3D reconstruction. A) Euler angle coverage (gamma versus beta). Thicker dots indicate a larger number of particles per angle pair. B) FSC half bit threshold (blue line) and FSC curve of the last 3D reconstruction. The two curves cross at a 1/resolution of 0.03, or about 33 $\AA$ resolution.

the 3D map, but we cannot tell in which configuration. We think, however, that the spike-like feature protruding from the kinases is indeed the E3/K3 coiled-coil. In any case, a map with higher resolution is needed in order to have a more detailed view of the way the kinase domains interact in an inactive dimer.

For this reason, only the crystal structures of the extracellular domains were used for docking into the electron density map. They fit extremely well into the ectodomain, as can be seen in Figure 33, especially into the left one. However, as it was mentioned before, the junction between domain II and III forms a different angle as the one seen in the crystal structure. This new angle allows the domain IV of the left ectodomain to wrap around and behind the right one, and form an extensive interaction interface. In the case of the right ectodomain, it seems to adopt a semi-extended configuration, as

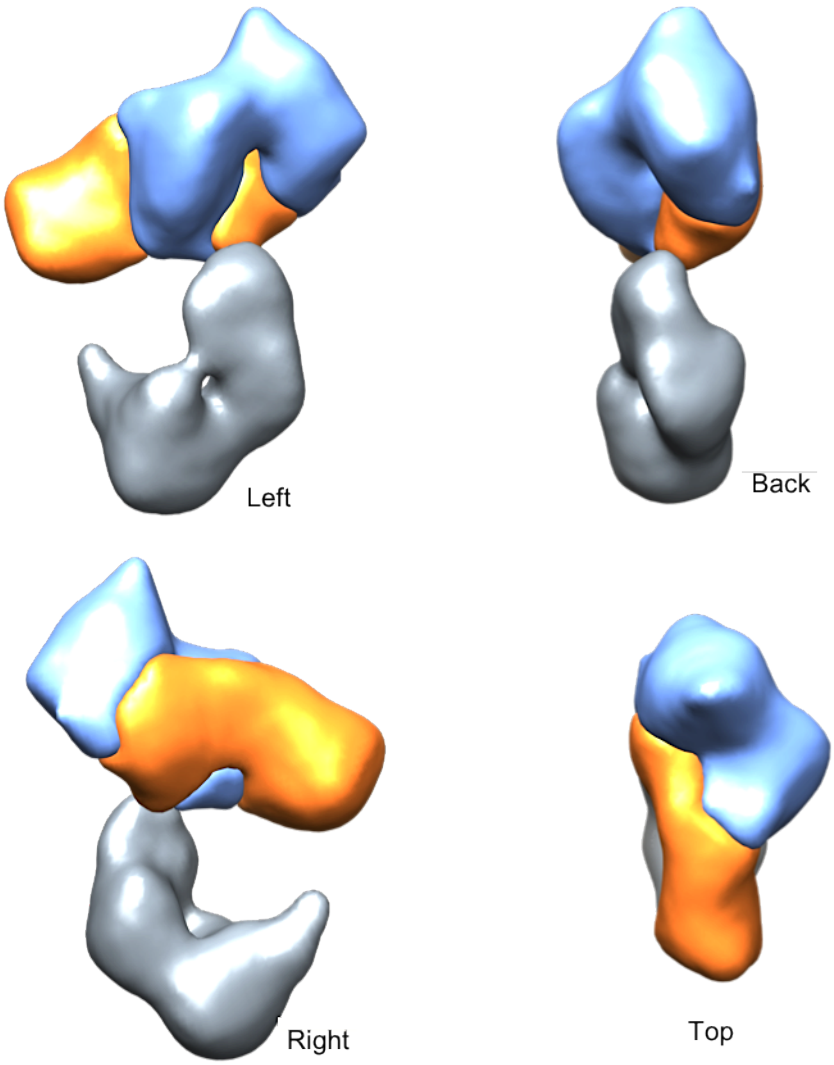


Figure 32: Contoured ectodomains of the final 3D reconstruction. The final 3D map was segmented with Chimera, and those segments corresponding to the two ectodomains were fused and coloured (initial number of segments per ectodomain was 2). Left domain, blue. Right domain, orange.

mentioned above. This can be shown by the docking of the crystal structures in Figure 33.

The crystal docking shows that the left ectodomain is in a tethered conformation, very similar to the one found in the crystal structure 1NQL. However, and as mentioned before, the different domains seem to exist in a more relaxed state. This is especially true for domain IV, which appears to form a different angle with the rest of the molecule. In the crystal structure, domain IV is aligned with domains I, II and III, whereas in the electron density map domain IV forms an angle of about 90º (Figure 33, back view). This domain also seems to closely interact with domain IV of the right monomer. The 3D model also shows that the boundary between domain IV and transmembrane domain of the left monomer (blue) forms a strong angle. In this sense, our structure differs from the proposed models in that the transition between these two domains is not smooth and linear, but rather very flexible. Although the resolution achieved in the final 3D model doesn't allow us to clearly identify the exact contacts between the ectodomains, we can say that they interact in an extensive manner and asymmetrically. The whole ectodomain of the left, constrained monomer seems to be involved, whereas for the right, extended one, only domains II, III and IV appear to interact with the tethered ectodomain.

Since the right domain seems to adopt an extended configuration but it doesn't appear too defined in the final 3D model, we wanted to test if this ectodomain is actually scanning different conformations. This would imply that what we see is actually an average of all these. For this purpose, a maximum-likelihood 3D (ML3D) run was done in Xmipp [160], using as a starting model the final 3D reconstruction shown in Figure 30. Three different ML3D structures were generated, of which two of them are shown in Figure 34.

The results of the ML3D show that the right ectodomain is able to adopt different

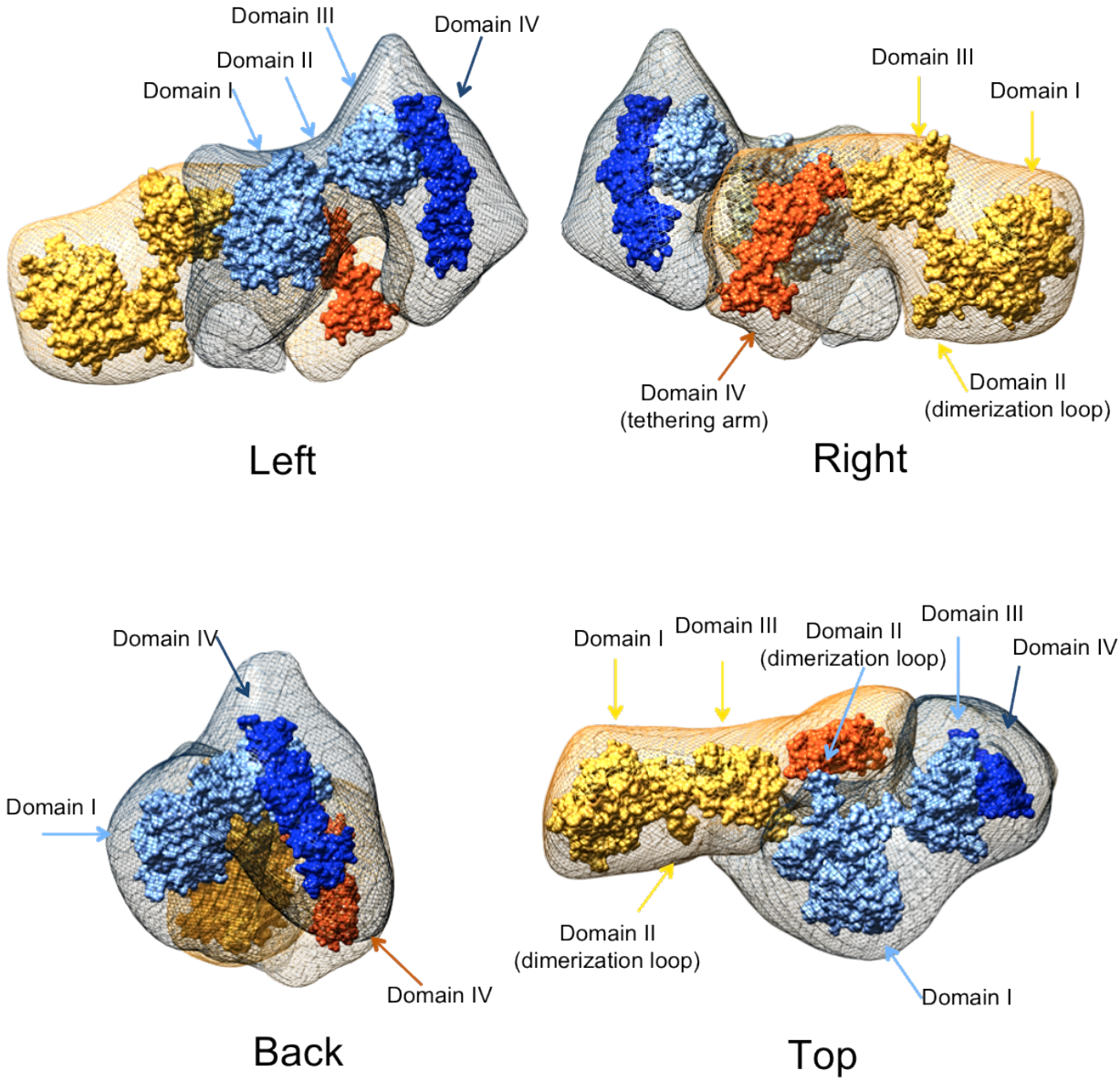


Figure 33: Crystal docking of EGFR ectodomains (1NQL) into the ectodomain electron density maps. The PDB file was divided into two parts: domains I-III and domain IV. The left, tethered ectodomain is shown in shades of blue. The right, extended one, in shades of orange. It is very easy to see that if domain IV pushes domain III forward in the right ectodomain (orange), this will be closer to domain I, creating a high affinity site.

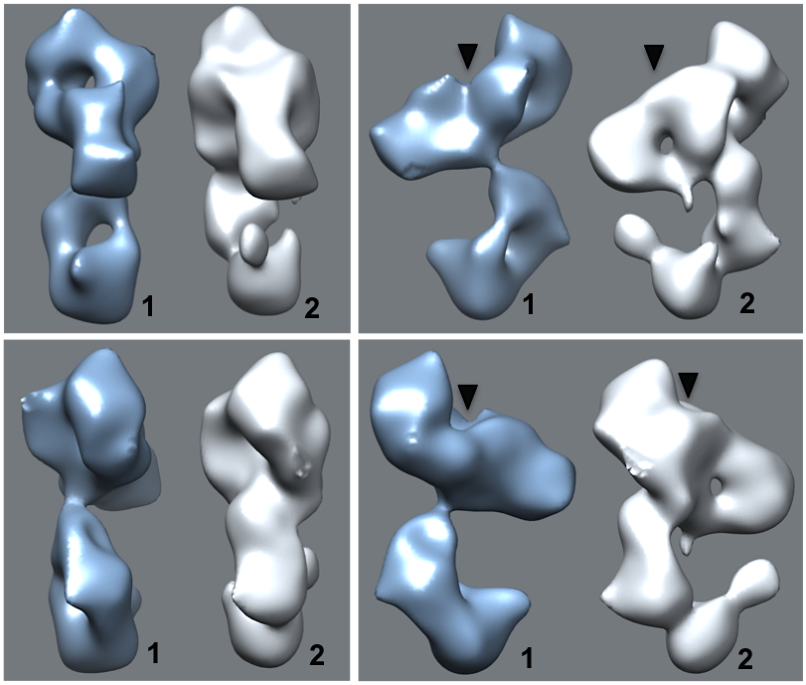


Figure 34: ML3D results of the final 3D model. Panels show front (top left), left (top right), back (bottom left) and right views (bottom left). The area of the right ectodomain that shows the highest difference in density in the two maps is marked with black arrows.

configurations, while the left one only shows minor changes. In this respect, the most prominent change comes from domains II, III and IV of the right ectodomain. In one of the structures, these seem to make extensive contacts with the left ectodomain, while in the other the interface between the two ectodomains seems to be reduced to contacts through domain III of the right monomer and domain I of the left one (Figure 34, black arrows). In the ML3D 1 structure, the right ectodomain seems to be pushed forward, while in structure 2 it interacts more closely with domain IV of the left ectodomain. Unfortunately, and as was the case with the final 3D reconstruction, the resolution is not enough to see changes in the kinase domains. For clarity reasons, the maps were contoured and coloured according to their volume density and are shown in Figure 35.

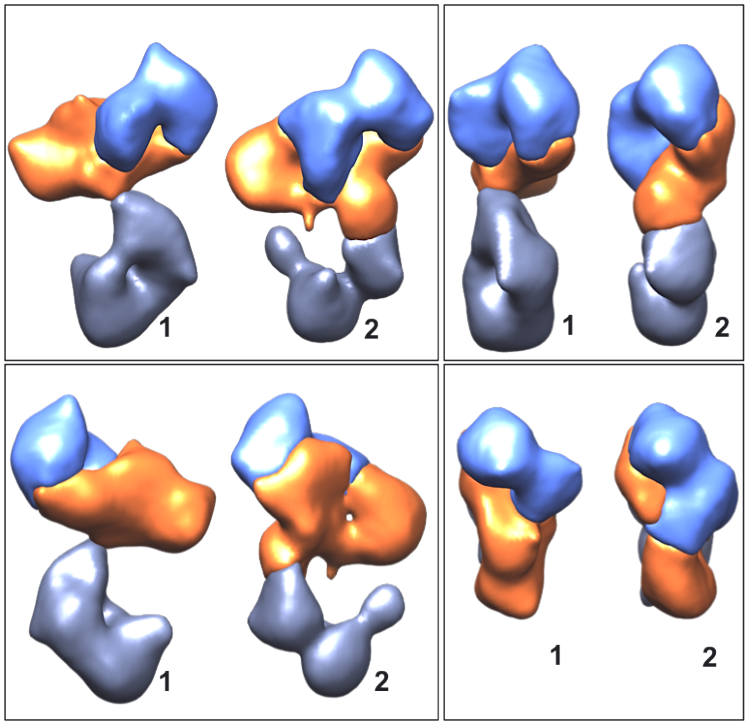


Figure 35: Contoured ectodomains of the two ML3D maps. The most prominent difference between the two structures can be seen in the right ectodomain (orange).

4.3.3 Binding of EGF induces large structural re-arrangements in EGFRΔ998 E3/K3 dimers

From the biochemical analysis, we know that the E3/K3 induced dimer is able to bind ligand (Figure 19). We wanted to test if ligand binding had an effect in the structure of an EGFRΔ998 E3/K3 dimer. For this purpose, after a tandem affinity purification, the samples were treated on ice with 20 μM EGF for 15 minutes and 0.1 mM ATP for 15 minutes. Afterwards, the sample was loaded onto a Superdex200 column. The negative staining and image acquisition were done following the same protocol as for the unliganded EGFR E3/K3 dimers. About 12000 particles were manually picked from the micrographs and an initial, reference-free ML2D performed in Xmipp [156–158]. As

with the unliganded dimer, the initial characterization showed several classes with a defined shape, and different from the unliganded dimers, which are shown in Figure 36.

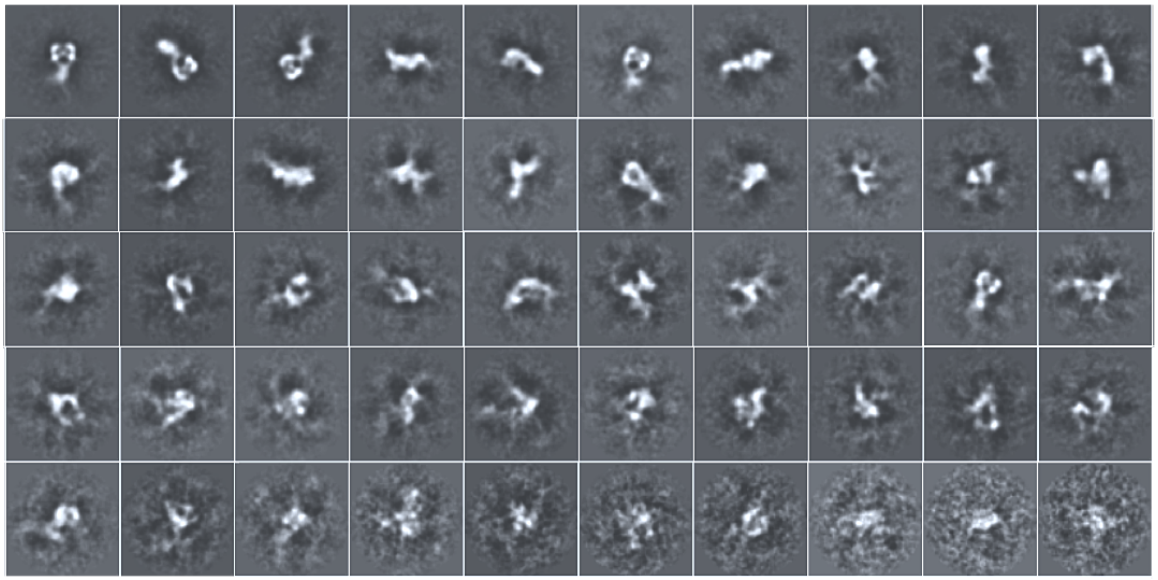


Figure 36: First 50 ML2D classes of liganded EGFRΔ998 E3/K3, sorted by number of members per class (from top left to bottom right). ∼12000 particles were manually picked from the micrographs, extracted and used in an initial reference-free ML2D round.

Some of the classes (especially 1, 2 and 3) show a very distinctive view of the activated EGFR dimer, as has been published before [81, 82]. The structural change from unliganded to liganded dimers is also obvious, as compared to Figure 26. Not only are the unliganded EGFR E3/K3 dimers able to bind ligand, but also this event triggers a major conformational change in the ectodomains as has been described in the literature [81, 82]. After the initial ML2D round, the classes were imported to Imagic and used as references for subsequent rounds of MRA/MSA, following the same protocol as for the unliganded dimers. The results of the third MRA/MSA round showing the top 70 classes (of a total of 200 classes) are presented in Figure 37.

Some of the classes show the distinctive view of a liganded EGFR dimer, as was the

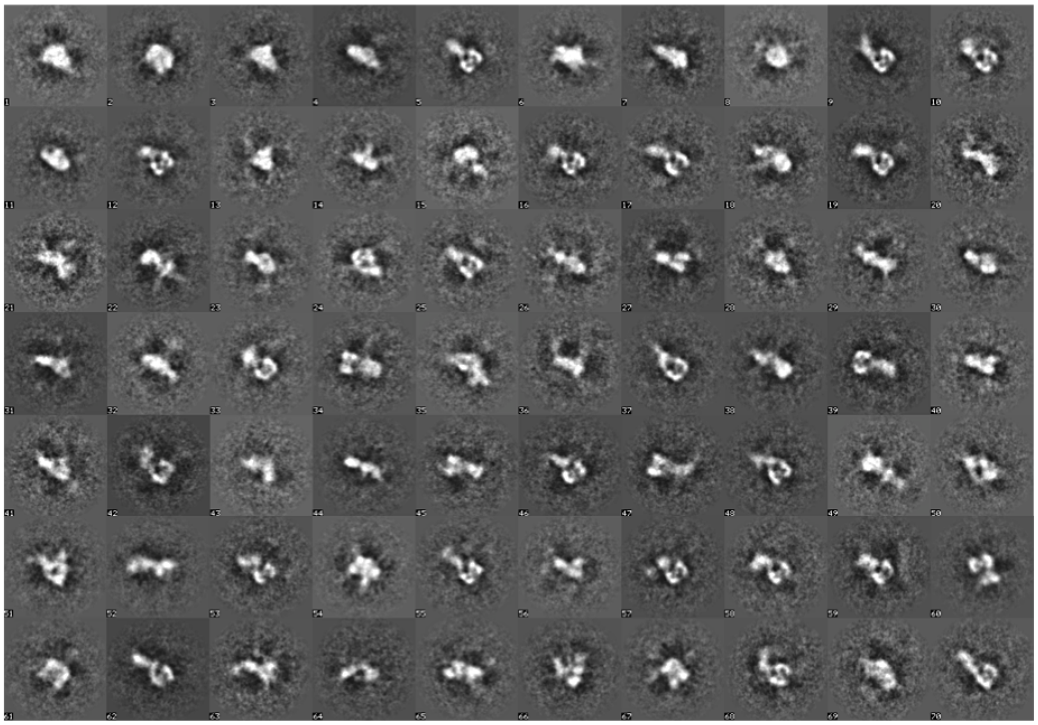


Figure 37: Top 70 classes of liganded EGFR E3/K3 dimers. The 50 initial classes obtained by free ML2D were used as references for subsequent rounds of MRA/MSA. After the third MRA/MSA iteration, 200 classes were generated and sorted by score. Only the top 70 are shown in this figure.

case for the initial ML2D round. However, and even though the ectodomains seem to be very defined, the kinase domains appear as a blur in most of the cases. It also seems that the number of views is somewhat reduced as compared to the unliganded dimers, possibly because of a different interaction with the carbon grid after ligand binding. A 3D reconstruction of a liganded EGFR dimer was not possible, even though several approaches were tried. Contouring the classes and even trying to process the kinases and ectodomains separately yielded no results. This is probably because the kinase, once the dimer is active, becomes extremely flexible and mobile. Hence, the ectodomains appear very well defined, while the kinase domains do not. A different approach is needed in

order to obtain the structure of a liganded dimer, possibly involving reduction in the flexibility of the kinase domains.

4.4 Biochemical characterization of dimeric EGFRΔ998

4.4.1 Unliganded EGFRΔ998 dimers are able to phosphorylate and trigger downstream signalling upon ligand binding

Different experiments were performed to check if the E3/K3-induced EGFRΔ998 homodimers were able to behave as the wild type receptor. This includes plasma membrane localization and activation and signalling upon ligand binding. To check the localization of the preformed dimers, CHO-TetOn cells stably expressing EGFRΔ998 E3/K3 homodimers were treated with EGF-Rhodamine (50 ng/mL) on ice after overnight expression and serum starvation. For the initial localization of the receptors, cells were immediately methanol-fixed after EGF-Rhodamine treatment and washing (Figure 38, left panels). The ability of the receptors to be internalized after ligand binding was tested by placing the EGF-Rhodamine treated cells at 37°C for 15 minutes. After this, cells were methanol-fixed and imaged with a confocal microscope (Figure 38, right panels).

Preformed EGFRΔ998 E3/K3 dimers are correctly localized at the plasma membrane, and become internalized upon ligand binding, as can be clearly seen in Figure 38. Next, I wanted to test if the preformed dimers are also able to retain their phosphorylation activity and trigger signalling after ligand binding. For this experiment, cells expressing EGFRΔ998 E3/K3 dimers were induced and starved overnight, and treated with or without EGF (100 ng/mL) on the next day. After this treatment, cells were

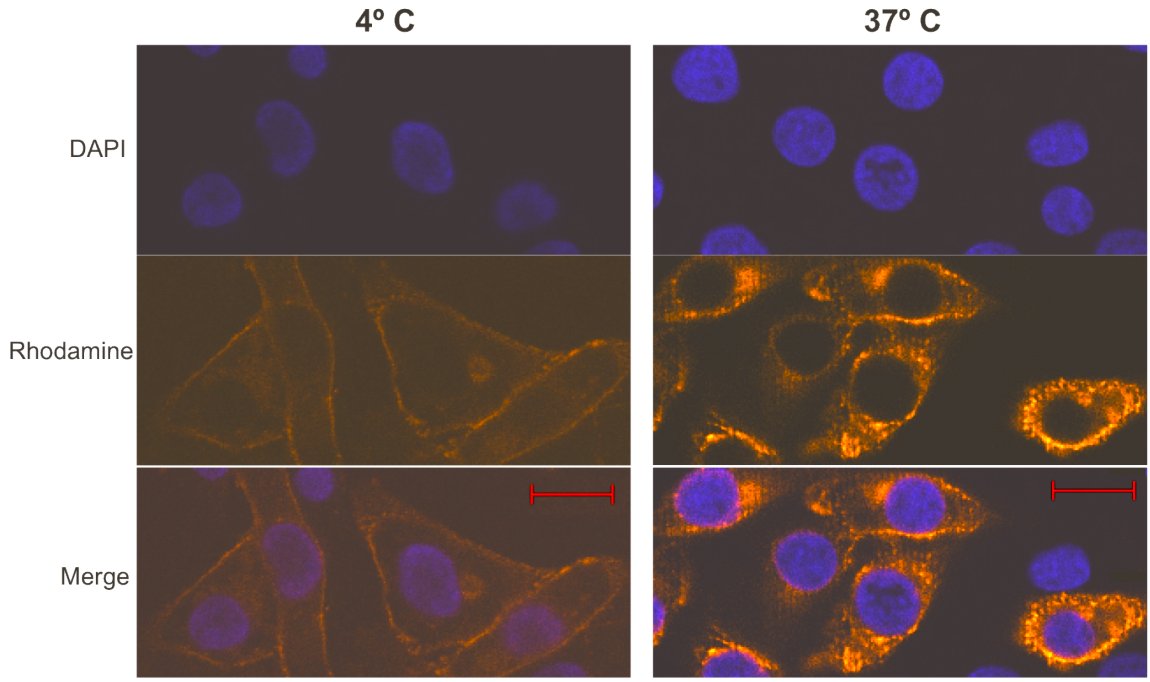


Figure 38: EGFRΔ998 E3/K3 are present at the plasma membrane and internalized upon ligand binding. CHO TetOn cells stably expressing EGFRΔ998 E3/K3 dimers were treated with 50 ng/mL EGF-Rhodamine on ice and either methanol-fixed and imaged directly (left panels) or transferred to 37°C for 15 minutes before fixation and imaging (right panels). Scale bar is 20 μm.

directly lysed in 700 μL of 2.5X Laemli buffer supplemented with 1 mM Na_3VO_4. These lysates were then used to probe the phosphorylation state of the cells. As a positive control, A431 cells were used (A431 cells overexpress EGFR receptors) (Figure 39).

In order to rule out the possibility that the increase in phosphotyrosines comes from the overexpression of one of the monomers, an ECL quantification of the anti-His and anti-Strep signals was made. Since we know that the purified molecules elute as a dimer, the ratio between the anti-His and anti-Strep signals of an affinity-purified and gel-filtrated sample was used as reference of a 1:1 ratio. This was then compared to the signal ratio of the anti-His and anti-Strep of cell lysates, and the results are shown in

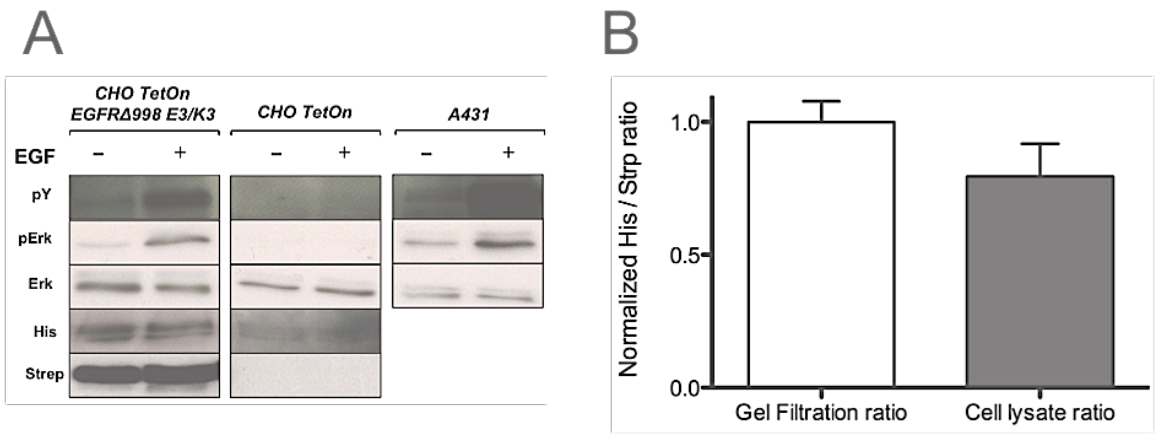


Figure 39: EGFRΔ998 E3/K3 dimers are able to phosphorylate and trigger downstream signalling upon ligand binding. Cells were seeded in 6-well plates and left to reach 80-90% confluence. CHO TetOn cells stably expressing EGFRΔ998 E3/K3 dimers and unmodified CHO TetOn cells were washed twice with PBS and expression induced for 22h in DMEM serum-free medium containing 1 μg/mL doxycycline. A431 cells were starved in DMEM serum-free medium for the same period of time. Cells were then treated with (+ lanes) or without (- lanes) 100 ng/mL EGF in DMEM serum-free plus 0.1% BSA for 8 minutes at 37°C. Wells were immediately washed twice with ice cold PBS and collected in 700 μL 2.5X Laemli buffer supplemented with 1 mM Na_3VO_4, boiled for 5 minutes and analyzed by western blot. B) Western blot His/Strp ratio of gel filtration samples versus total cell lysates. The gel filtration ratio represents a 1:1 interaction. EGFRΔ998 E3 seems to be about 20% more abundant than EGFRΔ998 K3 in cells, which implies that about 80% of the molecules are dimeric. This result demonstrates that the increase in the anti-pY signal upon ligand binding comes from the preformed dimers and not from an overexpression of one of the monomers.

Figure 39 B. About 80% of the molecules *in vivo* are present as dimers, which confirms that the increase in phosphotyrosines comes from the activation of the preformed dimers after ligand binding. The results indicate that preformed EGFRΔ998 E3/K3 dimers are able to be phosphorylated after ligand binding, as well as to trigger downstream signalling, as demonstrated by the anti-pErk lane. However, this truncated form of the receptors is lacking most of the canonical autophosphorylation sites of EGFR, which are mainly located at the C-terminal tail. Other phospho-tyrosine sites, however, are thought to be phosphorylated by intracellular proteins upon ligand binding [33–36]. According to the literature, the truncated form of EGFR receptors used in this study contains one tyrosine autophosphorylation site, Y992. Nevertheless, it has been recently shown that EGFR receptors are also able to autophosphorylate at Y845 ([153] and this study), although the phosphorylation of this site has been traditionally attributed to Src. For this reason, we wanted to test if these sites become phosphorylated after ligand binding (Figure 40).

The analysis of the phosphorylation state of both Y845 and Y992 of EGFRΔ998 E3/K3 in stably transfected CHO-TetOn cells revealed that, in the absence of ligand, these sites have a very high basal phosphorylation. The levels of this basal activity are not changed upon ligand binding (Y845) or only slightly (Y992). The analysis of A431 cells indicates that the antibodies are working properly. However, and even though these sites seem to be highly phosphorylated before ligand binding, downstream signalling and overall phosphotyrosine levels remain responsive to ligand addition. This result could indicate that, either predimerization causes the phosphorylation of Y845 and Y992, or that some cellular proteins in CHO cells are phosphorylating these sites in the absence of ligand. For this reason, CHO-TetOn cells were transiently transfected with full-length monomeric EGFR and the phosphorylation state analyzed in the same way. These results show that the basal phosphorylation of Y845 is CHO-cell dependent, as the

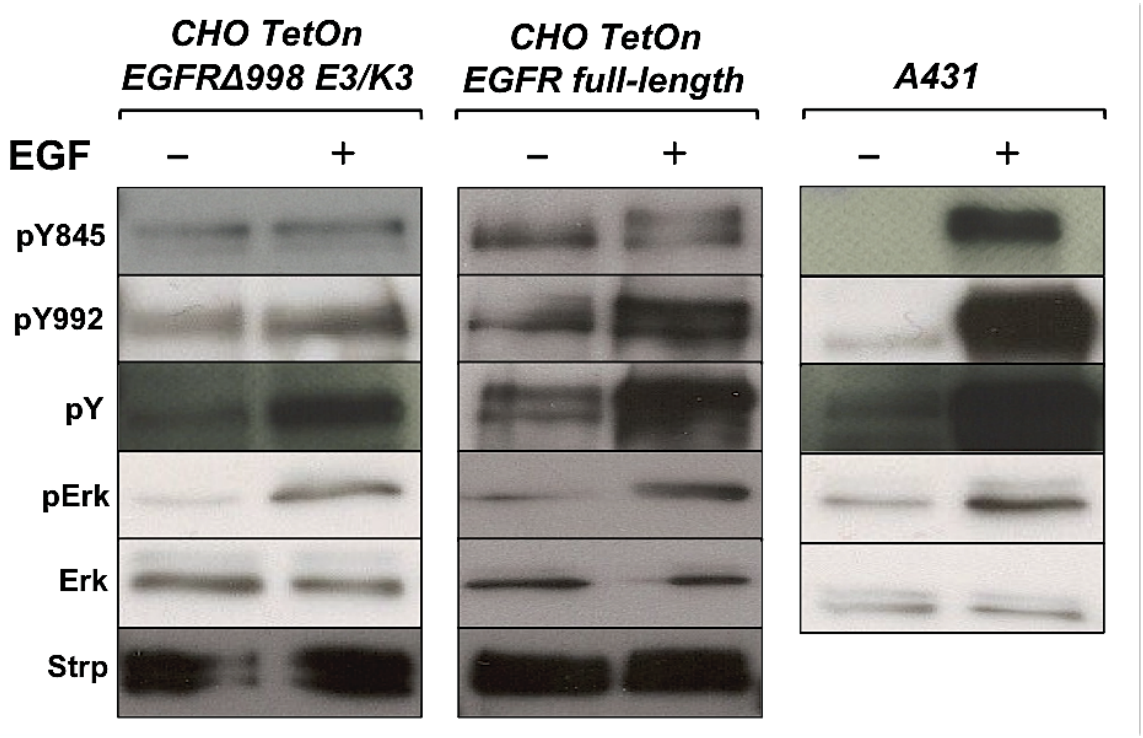


Figure 40: Phosphorylation analysis of EGFRΔ998 E3/K3 and full-length EGFR reveals higher basal signals in CHO TetOn cells. Cells were seeded in 6-well plates and left to reach 80-90% confluence. CHO TetOn cells stably expressing EGFRΔ998 E3/K3 dimers and full-length EGFR transiently transfected CHO TetOn cells were washed twice with PBS and expression induced for 22h in DMEM serum-free medium containing 1 μg/mL doxycycline. A431 cells were starved in DMEM serum-free medium for the same period of time. Cells were then treated with (+ lanes) or without (- lanes) 100 ng/mL EGF in DMEM serum-free plus 0.1% BSA for 8 minutes at 37°C. Wells were immediately washed twice with ice cold PBS and collected in 700 μL 2.5X Laemli buffer supplemented with 1 mM Na_3VO_4, boiled for 5 minutes and analyzed by western blot.

full-length EGFR is highly basally phosphorylated, and only a slight increase in pY845 is seen upon ligand binding. The same is true for Y992 and total phosphotyrosine, although the activation levels are higher than Y845. Downstream signalling doesn't appear to be affected by this basal activity. Since the results were similar to those obtained with the truncated and dimerized version of EGFR, we can conclude that

CHO-TetOn cells, and not predimerization, induce a higher basal phosphorylation of EGFR in our experimental model. The increase in tyrosine phosphorylation seen after ligand binding in EGFRΔ998 E3/K3 dimers is not due to Y845 or Y992. Hence, there are two possibilities left: either EGFR receptors are able of autophosphorylating other, as of yet undescribed tyrosines; cellular proteins recognize and phosphorylate other tyrosines only after ligand binding; or a combination of both. It is known that several other residues are recognized and phosphorylated by cellular proteins [33–36]. The question is, how are these cellular proteins able to phosphorylate these tyrosines only after ligand addition, and not before? In our model system, the EGFR kinase domains are predimerized. The predimerization would bring the kinases to interact in a definite way, which could hide or sterically block tyrosines that will only be phosphorylated upon ligand binding. This event produces a structural change in the extracellular domains, which is then transmitted to the kinase domains. This will also change the way they interact, becoming able to be autophosphorylated and exposing the blocked tyrosines as targets for phosphorylation.

4.4.2 Tyrosine kinase activity of purified unliganded EGFRΔ998 dimers is independent of ligand addition

The phosphorylation capabilities of the EGFRΔ998 E3/K3 dimers was tested *in vitro*. For this experiment, two different approaches were followed: western blot analysis of the phosphotyrosine content of *in vitro* activated EGFRΔ998 E3/K3 dimers, and ability of the dimers to phosphorylate an exogenous synthetic substrate. For the first approach, EGFRΔ998 E3/K3 dimers were purified and loaded onto a Superdex200 column. Fractions containing the dimers were split into two, and treated either with ATP only or with EGF and ATP, and the phosphorylation state was analyzed by western blot (total

phosphotyrosine, pY845 and pY992) (Figure 41 B). For the second approach, a Universal Tyrosine Kinase Assay (Takara, MK410) was used. Gel filtration fractions were split into two, and loaded onto a well of a 96-well plate coated with a synthetic poly-(Glu-Tyr) substrate, with a wide specificity spectrum. Wells were treated with or without 20 μM EGF on ice for 20 minutes, and the reaction was started by adding 20 μM ATP. The detection of the phosphorylated substrate was done following the addition of an anti-phosphotyrosine antibody and colorimetric quantification (ELISA) (Figure 41 C). Surprisingly, the phosphorylation state and the ability of the EGFRΔ998 E3/K3 dimers to phosphorylate an exogenous peptide didn't show any change upon ligand binding, in contrast to what happened *in vivo*. This means that the kinase is active *in vitro* even in the absence of ligand. Since the original samples had no detectable levels of phosphorylation (Figure 41 A) the subsequent tyrosine phosphorylation comes from the kinases themselves. We hypothesised that the kinases are active even in the absence of ligand because of the solubilisation from the plasma membrane. The anchoring to the plasma membrane would prevent the kinases from adopting an active conformation in the absence of ligand. A recent publication [120] strongly supports this hypothesis. Purified EGFR molecules have tyrosine kinase activity in the absence of ligand, just as we see in this study. However, when these EGFR molecules are inserted into a special mixture of lipids, thought to resemble the composition of certain regions of the plasma membrane, the kinase activity is only seen upon ligand binding.

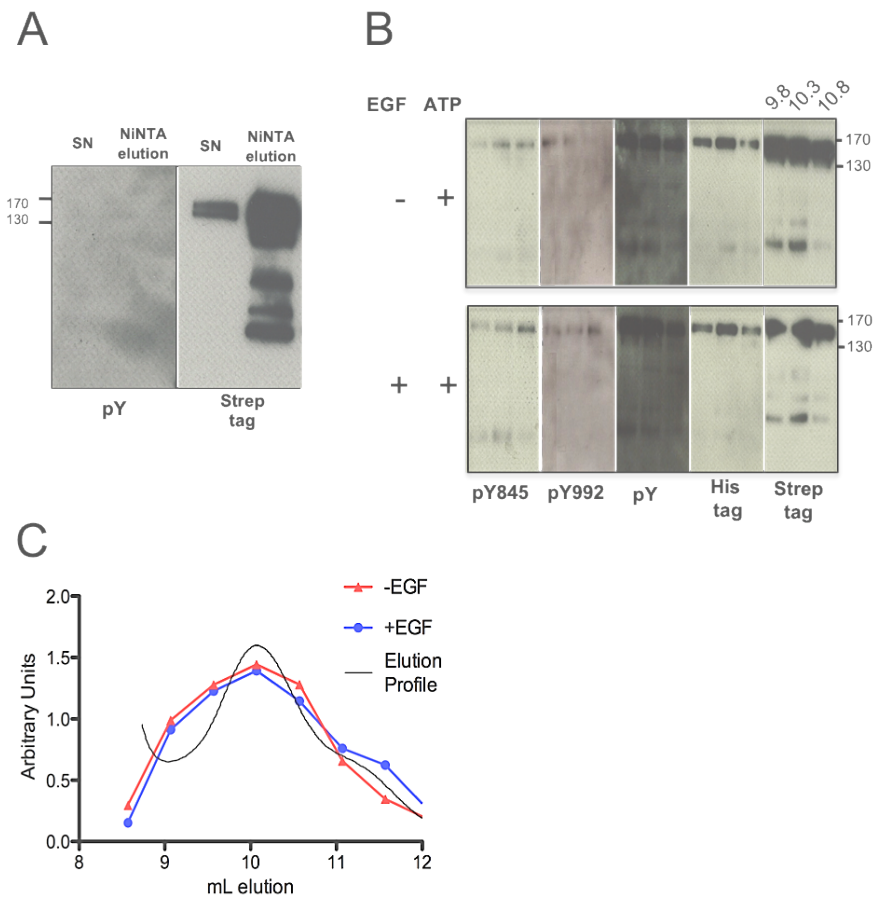


Figure 41: *In vitro* kinase activity of purified EGFRΔ998 dimers. The eluted fractions from a gel filtration were divided into two aliquots, and treated with ATP in the presence or absence of EGF. A) Initial phosphorylation state of the cell lysate (SN) and after NiNTA elution, showing that the proteins had no detectable phosphorylation prior to the addition of ATP. B) Phosphorylation state of the eluted fractions after ATP treatment with or without EGF. Fractions were treated with or without 20 μM EGF for 20 minutes on ice, and the phosphorylation reaction was started by adding ATP at a final concentration of 20 μM. The samples were at room temperature for 15 minutes, and the reaction was stopped by the addition of 1X Laemli buffer and subsequent boiling. Above the last panel, the mL of elution of the gel filtration sample are shown. C) Colorimetric assay showing the tyrosine kinase activity profile of the eluted dimers in the absence (red) or presence (blue) of EGF, treated as in B). The gel filtration profile is shown as a black line. See text for more details on the assay.

4.4.3 EGFRΔ998 dimers bind ligand with high and low affinities

The ability of EGFR receptors to bind ligand with two different affinities *in vivo* has been unexplained for more than two decades [86, 95, 104–109]. When EGFR is purified, the apparent two-site affinity disappears [110, 120]. It has been suggested that preformed EGFR dimers are responsible for the high-affinity binding of EGF [45, 111], or that the receptors achieve the high-affinity conformation by interacting with cellular components [87, 113]. These components would be lost during the purification procedure, and with them, the high-affinity binding. We wanted to test if the predimerization of EGFR has an effect in maintaining the high-affinity binding seen *in vivo*. For this reason, a ligand saturation curve experiment was performed, using europium-labelled EGF (Eu-EGF) as ligand. Gel filtration purified samples were pooled and added to a 96-well NiNTA coated plate. Eu-EGF was added to each well at concentrations ranging from 0.01 nM to 70 nM, per triplicate, and the time-resolved fluorescence (TRF) of the samples was measured after the reaction reached equilibrium. Data were analyzed with GraphPad Prism, and a comparison of fits between one-site and two-site binding was performed. A two-site binding was preferred ($P = 0.0003$). Surprisingly, the purified dimeric EGFRΔ998 E3/K3 dimers still retain both high- and low affinity sites, as evidenced by the biphasic, concave-up Scatchard plot obtained (Figure 42), characteristic of binding assays performed *in vivo*.

To our knowledge, this is the first time that purified EGF receptors are able to reproduce the ligand binding behaviour seen *in vivo*. The high- and low affinity constants were calculated and have values of $K_{d1} = 0.22 \pm 0.15$ nM and $K_{d2} = 30.8 \pm 8.1$ nM. The affinities are somewhat higher than the ones reported in the literature for *in vivo* assays, but this could be due to the fact that the receptors are not in their native,

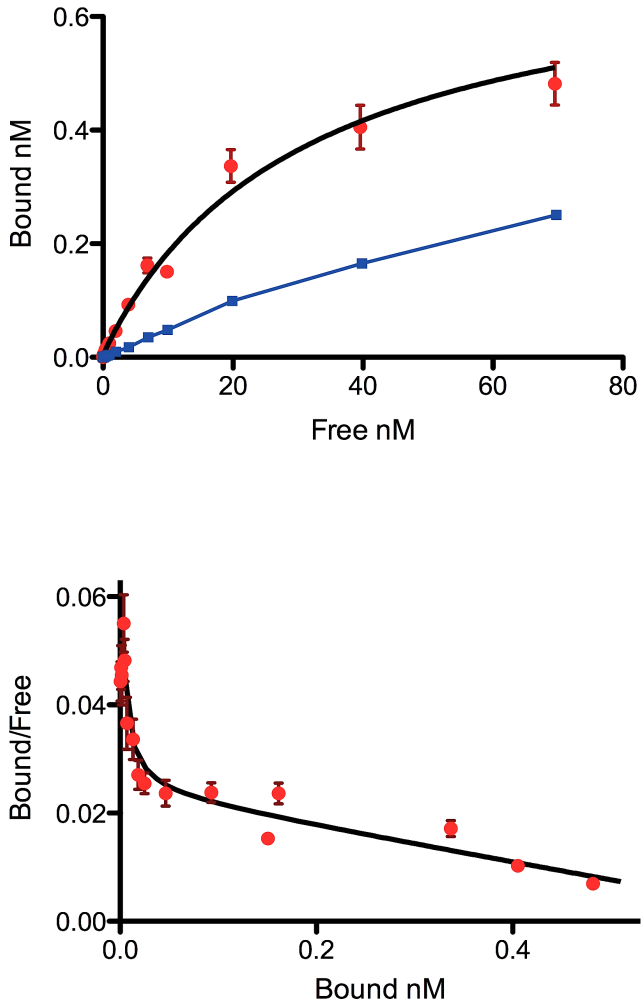


Figure 42: Ligand saturation analysis of purified EGFRΔ998 E3/K3 dimers. Fractions containing EGFRΔ998 unliganded dimers were pooled and bound to black nickel-coated 96 well plates per triplicate. Europium-EGF (Eu-EGF) was added to the wells at concentrations ranging from 0.01 to 70 nM. Non-specific binding was calculated per duplicate by adding 3 μM unlabeled EGF at each concentration of Eu-EGF. Fluorescence levels were measured by TRF (time resolved fluorescence) after incubation for two hours at room temperature. A) Experimental data (red dots) were fitted to a two-site binding curve (black). Non-specific binding is shown in blue. Kd1 = 0.22$\pm$0.15 nM. Kd2 = 30.8$\pm$8.2 nM. Percentage of high affinity sites, 0.83$\pm$0.48 B) Scatchard transformation of the experimental data (red dots) and the fitted curve in A) (black line). Bars denote standard error.

membranous environment, or that the induced E3/K3 dimerization is able to mimic the interaction only to some extent. This could also explain the fact that only a small percentage of the molecules (about 1%) seem to retain high affinity binding. In any case, the affinity ratio (about 150) is similar to wild type receptors.

5

Discussion

Ligand-mediated signal transduction via EGFR receptors has been extensively studied in the last decades. Much is known about the regulation, activation and downstream signalling following EGF binding, as well as about the structural features of the separate domains of EGFR. However, we are still lacking a detailed view of how the signal is transmitted from the extracellular medium to the cytoplasm. EGFR is known to undergo extreme structural changes upon ligand binding, and these changes in the ectodomain must somehow be transmitted to the kinase domains. There are many studies showing that a very important part of EGFR signal transduction is mediated by preformed EGFR dimers at the plasma membrane. In this study, I have described how this dimer associates, exhibiting ectodomain asymmetry and giving a possible basis for the high and low affinity sites observed in these receptors for more than two decades. To our knowledge, this structure represents the first three-dimensional reconstruction of a near full-length receptor tyrosine kinase. This was possible by the introduction of a new purification tool, a heterodimeric coiled-coil, which allowed the preformed dimers to resist the purification procedure.

5.1 The E3/K3 coiled coil as a universal dimerization and purification tool and its applications for signal transduction studies

The E3/K3 coiled-coil interaction has been described extensively in the past, but its components were always free peptides [141, 146, 147]. In this study I have showed how a synthetic, *de novo* designed heterodimeric coiled-coil can be used as a dimerization and purification tool, with the attachment of an affinity tag at its C-terminus. By using FRET measurements and gel filtration assays, I have illustrated how these two peptides perform when attached to a protein. The general behaviour is very similar to what other studies have shown, but there are some main differences. In the case of the E3 peptide alone, its behaviour in solution is virtually monomeric, up to the concentrations tested. However, the K3 peptide shows a very strong dependency on the surrounding residues, as I have demonstrated with the KLAATLE and polyglycine experiments. This is probably due to the fact that, because of sequence similarities, the pET21 linker behaves as a fourth heptad repeat. The number of heptad repeats has been shown to be important for coiled-coil design [141]. In any case, and even when followed by a polyglycine arm, the K3 peptide has an intrinsic concentration-dependent tendency towards homodimerization. In terms of specificity, I have shown that the E3/K3 interaction performs impeccably. Virtually all E3-tagged molecules are recruited by the K3 peptide, and no other natural protein, from prokaryotic or eukaryotic origin, binds any of the pair members. It also resists the most common reagents used in protein purification. However, the pH dependency when attached to a protein seems to differ from the behaviour observed in the free peptides [147]. The association appears to be effectively disrupted at pH5, but instead of E3 homotrimers, the proteins seem to behave

as monomers in solution. It is not probable that this effect is due to a specific property of Venus and ECFP. We hypothesise that the absence of homotrimers can be explained by both steric hindrance effects with the much bulkier tagged proteins, and the dependence on the surrounding residues. In any case, all the features mentioned above make of the E3/K3 tag a very powerful tool for those experiments that need a strong association of two target proteins, yet avoiding the use of a bulky dimerization tag. Although other coiled-coils have been used in protein purification [138, 139, 143, 161], we have chosen the E3/K3 pair both because of its small size (21 aminoacids) and because its properties were quite promising in a previous study [141]. Recently, it has also been used as a fluorescent marker to study EGFR internalization [162]. The induced dimerization of EGFR receptors has been tried before, but in different experimental setups. Split-protein systems, such as split luciferase and split Venus have been used in cells to study conformational changes and the association state of EGF receptors [76, 163]. The TEV protease split system has also been used to monitor protein association states [164]. The dimerization domains FRB and FKBP have also been used in EGF receptors with the same purposes [163, 165–167]. However, these systems are not suitable for purification, because they were designed to observe natural interactions between protein subunits, not to maintain them. It is true that another coiled-coil pair, the GCN4, has been used as a tool for the purification of VEGF receptor ectodomains [168] and in EGFR kinase domain studies [63]. Although this protein is used as an archetypical coiled-coil interaction in protein design experiments (for example [145]), it is based on the natural activator GCN4. I wanted to have a completely synthetic pair in order to avoid any interaction with cellular proteins. Besides, the GCN4 interaction is homomeric. In light of the results, the E3/K3 pair can be used for the study of EGF receptor dimers, and could be easily translated to other cellular receptor systems.

5.2 Phosphorylation and activity of EGFR preformed dimers *in vivo* and *in vitro*

I have shown that induced EGFR E3/K3 dimers are able to generate a proper downstream signal only upon ligand binding. The total levels of phosphotyrosine rise upon ligand binding, just as expected. The fact that about 80% of the expressed EGFR receptors exist in a dimeric form shows that the E3/K3 preformed dimers are not activated before ligand binding. This has also been shown in the literature before [165]. Two canonical EGFR autophosphorylation sites, Y992 and Y845, were also assayed *in vivo*. Y845 has recently been shown to be an autophosphorylation site of EGFR [153] as well as in this study, although its phosphorylation is also dependent on Src [35]. No other autophosphorylation sites were analyzed because the constructs lack the C-terminal tail, where the vast majority of autophosphorylation sites are clustered. The analysis of Y845 and Y992, however, showed that the total amount of these two phosphotyrosines remained almost constant after ligand binding. This fact can mean that, either the EGFR preformed dimers are maximally phosphorylated at these two residues before ligand binding, or that only a small portion of dimers are phosphorylated and no further phosphorylation happens after ligand binding. The results *in vitro* show that, starting from a non-phosphorylated kinase, the EGFR receptors are able to autophosphorylate and phosphorylate an exogenous substrate. For this reason, we think that the EGFR receptors *in vivo* are maximally phosphorylated at these two residues before ligand binding. The other possibility would imply that the kinases are not active, which goes against the observations *in vitro*. But why are the kinases further phosphorylated after ligand binding *in vivo*, while Y992 and Y845 are already phosphorylated? We think that, before ligand binding, the kinases in a preformed dimer are interacting in a way that avoids the exposure of certain surfaces to cellular proteins, hence inhibit-

ing the phosphorylation of key tyrosine residues involved in signalling. Upon ligand binding, the structural changes in the extracellular domains are transferred to the kinase domains, which would expose both the already phosphorylated tyrosines and the surfaces that were hidden before ligand binding. This would now allow the docking of cellular proteins onto these sites and the phosphorylation of other tyrosine residues. These results show that phosphorylation must be accompanied by a structural change for signal transduction to occur. On the other hand, the phosphorylation state of Y845 doesn't seem to be very important for EGFR signal transduction [35, 90]. In this sense, EGFR is different to other protein tyrosine kinases, in which the phosphorylation of the tyrosine at the activation loop is essential for their activity. I have also shown that CHO-TetOn cells seem to have a higher background of phosphorylated Y845 and Y992. Since the phosphorylation of these residues prior to ligand binding does not trigger a downstream signal, we can conclude that both a structural change and phosphorylation are needed in order to activate a downstream pathway. It is also important to note that basal phosphorylation of Y845 has been reported in the literature, and it does not trigger a downstream signal, just as reported in this work [169].

5.3 Ectodomain asymmetry of preformed EGFR dimers and curvilinear Scatchard plots

Curvilinear Scatchard plots of ligand binding to EGFR receptors were seen for the first time almost three decades ago [104, 105]. How a single receptor with identical sequence can generate high and low affinity sites for its ligand has remained unknown for a long time. In this study, I have shown that the ectodomain architecture of preformed EGFR dimers is far away from the symmetry that previous studies and mathematical models have assumed [87, 112]. One of the ectodomains adopts a tethered conformation, albeit

with a different angle between domain IV and the rest of the molecule, while the other ectodomain is in an open conformation and probably has a high flexibility. This fact points to the possibility that the two binding sites are not equivalent, since they are structurally different, and could translate into the high and low affinity states of the EGFR receptors. However, the amount of high-affinity sites is about 1% of the total number of sites. If we took the structural model at its simplest, we would expect 50% of the sites being high-affinity, and 50% of the sites being low affinity. This is very far from the experimental observation, and hence a different mechanism must occur in order to explain the high and low affinity sites. In recent studies [107, 170], the high and low affinity sites have been explained as a result of negative cooperativity in preformed dimers. Insulin receptors, related to EGF receptors, also show negative cooperativity [117, 118]. Another fact that supports the negative cooperativity hypothesis is the recent structural study on *Drosophila* EGF receptor [116]. This study shows that the binding of one Spitz ligand to the predimerized EGF receptor provokes a conformational change in the second site, closing it and reducing its affinity for the ligand. The difference between these studies and the present one is that the preformed, unliganded dimers are assumed or observed to be structurally equivalent. The structure that I have presented in this study reveals an unmistakable ectodomain asymmetry prior to ligand binding. We have not proved if this structure presents any kind of negative cooperativity, but the fact that both ectodomains are interacting closely could point towards this explanation. The negative cooperativity could arise from a structural change generated by the interaction between the two ectodomains prior to ligand binding. Besides, a one-site equation with Hill coefficient can be very nicely fitted to our data. When the fits of a two-site and a one-site with Hill coefficient are compared, a two-site fitting is preferred, but with P=0.02. If we decrease the threshold to P=0.01, then a one-site with Hill coefficient fitting is preferred, with $n_H = 0.88 \pm 0.02$ (Figure 43).

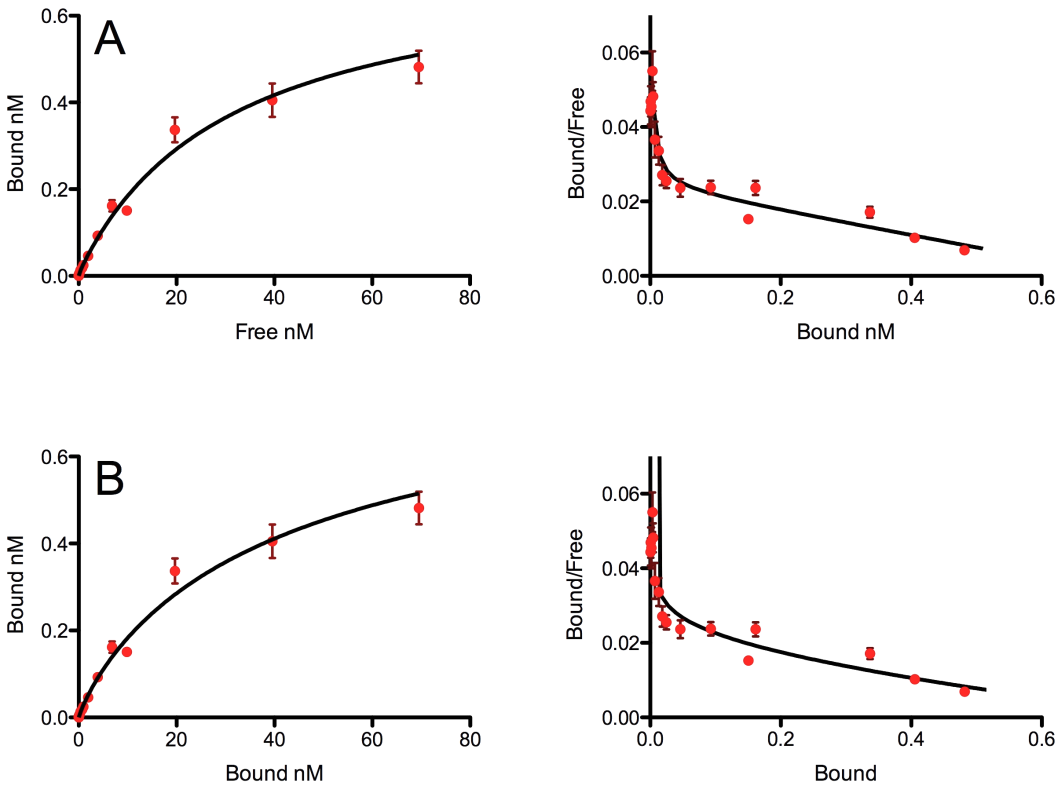


Figure 43: Comparison between two-site binding and one-site binding with Hill coefficient. A) Left panel, two-site binding saturation curve (black line) fitted to the experimental data (red dots). Right panel, Scatchard transformation. B) Left panel, one-site binding saturation curve with Hill coefficient (black line, $n_H = 0.88 \pm 0.02$) fitted to the experimental data (red dots). Right panel, Scatchard transformation.

The fact that only 1% of the sites is able to bind ligand with high affinity could be explained by the flexibility of the right ectodomain. This would scan different conformations, of which only a reduced number corresponds to the high affinity state:

L-L⇌H-L,

where H is high affinity site and L, low affinity site.

In this sense, both ligand binding sites would represent a low affinity state, but one

of them would be in a dynamic equilibrium because of the interaction with another EGFR molecule and probably some intracellular components. In some configurations, one of the low affinity states would change and form a high affinity site. From our model, we think that the right ectodomain (depicted in orange in the figures) is the one that undergoes these changes, since it seems to be more flexible than the left one (in blue). The right ectodomain seems to adopt the dynamic configuration of a monomeric EGF receptor seen in previous studies [159] and monomeric receptors are thought to have high affinity binding *in vivo* [107]. On the other hand, the left ectodomain presents a tethered configuration, more similar to the one seen in crystal structures [24–26]. This fact could mean that the right ectodomain hinders the inherent flexibility of the left one, poising it in a low affinity state. The fact that experiments with isolated ectodomains have failed to show two binding sites highlights the importance of describing the interactions that take place between two EGFR monomers in terms of a full-length molecule (or at least near full-length, like the one used in this study).

However, we cannot rule out the possibility that the high and low affinity states are due to slight structural changes that are too subtle to be seen at 33 $\mathring{A}$ resolution, in which case any of the ectodomains could serve as a high affinity site. Having two structurally different sites suggests also some kind of hierarchy in ligand binding, which could have consequences for downstream signalling, as will be commented on later. Other studies have shown that the intracellular domains, particularly the juxtamembrane, have a role in the maintenance of the high-affinity state [95, 108]. Unfortunately, the resolution achieved doesn't allow us to infer any structural conformations of this part of the molecule. However, the purified EGFR dimers are lacking a membranous environment, which makes their kinases active even in the absence of ligand. The juxtamembrane region is known to be also involved in regulating EGFR kinase activity [63, 96, 97]. It is very possible that the juxtamembrane region of the purified EGFR

receptors is not in its native configuration. The kinase domains are also known to be involved in the establishment of preformed EGF receptor dimers [149], and in our experimental system we know that the kinase domains are active even in the absence of ligand after solubilisation. This implies that the kinase domains may not be in their normal configuration, which could also have an effect on the way the ectodomains are interacting. All these effects could make the high-affinity binding state unstable and hence decrease the observed high affinity sites to about 1%.

Why is one of the ectodomains tethered while the other seems to be in an extended conformation, or even scanning different configurations? What are the molecular mechanisms that allow this to happen? The tethered conformation involves a dimerization loop in domain II that interacts with domain IV. However, in this structure I have shown that domain IV of the tethered ectodomain forms a different angle with the rest of the molecule, and the dimerization loop is too far away to interact with domain IV. Nevertheless, this ectodomain seems to be constrained to this tethered conformation. On the other hand, domain IV of both molecules seems to be interacting closely, especially at the C-terminal stretch. We think that the dimerization loop of the tethered ectodomain is no longer interacting with domain IV of the same molecule, but with a part of the extended ectodomain. Unfortunately, at this resolution we cannot unequivocally identify which region(s) of the extended ectodomain are involved. However, mutation studies have shown that mutation of Y246, a residue present in the dimerization loop, abolishes dimerization of EGF receptors, and most interestingly, the mutated receptor presents only low affinity binding [86, 107]. A recent study has identified a region in the tethering arm (residues 571-593, at the C-terminal portion of domain IV), that when deleted abolishes any kind of cooperativity, and hence the receptors have only one affinity state [170]. These results support our hypothesis of an involvement of the tethering arm in the maintenance of ectodomain asymmetry, and

hence, the high affinity state. From our 3D reconstruction, domain IV seems to form an angle with the transmembrane region, and not just a simple linear continuation of the polypeptide chain. This would imply that the hinge domain between domain IV and transmembrane is very flexible, which would also support our idea that domain IV of the right, extended ectodomain is able to move back and forth, changing the distance between domains I and III and generating a high affinity site. In a recent publication [51] this has been shown to be the case. The 7 aminoacids connecting domain IV with the transmembrane region seem to be extremely flexible. In some molecular dynamics studies [171, 172] the EGFR dimers are not standing proud off the membrane, but lying on it, which would also imply a high flexibility in this region. These MD simulations are backed up by some *in vivo* studies [130]. Again, it seems to be very important that the interaction between two EGFR monomers is described with full-length molecules, as it seems very probable that this back- and forth movement of the right ectodomain is triggered by the tethered, left molecule. That would also explain why we do not see the enantiomer (left ectodomain extended, right ectodomain tethered). There seems to be a very well defined interaction surface between the two ectodomains that involves different sides of the EGFR molecules, hence yielding the asymmetric structure I have presented. In order to prove these hypotheses, the above mentioned mutants should be structurally and biochemically characterized with the E3/K3 system, and an increase in resolution would allow a better definition of the interaction sites between the two ectodomains.

5.4 Biological implications of ectodomain asymmetry in EGFR preformed dimers

The fact that preformed EGF receptors present ectodomain asymmetry has a number of different implications, the most obvious being a hierarchy in ligand binding. This would imply that one of the ectodomains (high-affinity) binds ligand before the other one (low affinity), and hence transmits this change to its kinase domain. Following this line of thinking, there is the intriguing possibility that single-liganded dimers are able to signal. In this case, one of the kinases would phosphorylate its partner, showing a different interface and phosphotyrosine pattern than the double-liganded dimer, or maybe generating a different order of tyrosine phosphorylation. This idea has been commented by [107] and this work shows additional pieces of evidence supporting it. In this way, single-liganded dimers would mediate different responses than double-liganded dimers. The cell could respond differently to low and high concentrations of the same ligand, without the necessity of generating a different receptor. By controlling where the receptors are located, and hence displacing the equilibrium towards preformed dimers of EGFR, the cell could also respond asymmetrically to ligand addition. These two ideas have been studied in two recent publications [111, 173], in which they show that EGF receptors generate different downstream signals depending on the ligand concentration, and that EGF receptors can be located in membrane in an asymmetric manner.

I propose an activation model that implies the association of two unliganded EGFR monomers, mainly via their ectodomains and regulated by their kinase and juxtamembrane regions. The association of the ectodomains is asymmetric, and generates two ligand-binding sites that are not equivalent, thus making possible the existence of both high and low-affinity sites within an EGFR dimer. This would allow cells to respond differently to diverse concentrations of ligand (Figure 44).

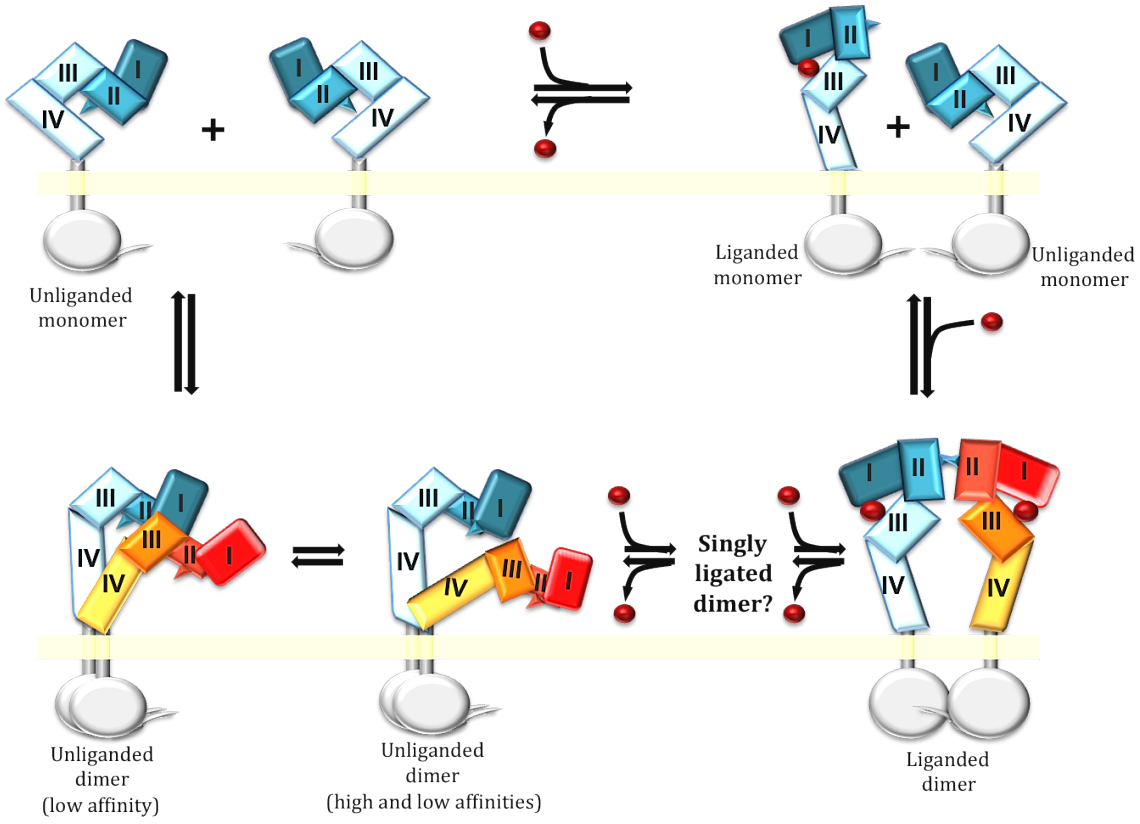


Figure 44: Proposed activation pathway for the EGF receptor. Top panel, classic, sequential activation model for EGF receptors. They exist in a tethered, monomeric state and ligand binds sequentially to generate an active, doubly liganded dimer. Bottom panel, activation pathway involving preformed EGF receptor dimers. In the majority of configurations, the preformed EGF receptor dimer has low affinity sites, but one of the ectodomains is able to scan different conformations, of which one corresponds to the high affinity state. The other ectodomain retains low affinity, and represents the second binding event to generate a doubly liganded dimer.

EGFR is involved in many cellular pathways and its abnormal behaviour can cause a plethora of illnesses, as has been commented on before. It is well true that we have a lot of information regarding the mode of action, activation and regulation of ErbB receptors, but we are lacking a complete picture of these processes. The first step in ErbB receptor activation is ligand binding, a seemingly simple mechanism that has proved to be very difficult to untangle. With this work I have added a new point of view that hopefully will serve as the basis for future work, not only no the fields of basic research, but also applied to the development of new molecules and therapies that could serve as blockers for aberrant EGFR-mediated signalling.

6

Materials and Methods

6.1 Materials

6.1.1 Constructs used in the characterization of the E3/K3 interaction

- **pET21a-Stag-Venus-K3-KLAATLE-His**: bacterial expression of N-terminally S-tagged Venus followed by the K3 peptide. The K3 sequence is followed by the endogenous pET21 linker before the hexahistidine tag (KLAATLE).
- **pET21a-Stag-Venus-K3-pG-His**: bacterial expression of N-terminally S-tagged Venus followed by the K3 peptide. A polyglycine linker (GGSG) was introduced between the K3 and the his tag, the final sequence being GGSGLE.
- **pET21a-Stag-Venus-TEV-K3-pG-His**: bacterial expression of N-terminally S-tagged Venus followed by the K3 peptide. A TEV cleavage site (ENLYFQG) was introduced after the Venus coding sequence and followed by a polyglycine linker (GGSG), the K3 sequence, a polyglycine linker (GGSG) and a hexahistidine tag.
- **pET21a-Stag-Venus-TEV-K3-PPG-His**: variation of the pET21a-Stag-Venus-

TEV-K3-pG-His construct including a diproline motif (PP) after the K3 sequence and before the polyglycine linker (final sequence is PPGSGLE).

- **pET21a-ECFP-E3-pG-Strp**: bacterial expression of ECFP followed by the E3 peptide. A polyglycine linker (GGSG) was introduced between the E3 and the Strep tag II (WSHPQFEK). A stop codon immediately follows the Strep tag.

- **pET21a-ECFP-TEV-E3-pG-Strp**: variation of the pET21a-ECFP-E3-pG-Strp construct in which a TEV cleavage site (ENLYFQG) was introduced after the ECFP coding sequence, followed by a polyglycine linker (GGSG), the E3 sequence, a polyglycine linker (GGSG) and a Strep tag II.

6.1.2 Constructs used for the purification and characterization of the ErbB receptor pairs

- **pBI-1delta-E3/1delta-K3**: co-expression of EGFR-delta998-TEV-K3-pG-His and EGFR-delta998-TEV-E3-pG-Strp from a single, bidirectional promoter in CHO-K1 TetOn cells (Clontech). The parent plasmid was from Addgene (Plasmid 16623: pBI-GFP).

- **pBI-1FL**: expression of full-length EGFR-TEV-Strep from a single promoter in CHO-K1 TetOn cells (Clontech).

- **pBI-1delta-E3/2-delta-K3**: co-expression of EGFR-delta998-TEV-E3-pG-Strp and ErbB2-delta1007-TEV-K3-pG-His from a single, bidirectional promoter in CHO-K1 TetOn cells (Clontech). The polyglycine arm variant includes a 10-residue linker $(GGGSG)_2$ after the TEV sequence.

6.2 Methods

6.2.1 Cloning of the Venus/ECFP E3/K3 constructs

The ORFs for Venus and ECFP were amplified from the plasmids pPWV and pPWC, a kind gift from Leonie Ringrose Group (IMBA, Vienna), and cloned between the NdeI and HindIII sites of a pET-21a vector. In the case of ECFP, a Strep-tag II (WSHPQFEK) preceded by a polyglycine (pG) linker and followed by a stop codon was added just before the vector's his tag sequence. An S-tag (KETAAAKFERQHMDS) coding sequence was added at the 5' end of Venus, in order to increase its molecular weight and differentiate it from ECFP in subsequent analyses. Two complementary oligos coding for K3 and E3 peptides, preceded by a polyglycine linker (G<u>GS</u>G, in which the underlined residues are the translation of a BamHI site) and including the overhangs for Bsp1407I and HindIII, were annealed and cloned into these sites of the pET-21-S-Tag-Venus and pET-21- ECFP, respectively. These vectors were named pET-21-SVenus-K3-KLAATLE-His and pET-21-ECFP-E3-Strp. KLAATLE is the pET-21a sequence between the HindIII site and the His tag. The K3-polyglycine variant was generated as above, but with the polyglycine coding sequence just after the K3 sequence and inserted between the Bsp1407I and XhoI sites of pET-21S-Tag-Venus. The linker between the K3 peptide and the his tag in this case is GGSGLE. The TEV variants were generated by annealing two complementary oligos coding for the TEV cleavage site (ENLYFQG) and cloning them into the Bsp1407I and BamHI sites of pET-21S-Tag-Venus-K3-pG-His and pET- 21-ECFP-E3-Strp, yielding pET-21-SVenus-TEV-K3-pG-his and ECFP-TEV-E3-Strp.

6.2.2 Expression, purification and TEV cleavage of E3/K3–tagged ECFP /Venus

The different constructs were transformed into BL21 Rosetta and grown in LB plates supplemented with ampicillin, single colonies picked and transferred to 1 L flasks in the presence of ampicillin and chloramphenicol (100 μg/mL and 25 μg/mL, respectively). Expression was induced overnight at OD 600 = 0.7 with a final concentration of 1 mM IPTG, at 25^0C. Cells were then harvested at 6000xg, 15 minutes and the pellet resuspended in 70 mL lysis buffer (Tris 20 mM, NaCl 150 mM, imidazole 10 mM, pH 7.5). The suspension was sonicated 3 times, 1 minute and centrifuged at 20.000xg, 25 minutes, 4^0C. The supernatant was then filtered with a 0.25 μ pore size membrane and applied to either a HisTrap 1mL or a StrepTrap 5mL (GE Healthcare) using an ÄKTA HPLC chromatography system. For HisTrap purification, the column was washed with lysis buffer plus 20 mM imidazole and eluted in lysis buffer 300 mM imidazole. For StrepTrap purification, the column was washed with lysis buffer and eluted with lysis buffer plus 2.5 mM D-desthiobiotin. Eluted fractions were pooled and its concentration measured with the Bradford method (20 μL sample in 1 mL Bradford reagent) per triplicate. For the tandem affinity purification, the ECFP-TEV-E3-Strp lysate was treated with 500 μL NiNTA magnetic beads (Qiagen NiNTA Superflow) in order to clear the sample from any proteins whose synthesis may have not finalized after the stop codon and continued through the his tag. Both SVenus-TEV-K3-pG-his and ECFP-TEV-E3-Strp lysates were mixed and a sequential affinity purification (StrepTrap after HisTrap) carried out as described above. The purity of the samples was assayed by SDS-PAGE and Coomassie staining. For the gel filtration assays, purified samples were brought to the desired concentration in 250 μL lysis buffer and loaded onto a Superdex 200 10/300 GL column (GE Healthcare, bead volume 24 mL) at a flow rate of 0.6 mL/

min, previously equilibrated with 35 mL lysis buffer. Eluted fractions were collected in 250 μL aliquots, tested for fluorescence levels and the concentration of the highest peak measured with a Bradford assay (20 μL sample in 1 mL Bradford reagent). The purity of the samples was assayed by SDS-PAGE and Coomassie staining. For the TEV cleavage, purified, his-tagged TEV (a kind gift from Westermann group, IMP) was added to the purified Venus and ECFP samples, and left overnight at 4^{o} in lysis buffer plus 1 mM DTT. On the next day, samples were loaded onto a NiNTA column and the flowthrough collected.

6.2.3 Gel filtration standard curve

All calculated molecular weights, in both the Venus/ECFP as well as in the EGFR experiments, were extrapolated from a standard curve of a mixture of gel filtration molecular weight markers (BioRad). 100 μL of these mixture were brought to a volume of 250 μL in PBS buffer and loaded onto a Superdex200 column (GE Healtcare) connected to an ÄKTA HPLC system. The markers used and the observed elution volumes are:

- **Thyroglubulin**: 670 kDa, 9.30 mL.
- **Gamma–globulin**: 158 kDa, 12.81 mL.
- **Ovalbumin**: 44 kDa, 15.37 mL.
- **Myoglobin**: 17 kDa, 17.33 mL.

The molecular weight is given by the equation $Log(MW, Da) = -0.2014x + 4.7379$, where x is the observed elution volume of a sample. R^2=0.99829.

6.2.4 Fluorescence measurements, FRET efficiency and distance estimation

A final concentration of 2 μM for each species was used (SVenus-TEV-K3-pG-His only, ECFP-TEV-E3-Strp only, manually mixed equimolar amounts of the TEV constructs and tandem affinity purified heterodimer). The samples were diluted to the desired concentration in a final volume of 200 μL lysis buffer and their fluorescence measured in a Tecan GeniosPro fluorescence reader (manual gain 25, four measurements per well). Samples were assayed per triplicate. In the case of FRET efficiency buffer screening, the samples were dialysed against the desired buffer before mixing (except for the Triton X100 buffer, in which the samples were brought to a final detergent concentration of 1% by mixing them with a 2X concentrated buffer). Venus channel, excitation 505 nm emission 535 nm, FRET channel, excitation 420 nm/emission 535 nm, and ECFP channel, excitation 420 nm/emission 465 nm. The error bars show the standard error (σ/n) of the three measurements. This value was used for the calculation of FRET efficiency and distance estimation errors (error propagation). The FRET efficiency (E) was calculated with the fluorescence reduction of the donor (ECFP) in the presence of the acceptor (Venus), according to the following relationship:

$E = 1 - \frac{R_{\text{mix}}}{R}$ *Equation 1*

[174], where R is the fluorescence ratio of an ECFP-only sample over a Venus-only sample , R_{mix} is the fluorescence ratio of ECFP over Venus when both proteins are mixed, subtracting the expected spill over of each protein in other fluorescence channels when measured alone, as follows:

$$R_{\text{mix}} = \frac{ECFPChannel_{\text{mixed}} - ECFPChannel_{\text{Venus}}}{VenusChannel_{\text{mixed}} - VenusChannel_{\text{ECFP}}}$$

And once the efficiency is known, the distance can be calculated with the following equivalence:

$$d(\text{Å}) = 50 \times \sqrt[6]{\frac{1}{E} - 1} \qquad \textit{Equation 2}$$

50 $\mathring{A}$ is the Förster distance for the pair ECFP/Venus. http://www.microscopyu.com/tutorials/java/fluorescence/fpfret/index.html

6.2.5 Isothermal Titration Calorimetry

A VP–ITC Microcalorimeter (MicroCal, LLC) was used. ECFP-TEV-E3-Strp was brought to a final concentration of 100 μM in the syringe and SVenus-TEV-K3-pG-His to 10 μM in the chamber, both in lysis buffer. 30 injections of 10 μL each, with a delay of 1 minute between injections, were carried out at a temperature of 25^0C. The results were analyzed and plotted using MicroCal Origin software.

6.2.6 Mammalian cell culture

Adherent cells were maintained in complete-DMEM (high glucose) medium from IMBA/ IMP media kitchen, supplemented with 7% FCS (Sigma Aldrich), 2 mM glutamine (Sigma Aldrich) and 1X non-essential aminoacids (Sigma Aldrich), in a humid incubator at 37^0C and 5% CO2. For the maintenance of CHO-K1 TetOn and its derived stable cell lines, a final concentration of 150 μg/mL G418 was added to the medium. For induction of CHO-K1 TetOn, complete-DMEM was exchanged for serum free-DMEM plus 1μg/mL doxycycline, having the same formulation as above but containing no FCS.

6.2.7 Cloning of EGFR and ErbB2 E3/K3 constructs

EGFRΔ998 and ErbB2Δ1007 were amplified from the parental plasmids pCMV-HER1 and pMMTV-HER2, a kind gift from Erwin Wagner's lab, IMP. The plasmid backbone was pBI-GFP (Addgene plasmid number 16623). All forward oligos contained a mammalian ribosome binding site just before the ATG start codon (GCCACC–ATG).

- **Generation of pBI-EGFRΔ998–E3Strp/EGFRΔ998–K3pGStrp**: pBI-GFP was cut with NotI and SalI. The coding sequence for EGFRΔ998 was amplified from pCMV-HER1 including both restriction sites and adding a SpeI site just before the SalI site. This was cloned into pBI-GFP, generating pBI–EGFRΔ998–SpeI–SalI/GFP. This plasmid was cut with NheI and Bsp1407I, removing the GFP sequence from the parental plasmid. The coding sequence for EGFRΔ998 was amplified from pCMV-HER1 including both restriction sites NheI and Bsp1407I and adding an AflII site before the Bsp1407I site. The produced plasmid was pBI–EGFRΔ998–SpeI–SalI/EGFRΔ998–AflII–Bsp1407I. Two complementary oligos coding for TEV-E3-Strp-Stop were annealed, including the overhangs for SpeI and SalI, and cloned into pBI–EGFRΔ998–SpeI–SalI/EGFRΔ998–AflII–Bsp1407I. The generated plasmid was pBI–EGFRΔ998–TEV-E3-StrpI/EGFRΔ998–AflII–Bsp1407I. This was finally digested, and two complementary oligos coding for TEV-K3-pG-His-Stop were annealed, including the overhangs for AflII and Bsp1407I, and cloned into the plasmid above. The final plasmid was pBI–EGFRΔ998–TEV-E3-StrpI/EGFRΔ998–TEV-K3-pG-His.

- **Generation of pBI–EGFRΔ998–TEV-E3-StrpI/ErbB2Δ998–TEV-(polyGly)–K3-pG-His**: The same protocol as above was followed, but exchanging the EGFRΔ998–AflII–Bsp1407I sequence for an ErbB2Δ1007–AflII–Bsp1407I. This sequence was amplified from pMMTV-HER2. For the polyglycine constructs, a sequence coding for $(GGGSG)_2$ was added just before the K3 coding sequence. The final plasmid was then pBI–EGFRΔ998–TEV-E3-StrpI/ErbB2Δ998–TEV-(polyGly)–K3-pG-His.

- **Generation of pBI–EGFR–Full Length–Strp/GFP**: EGFR-Full Length was amplified from pCMV-HER1 including the NotI and SalI restriction sites and a C-terminal Strp tag followed by a stop codon, and cloned into pBI–GFP. The plasmid generated was pBI–EGFR–Full Length–Strp/GFP.

6.2.8 Preliminary purification tests of EGFR constructs

Lipofectamine 2000 (Invitrogen) was used to transfect 80-90% confluent CHO-K1 TetOn cells according to the manufacturer's instructions. 3,5 μg of plasmid (pBI-1deltaE3/ 1deltaK3 or pBI-1FLE3/1FLK3) were mixed with 8,5 μL of Lipofectamine 2000 reagent. This mix was added to CHO-K1 TetOn cells and removed after 5-6 hours incubation at 37^0C. The medium was then exchanged to serum free-DMEM plus 1μg/mL doxycycline. After 18-21 hours of induction, cells were washed 3X with ice-cold PBS and solubilised in Solubilisation Buffer (Tris 20 mM, NaCl 150 mM, Triton X100 0.2% v/v, EDTA 1mM, protease inhibitors (Roche Complete EDTA-free), 0.5 mM beta-mercaptoethanol (2-ME)) for one hour on ice. Lysates were cleared by centrifugation at 20000g, 30 minutes at 4^0C. The supernatant was then applied to a Strep-Tactin spin column (IBA Germany) and the purification was done following the manufacturer's instructions, but supplementing the kit's wash and elution buffers with 0.1% Triton X100. The eluted fraction from the Strep-Tactin spin columns was loaded onto a NiNTA-agarose spin column (Qiagen), washed with wash buffer supplemented with 0.1% Triton X100 and eluted in Tris 20 mM, NaCl 150 mM, Triton X100 0.1%, 300 mM imidazole. All buffers contained protease inhibitors. The order of tandem affinity purification can be reversed in some experiments, and it will be stated so in the text. Samples were analysed by western blot. In the case of EGFR/ErbB2 heterodimer tests, geldanamycin was added to the solubilization buffer at a final concentration of 10μM at 45 minutes after initiating the solubilization, and incubated for further 15 minutes on ice.

6.2.9 Generation of CHO-K1 TetOn stable cell lines

A 10-cm Petri dish with 80-90% confluent CHO-K1 TetOn cells was used. pBI-1deltaE3/1deltaK3 and pTK-Hyg (carrying a hygromycin resistance gene) were transfected as described above (molar ratio 10:1), substituting the serum free-DMEM with complete-DMEM. After 24 hours, cells were trypsinized and split 1:10 in complete-DMEM containing 150 μg/mL G418 and 250 μg/mL hygromycin (Invitrogen). When massive cell death occurred (about 5 days later), the plates were washed with PBS and the medium exchanged to new complete-DMEM supplemented with 150 μg/mL of both G418 and hygromycin. Cells were allowed to grow to confluence and then a serial dilution was performed, in order to achieve about 2-4 cells/mL. 250 μL of this dilution were seeded into 96-well plates and single colonies allowed to grow. These colonies were then transferred to 6-well plates, per duplicate. One of the duplicates was induced in serum free-DMEM with 1μg/mL doxycycline for 18-21 hours, lysed in SDS-PAGE sample buffer and assayed for expression by western blot using anti-hexahistidine (Qiagen) and anti-NWSHPQFEK (Genscript) antibodies. Positive colonies for both antibodies were allowed to grow and frozen in complete-DMEM containing 10% DMSO and no selection drugs. These cell lines were labelled CHOK1-TetOn-1deltaE3/1deltaK3.

6.2.10 Large scale expression and purification of EGFR constructs from CHO-K1 TetOn cells

CHOK1-TetOn-1deltaE3/1deltaK3 were grown in 175 or 500 cm^2 plates to 80-90% confluence in complete-DMEM containing 150μg/mL G418 and 250 μg/mL hygromycin. A typical expression assay consists of about 25-30 x 175 cm^2 plates. Cells were washed 2X with PBS to remove all traces of serum and induced in serum free-DMEM containing no selection drugs and 1μg/mL doxycycline for 18-21 hours. Plates were put on ice,

washed 3X with ice-cold PBS and carefully scraped off the plate in PBS containing 1 mM EDTA and protease inhibitors (Roche Complete, 3 tablets/250 mL), using 5 mL per plate. Cells were pelleted at 800g, 6 minutes, 4^0C. The PBS was removed and the pellet immediately frozen at -80^0C. Pellets from 25-30 x 175 cm^2 plates were not allowed to thaw and immediately solubilised in ice-cold Solubilisation Buffer (Tris 20 mM, NaCl 150 mM, Triton X100 0.2% v/v, EDTA 1mM, protease inhibitors (Roche Complete EDTA-free), 4 tablets per 80 mL and 0.5 mM 2-ME), for 45 minutes on ice. Lysates were cleared at 50000g, 4^0C, 30 minutes, and the supernatant filtered with a 0.25 μ pore size membrane. This supernatant was applied to a StrepTactin-5mL column (GE Healthcare) in an ÄKTA system. The column was washed with 60 mL of Wash Buffer (Tris 20 mM, NaCl 150 mM, Triton X100 0.1%, 2-ME 0.5 mM, protease inhibitors) and eluted in Wash Buffer plus 2.5 mM desthiobiotin (Sigma) and protease inhibitors. The eluted fractions were pooled and applied to a 1mL NiNTA-agarose column (GE Healthcare) in an ÄKTA system, washed with 10 mL Wash Buffer and eluted in Wash Buffer plus 300 mM low-absorbance imidazole plus protease inhibitors. Peak fractions were confirmed with a Bradford assay, pooled and applied onto a Superdex-200 column (GE Healthcare). The gel filtration run was performed in DDM-150 buffer (20 mM Tris, 150 mM NaCl, 0.4 mM dodecylmaltopyranoside, (DDM)).

6.2.11 Fluorescence microscopy of *in vivo* ligand binding

Stable CHO-TetOn cell lines expressing EGFRΔ998 E3/K3 dimers were grown to 80-90% confluence on glass coverslips and expression induced for 22 hours with 1 μg/mL doxycycline in serum-free medium. Cells were then washed 2X with ice-cold PBS and treated with 100 ng/mL EGF-Rhodamine (Invitrogen) in serum-free medium plus 0.1% BSA on ice for 15 minutes. After this, cells were washed twice with ice-cold PBS and either fixed directly with methanol or brought to 37^0 C for 10 minutes to induce receptor

intake before fixation. Mounting medium contained DAPI (Vectashield). Cells were then imaged in a LSM700 confocal microscope (Zeiss) under laser excitation 405 nm (DAPI) and 555 nm (Rhodamine) and emission filters 490 nm and 560 nm, respectively.

6.2.12 Kinase assays of purified EGFR dimers

Western-blot analysis: EGFRΔ998 E3/K3 dimers were purified as in 6.2.10 and loaded onto a Superdex200 column. Fractions containing the dimers were split into two, and treated either with 20 μM ATP only or with 20 μM EGF on ice for 20 minutes and 20 μM ATP in buffer supplemented with 1mM Na_3VO_4, and the phosphorylation state was analyzed by western blot (total phosphotyrosine, pY845, pY992, his- and strep-tags). Colorimetric analysis: Gel filtration fractions were split into two, and loaded onto a well of a 96-well plate coated with a synthetic poly-(Glu-Tyr) substrate, with a wide specificity spectrum (Takara). Wells were treated with or without 20 μM EGF on ice for 20 minutes, and the reaction was started by adding 20 μM ATP in buffer supplemented with 1mM Na_3VO_4. The detection of the phosphorylated substrate was done following the addition of an anti-phosphotyrosine antibody and colorimetric quantification (ELISA).

6.2.13 Activation and signalling of EGFR constructs *in vivo*

Cells were seeded in 6-well plates and left to reach 80-90% confluence. CHO TetOn cells stably expressing EGFRΔ998 E3/K3 dimers and unmodified CHO TetOn cells were washed twice with PBS and expression induced for 22h in DMEM serum-free medium containing 1 μg/mL doxycycline. A431 cells were starved in DMEM serum-free medium for the same period of time. Cells were then treated with or without 100 ng/mL EGF in DMEM serum-free plus 0.1% BSA for 8 minutes at 37^0C. Wells were

immediately washed twice with ice cold PBS and collected in 700 μL 2.5X Laemli buffer supplemented with 1mM Na_3VO_4, boiled for 5 minutes and analyzed by western blot.

6.2.14 Ligand saturation assays of purified EGFR dimers

Fractions containing purified EGFRΔ998 unliganded dimers were pooled and bound to black nickel-coated 96 well plates per triplicate. Europium-EGF (Eu-EGF, PerkinElmer) was added to the wells at concentrations ranging from 0.01 to 70 nM. Non-specific binding was calculated per duplicate by adding 3 μM unlabeled EGF at each concentration of Eu-EGF. Samples were left at room temperature for two hours. After washing and adding enhancement solution (HVD Life Science Vertriebs) according to the manufacturer's instructions, fluorescence levels of bound ligand were measured by TRF (time resolved fluorescence) in a TECAN GeniosPro fluorometer. Excitation wavelength 340 nm, emission wavelength 612 nm, manual gain 70, integration time 700 μs, lag time 175 μs, number of reads 10, multiple reads per well 3x3, square, mirror dichroic 3. A standard curve was calculated in the same way by adding known amounts of Eu-EGF per duplicate to wells containing enhancement solution only. The free ligand concentrations were calculated by subtracting the bound ligand from the added concentrations, plotted and analysed with Graph Pad Prism. The tables and figures below show the actual values obtained in the experiments.

Added nM	Free nM		Specific binding nM		Unspecific binding nM	
	Average	*SE*	*Average*	*SE*	*Average*	*SE*
0.01	9.575E-3	3.464E-05	4.247E-4	3.464E-05	1.990E-05	-
0.02	1.910E-2	7.731E-05	8.959E-4	7.732E-05	1.991E-05	-
0.04	3.826E-2	2.096E-4	1.740E-3	2.101E-4	2.488E-05	1.493E-05
0.07	6.635E-2	3.521E-4	3.653E-3	3.520E-4	8.959E-05	-
0.1	9.540E-2	3.694E-4	4.599E-3	3.696E-4	2.190E-4	9.953E-06
0.2	0.193	9.09465E-4	7.064316E-3	9.26948E-4	5.77354E-4	1.79179E-4
0.4	0.38698112	1.443946E-3	1.301888E-2	1.446423E-3	7.61506E-4	8.46085E-05
0.7	0.681551096	1.820102E-3	1.8448904E-2	1.833225E-3	1.552887E-3	2.18957E-4
1	0.975120423	1.889994E-3	2.4879577E-2	1.890649E-3	2.070492E-3	4.97671E-05
2	1.953760258	4.674579E-3	4.6239742E-2	4.683795E-3	4.75321-3	2.93675E-4
4	3.907073702	7.10491E-3	9.2926298E-2	7.10491E-3	8.978976E-3	-
7	6.838420235	1.3050545E-2	0.161579765	1.3050545E-2	1.765932E-2	-
10	9.849497701	9.16129E-3	0.150502299	9.16129E-3	2.4000478E-2	-
20	19.66300782	2.8683376E-2	0.336992176	2.887428E-2	4.9623126E-2	3.314818E-3
40	39.59463462	3.5228786E-2	0.405365377	3.8543801E-2	8.2581825E-2	1.5638326E-2
70	69.51823598	3.7375979E-2	0.481764021	3.7386395 E-2	0.125379263	8.82458E-4

Table 6.1: Ligand saturation curve data.

Added nM	Fluorescence	Blank corrected
0	0.66667	0
0.0005	6.611	5.944
0.001	12.888	12.222
0.005	63.333	62.667
0.01	119.44	118.773
0.05	528	527.333
0.1	1177.4	1176.733
0.5	5329	5328.333
1	11284	11283.333

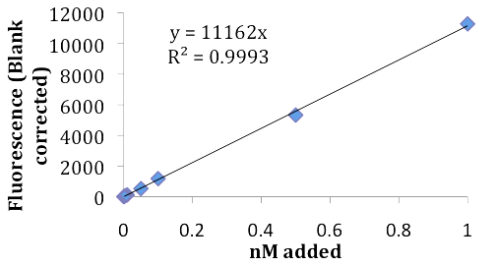


Figure 45: Standard curve for TRF.

6.2.15 Negative staining of purified EGFR dimers

For negative staining, fractions from purified 1deltaE3/1deltaK3 subjected to size exclusion chromatography were used (fraction C4 of a normal run). Carbon-coated grids were glow-discharged for 60 sec and a 5 μL sample was allowed to settle for 30-180 seconds, depending on protein concentration. The drop was blotted from a side and the grid was washed 3 x 5μL water plus a final wash of 5 μL water for about 10 sec. The water droplet was blotted from a side and the grid was quickly washed with 5 μL

of stain (Nanovan; 0.75% uranyl formate; or 0.5% uranyl formate, depending on the experiment). Finally, a 5μL droplet of stain was added to the grid for 30 seconds and excess stain removed by blotting from a side. Grids were allowed to air-dry and either used immediately or stored in a dehumidified vacuum chamber.

6.2.16 Image acquisition and processing

Negative-stain EM data was collected in an FEI Tecnai Polara at 300kV on a Gatan Ultrascan 4000 UHS CCD (4k x 4k, pixel size of 15 micron) using the Leginon software package for automated acquisition [175]. The final pixel size was 2.08 $\mathring{A}$ (71,949-fold magnification), with the defocus ranging from -1.5μm to -4.5μm. The electron dose used was between 40 to 60 electrons per $\mathring{A}^2$ [176]. The mean defocus value of the image was estimated by CTFFIND3 and corrected by phase-flipping with IMAGIC-5 (Image Science Software GmbH, Germany). Particles were picked manually by using e2boxer from the EMAN2 package [177]. First, an unbiased multi-reference 2D alignment was performed by XMIPP using a maximum-likelihood target function [157, 158, 178]. The resulting 2D classes (only ignoring classes most likely representing junk) were used as initial references for several cycles of alignment-by-classification in IMAGIC-5 until convergence was reached. No particles were removed from the dataset in the process.

6.2.17 3D model generation and refinement

A rough initial 3D model was generated in UCSF Chimera [179] from available EGFR crystal structures. Selected forward projections of this model were matched to selected 2D classes of the particle by visual inspection, in order to obtain initial Euler angles, in IMAGIC-5. A first generated 3D model was reprojected, and gradually more classes were carefully included after visual inspection to improve the model. After a

few rounds of supervised projection matching of 2D classes, several more rounds of unsupervised projection matching of single particles were performed to refine the model. Datasets were also split in half in order to calculate the Fourier shell correlation (FSC). Furthermore, in order to analyse heterogeneity in the dataset, maximum likelihood 3D-refinement in XMIPP was performed [160].

References

[1] M. A. Lemmon and J. Schlessinger, "Cell signaling by receptor tyrosine kinases," *Cell*, vol. 141, no. 7, pp. 1117–34, 2010. eng. 9, 12

[2] A. Citri and Y. Yarden, "Egf-erbb signalling: towards the systems level," *Nat Rev Mol Cell Biol*, vol. 7, no. 7, pp. 505–16, 2006. eng. 9, 12, 17

[3] G. Manning, D. B. Whyte, R. Martinez, T. Hunter, and S. Sudarsanam, "The protein kinase complement of the human genome," *Science*, vol. 298, no. 5600, pp. 1912–34, 2002. eng. 9

[4] J. Schlessinger, "Common and distinct elements in cellular signaling via egf and fgf receptors," *Science*, vol. 306, no. 5701, pp. 1506–7, 2004. eng. 9

[5] C. W. Ward, M. C. Lawrence, V. A. Streltsov, T. E. Adams, and N. M. McKern, "The insulin and egf receptor structures: new insights into ligand-induced receptor activation," *Trends Biochem Sci*, vol. 32, no. 3, pp. 129–37, 2007. eng. 9, 27, 30

[6] M. A. Lemmon, "Ligand-induced erbb receptor dimerization," *Experimental Cell Research*, vol. 315, no. 4, pp. 638–648, 2009. eng. 10, 11, 12, 16, 21

[7] J.-P. Himanen and D. B. Nikolov, "Eph signaling: a structural view," *Trends Neurosci*, vol. 26, no. 1, pp. 46–51, 2003. eng. 10

[8] C. Wiesmann, G. Fuh, H. W. Christinger, C. Eigenbrot, J. A. Wells, and A. M. de Vos, "Crystal structure at 1.7 a resolution of vegf in complex with domain 2 of the flt-1 receptor," *Cell*, vol. 91, no. 5, pp. 695–704, 1997. eng. 10

[9] A. Ullrich, L. Coussens, J. S. Hayflick, T. J. Dull, A. Gray, A. W. Tam, J. Lee, Y. Yarden, T. A. Libermann, and J. Schlessinger, "Human epidermal growth factor receptor cdna sequence and aberrant expression of the amplified gene in a431 epidermoid carcinoma cells," *Nature*, vol. 309, no. 5967, pp. 418–25, 1984. eng. 10, 12

[10] J. Downward, Y. Yarden, E. Mayes, G. Scrace, N. Totty, P. Stockwell, A. Ullrich, J. Schlessinger, and M. D. Waterfield, "Close similarity of epidermal growth factor receptor and v-erb-b oncogene protein sequences," *Nature*, vol. 307, no. 5951, pp. 521–7, 1984. eng. 10, 17

[11] N. Moghal and P. W. Sternberg, "The epidermal growth factor system in caenorhabditis elegans," *Experimental Cell Research*, vol. 284, no. 1, pp. 150–9, 2003. eng. 10

[12] B. Z. Shilo, "Signaling by the drosophila epidermal growth factor receptor pathway during development," *Experimental Cell Research*, vol. 284, no. 1, pp. 140–9, 2003. eng. 10

[13] C. Qiu, M. K. Tarrant, S. H. Choi, A. Sathyamurthy, R. Bose, S. Banjade, A. Pal, W. G. Bornmann, M. A. Lemmon, P. A. Cole, and D. J. Leahy, "Mechanism of activation and inhibition of the her4/erbb4 kinase," *Structure*, vol. 16, no. 3, pp. 460–7, 2008. eng. 10, 16, 22

[14] A. Ullrich and J. Schlessinger, "Signal transduction by receptors with tyrosine kinase activity," *Cell*, vol. 61, no. 2, pp. 203–12, 1990. eng. 11, 12

[15] B. Linggi and G. Carpenter, "Erbb receptors: new insights on mechanisms and biology," *Trends Cell Biol*, vol. 16, no. 12, pp. 649–56, 2006. eng. 11, 12, 14, 16

[16] H. Fernandes, "Glycosylation-induced conformational modification positively regulates receptor-receptor association. a study with an aberrant epidermal growth factor receptor (egfrviii/delta egfr) expressed in cancer cells," *Journal of Biological Chemistry*, vol. 276, no. 7, pp. 5375–5383, 2001. eng. 11

[17] L. J. Slieker and M. D. Lane, "Post-translational processing of the epidermal growth factor receptor. glycosylation-dependent acquisition of ligand-binding capacity," *J Biol Chem*, vol. 260, no. 2, pp. 687–90, 1985. eng. 11

[18] A. M. Soderquist and G. Carpenter, "Glycosylation of the epidermal growth factor receptor in a-431 cells. the contribution of carbohydrate to receptor function," *J Biol Chem*, vol. 259, no. 20, pp. 12586–94, 1984. eng. 11

[19] C. J. Stroop, W. Weber, G. J. Gerwig, M. Nimtz, J. P. Kamerling, and J. F. Vliegenthart, "Characterization of the carbohydrate chains of the secreted form of the human epidermal growth factor receptor," *Glycobiology*, vol. 10, no. 9, pp. 901–17, 2000. eng. 11

[20] K. B. Whitson, S. R. Whitson, M. L. Red-Brewer, A. J. McCoy, A. A. Vitali, F. Walker, T. G. Johns, A. H. Beth, and J. V. Staros, "Functional effects of glycosylation at asn-579 of the epidermal growth factor receptor," *Biochemistry*, vol. 44, no. 45, pp. 14920–31, 2005. eng. 11

[21] Y. Zhen, R. M. Caprioli, and J. V. Staros, "Characterization of glycosylation sites of the epidermal growth factor receptor," *Biochemistry*, vol. 42, no. 18, pp. 5478–92, 2003. eng. 11

[22] J. T. Garrett, S. Rawale, S. D. Allen, G. Phillips, G. Forni, J. C. Morris, and P. T. P. Kaumaya, "Novel engineered trastuzumab conformational epitopes demonstrate in vitro and in vivo antitumor properties against her-2/neu," *J Immunol*, vol. 178, no. 11, pp. 7120–31, 2007. eng. 12

[23] H. Ogiso, R. Ishitani, O. Nureki, S. Fukai, M. Yamanaka, J.-H. Kim, K. Saito, A. Sakamoto, M. Inoue, M. Shirouzu, and S. Yokoyama, "Crystal structure of the complex of human epidermal growth factor and receptor extracellular domains," *Cell*, vol. 110, no. 6, pp. 775–87, 2002. eng. 12, 14, 19, 27

[24] S. Bouyain, P. A. Longo, S. Li, K. M. Ferguson, and D. J. Leahy, "The extracellular region of erbb4 adopts a tethered conformation in the absence of ligand," *Proc Natl Acad Sci USA*, vol. 102, no. 42, pp. 15024–9, 2005. eng. 12, 14, 19, 92

[25] H.-S. Cho and D. J. Leahy, "Structure of the extracellular region of her3 reveals an interdomain tether," *Science*, vol. 297, no. 5585, pp. 1330–3, 2002. eng. 12, 14, 19, 92

[26] K. M. Ferguson, M. B. Berger, J. M. Mendrola, H. S. Cho, D. J. Leahy, and M. A. Lemmon, "Egf activates its receptor by removing interactions that autoinhibit ectodomain dimerization," *Molecular Cell*, vol. 11, no. 2, pp. 507–17, 2003. eng. 12, 14, 19, 27, 92

[27] W. X. Schulze, L. Deng, and M. Mann, "Phosphotyrosine interactome of the erbb-receptor kinase family," *Mol Syst Biol*, vol. 1, p. 2005.0008, 2005. eng. 12, 13

[28] E. S. Henson and S. B. Gibson, "Surviving cell death through epidermal growth factor (egf) signal transduction pathways: implications for cancer therapy," *Cell Signal*, vol. 18, no. 12, pp. 2089–97, 2006. eng. 12, 17, 18

[29] R. N. Jorissen, F. Walker, N. Pouliot, T. P. J. Garrett, C. W. Ward, and A. W. Burgess, "Epidermal growth factor receptor: mechanisms of activation and signalling," *Experimental Cell Research*, vol. 284, no. 1, pp. 31–53, 2003. eng. 12, 17

[30] A. K. Koutras and T. R. J. Evans, "The epidermal growth factor receptor family in breast cancer," *Onco Targets Ther*, vol. 1, pp. 5–19, 2008. eng. 12, 17

[31] E. Witsch, M. Sela, and Y. Yarden, "Roles for growth factors in cancer progression," *Physiology (Bethesda)*, vol. 25, no. 2, pp. 85–101, 2010. eng. 12, 17, 18

[32] A. W. Burgess, H.-S. Cho, C. Eigenbrot, K. M. Ferguson, T. P. J. Garrett, D. J. Leahy, M. A. Lemmon, M. X. Sliwkowski, C. W. Ward, and S. Yokoyama, "An open-and-shut case? recent insights into the activation of egf/erbb receptors," *Molecular Cell*, vol. 12, no. 3, pp. 541–52, 2003. eng. 12, 14, 16

[33] T. B. Deb, L. Su, L. Wong, E. Bonvini, A. Wells, M. David, and G. R. Johnson, "Epidermal growth factor (egf) receptor kinase-independent signaling by egf," *J Biol Chem*, vol. 276, no. 18, pp. 15554–60, 2001. eng. 12, 76, 78

[34] J. A. Ewald, K. J. Coker, J. O. Price, J. V. Staros, and C. A. Guyer, "Stimulation of mitogenic pathways through kinase-impaired mutants of the epidermal growth factor receptor," *Exp Cell Res*, vol. 268, no. 2, pp. 262–73, 2001. eng. 12, 76, 78

[35] D. A. Tice, J. S. Biscardi, A. L. Nickles, and S. J. Parsons, "Mechanism of biological synergy between cellular src and epidermal growth factor receptor," *Proc Natl Acad Sci USA*, vol. 96, no. 4, pp. 1415–20, 1999. eng. 12, 22, 76, 78, 88, 89

[36] T. Yamauchi, K. Ueki, K. Tobe, H. Tamemoto, N. Sekine, M. Wada, M. Honjo, M. Takahashi, T. Takahashi, H. Hirai, T. Tushima, Y. Akanuma, T. Fujita, I. Komuro, Y. Yazaki, and T. Kadowaki, "Tyrosine phosphorylation of the egf receptor by the kinase jak2 is induced by growth hormone," *Nature*, vol. 390, no. 6655, pp. 91–6, 1997. eng. 12, 76, 78

[37] M. A. Olayioye, R. M. Neve, H. A. Lane, and N. E. Hynes, "The erbb signaling network: receptor heterodimerization in development and cancer," *EMBO J*, vol. 19, no. 13, pp. 3159–67, 2000. eng. 12

[38] Y. Yarden and M. X. Sliwkowski, "Untangling the erbb signalling network," *Nat Rev Mol Cell Biol*, vol. 2, no. 2, pp. 127–37, 2001. eng. 12

[39] G. Carpenter, "Erbb-4: mechanism of action and biology," *Experimental Cell Research*, vol. 284, no. 1, pp. 66–77, 2003. eng. 12

[40] A. Citri, K. B. Skaria, and Y. Yarden, "The deaf and the dumb: the biology of erbb-2 and erbb-3," *Experimental Cell Research*, vol. 284, no. 1, pp. 54–65, 2003. eng. 12, 14, 16, 17

[41] T. Domagala, N. Konstantopoulos, F. Smyth, R. N. Jorissen, L. Fabri, D. Geleick, I. Lax, J. Schlessinger, W. Sawyer, G. J. Howlett, A. W. Burgess, and E. C. Nice, "Stoichiometry, kinetic and binding analysis of the interaction between epidermal growth factor (egf) and the extracellular domain of the egf receptor," *Growth Factors*, vol. 18, no. 1, pp. 11–29, 2000. eng. 12

[42] M. A. Lemmon, Z. Bu, J. E. Ladbury, M. Zhou, D. Pinchasi, I. Lax, D. M. Engelman, and J. Schlessinger, "Two egf molecules contribute additively to stabilization of the egfr dimer," *EMBO J*, vol. 16, no. 2, pp. 281–94, 1997. eng. 12

[43] M. Odaka, D. Kohda, I. Lax, J. Schlessinger, and F. Inagaki, "Ligand-binding enhances the affinity of dimerization of the extracellular domain of the epidermal growth factor receptor," *J Biochem*, vol. 122, no. 1, pp. 116–21, 1997. eng. 12

[44] Y. Sako, S. Minoghchi, and T. Yanagida, "Single-molecule imaging of egfr signalling on the surface of living cells," *Nat Cell Biol*, vol. 2, no. 3, pp. 168–72, 2000. eng. 12, 30, 54

[45] Y. Yarden and J. Schlessinger, "Epidermal growth factor induces rapid, reversible aggregation of the purified epidermal growth factor receptor," *Biochemistry*, vol. 26, no. 5, pp. 1443–51, 1987. eng. 12, 27, 81

[46] C. P. Blobel, "Adams: key components in egfr signalling and development," *Nat Rev Mol Cell Biol*, vol. 6, no. 1, pp. 32–43, 2005. eng. 12

[47] H.-S. Cho, K. Mason, K. X. Ramyar, A. M. Stanley, S. B. Gabelli, D. W. Denney, and D. J. Leahy, "Structure of the extracellular region of her2 alone and in complex with the herceptin fab," *Nature*, vol. 421, no. 6924, pp. 756–60, 2003. eng. 14, 27

[48] T. P. J. Garrett, N. M. McKern, M. Lou, T. C. Elleman, T. E. Adams, G. O. Lovrecz, M. Kofler, R. N. Jorissen, E. C. Nice, A. W. Burgess, and C. W. Ward, "The crystal structure of a truncated erbb2 ectodomain reveals an active conformation, poised to interact with other erbb receptors," *Molecular Cell*, vol. 11, no. 2, pp. 495–505, 2003. eng. 14

[49] S. Li, K. R. Schmitz, P. D. Jeffrey, J. J. W. Wiltzius, P. Kussie, and K. M. Ferguson, "Structural basis for inhibition of the epidermal growth factor receptor by cetuximab," *Cancer Cell*, vol. 7, no. 4, pp. 301–11, 2005. eng. 14

[50] T. P. J. Garrett, N. M. McKern, M. Lou, T. C. Elleman, T. E. Adams, G. O. Lovrecz, H.-J. Zhu, F. Walker, M. J. Frenkel, P. A. Hoyne, R. N. Jorissen, E. C. Nice, A. W. Burgess, and C. W. Ward, "Crystal structure of a truncated epidermal growth factor receptor extracellular domain bound to transforming growth factor alpha," *Cell*, vol. 110, no. 6, pp. 763–73, 2002. eng. 14, 19, 21, 27

[51] C. Lu, L.-Z. Mi, M. J. Grey, J. Zhu, E. Graef, S. Yokoyama, and T. A. Springer, "Structural evidence for loose linkage between ligand binding and kinase activation in the epidermal growth factor receptor," *Mol Cell Biol*, vol. 30, no. 22, pp. 5432–43, 2010. eng. 14, 19, 26, 94

[52] X. Zhang, J. Gureasko, K. Shen, P. A. Cole, and J. Kuriyan, "An allosteric mechanism for activation of the kinase domain of epidermal growth factor receptor," *Cell*, vol. 125, no. 6, pp. 1137–49, 2006. eng. 14, 16, 22, 24, 47

[53] K. M. Ferguson, P. J. Darling, M. J. Mohan, T. L. Macatee, and M. A. Lemmon, "Extracellular domains drive homo- but not hetero-dimerization of erbb receptors," *EMBO J*, vol. 19, no. 17, pp. 4632–43, 2000. eng. 14

[54] D. Graus-Porta, R. R. Beerli, J. M. Daly, and N. E. Hynes, "Erbb-2, the preferred

heterodimerization partner of all erbb receptors, is a mediator of lateral signaling," *EMBO J*, vol. 16, no. 7, pp. 1647–55, 1997. eng. 14

[55] D. Karunagaran, E. Tzahar, R. R. Beerli, X. Chen, D. Graus-Porta, B. J. Ratzkin, R. Seger, N. E. Hynes, and Y. Yarden, "Erbb-2 is a common auxiliary subunit of ndf and egf receptors: implications for breast cancer," *EMBO J*, vol. 15, no. 2, pp. 254–64, 1996. eng. 14

[56] L. N. Klapper, S. Glathe, N. Vaisman, N. E. Hynes, G. C. Andrews, M. Sela, and Y. Yarden, "The erbb-2/her2 oncoprotein of human carcinomas may function solely as a shared coreceptor for multiple stroma-derived growth factors," *Proc Natl Acad Sci USA*, vol. 96, no. 9, pp. 4995–5000, 1999. eng. 14

[57] M. B. Berger, J. M. Mendrola, and M. A. Lemmon, "Erbb3/her3 does not homodimerize upon neuregulin binding at the cell surface," *FEBS Lett*, vol. 569, no. 1-3, pp. 332–6, 2004. eng. 14, 16

[58] J. P. Dawson, Z. Bu, and M. A. Lemmon, "Ligand-induced structural transitions in erbb receptor extracellular domains," *Structure*, vol. 15, no. 8, pp. 942–54, 2007. eng. 14, 21

[59] P. P. Di Fiore, J. H. Pierce, M. H. Kraus, O. Segatto, C. R. King, and S. A. Aaronson, "erbb-2 is a potent oncogene when overexpressed in nih/3t3 cells," *Science*, vol. 237, no. 4811, pp. 178–82, 1987. eng. 16, 17

[60] F. Meric-Bernstam and M.-C. Hung, "Advances in targeting human epidermal growth factor receptor-2 signaling for cancer therapy," *Clin Cancer Res*, vol. 12, no. 21, pp. 6326–30, 2006. eng. 16, 17

[61] D. Alvarado, D. E. Klein, and M. A. Lemmon, "Erbb2 resembles an autoinhibited invertebrate epidermal growth factor receptor," *Nature*, vol. 461, no. 7261, pp. 287–91, 2009. eng. 16

[62] P. M. Guy, J. V. Platko, L. C. Cantley, R. A. Cerione, and K. L. Carraway, "Insect cell-expressed p180erbb3 possesses an impaired tyrosine kinase activity," *Proc Natl Acad Sci USA*, vol. 91, no. 17, pp. 8132–6, 1994. eng. 16

[63] N. Jura, Y. Shan, X. Cao, D. E. Shaw, and J. Kuriyan, "Structural analysis of the catalytically inactive kinase domain of the human egf receptor 3," *Proc Natl Acad Sci USA*, vol. 106, no. 51, pp. 21608–13, 2009. eng. 16, 24, 26, 87, 92

[64] F. Shi, S. E. Telesco, Y. Liu, R. Radhakrishnan, and M. A. Lemmon, "From the cover: Erbb3/her3 intracellular domain is competent to bind atp and catalyze autophosphorylation," *Proceedings of the National Academy of Sciences*, vol. 107, no. 17, pp. 7692–7697, 2010. eng. 16

[65] J. Baulida, M. H. Kraus, M. Alimandi, P. P. Di Fiore, and G. Carpenter, "All erbb receptors other than the epidermal growth factor receptor are endocytosis impaired," *J Biol Chem*, vol. 271, no. 9, pp. 5251–7, 1996. eng. 16

[66] A. Sorkin and L. K. Goh, "Endocytosis and intracellular trafficking of erbbs," *Exp Cell Res*, vol. 314, no. 17, pp. 3093–106, 2008. eng. 16

[67] H. Waterman, I. Alroy, S. Strano, R. Seger, and Y. Yarden, "The c-terminus of the kinase-defective neuregulin receptor erbb-3 confers mitogenic superiority and dictates endocytic routing," *EMBO J*, vol. 18, no. 12, pp. 3348–58, 1999. eng. 16

[68] K. Aertgeerts, R. Skene, J. Yano, B.-C. Sang, H. Zou, G. Snell, A. Jennings, K. Iwamoto, N. Habuka, A. Hirokawa, T. Ishikawa, T. Tanaka, H. Miki, Y. Ohta, and S. Sogabe, "Structural analysis of the mechanism of inhibition and allosteric activation of the kinase domain of her2 protein," *Journal of Biological Chemistry*, vol. 286, no. 21, pp. 18756–18765, 2011. file://localhost/Users/fernandez/Desktop/Papers/2011/Aertgeerts/Journal16, 22

[69] N. Jura, N. F. Endres, K. Engel, S. Deindl, R. Das, M. H. Lamers, D. E. Wemmer, X. Zhang, and J. Kuriyan, "Mechanism for activation of the egf receptor catalytic domain by the juxtamembrane segment," *Cell*, vol. 137, no. 7, pp. 1293–307, 2009. eng. 16, 47

[70] J. Stamos, M. X. Sliwkowski, and C. Eigenbrot, "Structure of the epidermal growth factor receptor kinase domain alone and in complex with a 4-anilinoquinazoline inhibitor," *J Biol Chem*, vol. 277, no. 48, pp. 46265–72, 2002. eng. 16, 22

[71] E. R. Wood, A. T. Truesdale, O. B. McDonald, D. Yuan, A. Hassell, S. H. Dickerson, B. Ellis, C. Pennisi, E. Horne, K. Lackey, K. J. Alligood, D. W. Rusnak, T. M. Gilmer, and L. Shewchuk, "A unique structure for epidermal growth factor receptor bound to gw572016 (lapatinib): relationships among protein conformation, inhibitor off-rate, and receptor activity in tumor cells," *Cancer Research*, vol. 64, no. 18, pp. 6652–9, 2004. eng. 16, 22

[72] C.-H. Yun, T. J. Boggon, Y. Li, M. S. Woo, H. Greulich, M. Meyerson, and M. J. Eck, "Structures of lung cancer-derived egfr mutants and inhibitor complexes: mechanism of activation and insights into differential inhibitor sensitivity," *Cancer Cell*, vol. 11, no. 3, pp. 217–27, 2007. eng. 16

[73] F. Huang, D. Kirkpatrick, X. Jiang, S. Gygi, and A. Sorkin, "Differential regulation of egf receptor internalization and degradation by multiubiquitination within the kinase domain," *Molecular Cell*, vol. 21, no. 6, pp. 737–48, 2006. eng. 16

[74] H. S. Wiley, "Trafficking of the erbb receptors and its influence on signaling," *Experimental Cell Research*, vol. 284, no. 1, pp. 78–88, 2003. eng. 16

[75] G. Carpenter, "Nuclear localization and possible functions of receptor tyrosine kinases," *Curr Opin Cell Biol*, vol. 15, no. 2, pp. 143–8, 2003. eng. 17

[76] R.-H. Tao and I. N. Maruyama, "All egf(erbb) receptors have preformed homo- and heterodimeric structures in living cells," *Journal of Cell Science*, vol. 121, no. 19, pp. 3207–3217, 2008. eng. 17, 30, 54, 87

[77] M. L. Janmaat and G. Giaccone, "The epidermal growth factor receptor pathway and its inhibition as anticancer therapy," *Drugs Today*, vol. 39 Suppl C, pp. 61–80, 2003. eng. 17

[78] G. Pines, W. J. Kstler, and Y. Yarden, "Oncogenic mutant forms of egfr: lessons in signal transduction and targets for cancer therapy," *FEBS Lett*, vol. 584, no. 12, pp. 2699–706, 2010. eng. 17

[79] K. M. Ferguson, "Structure-based view of epidermal growth factor receptor regulation," *Annu. Rev. Biophys.*, vol. 37, pp. 353–73, 2008. eng. 17

[80] K. R. Schmitz and K. M. Ferguson, "Interaction of antibodies with erbb receptor extracellular regions," *Experimental Cell Research*, vol. 315, no. 4, pp. 659–670, 2009. eng. 17

[81] L.-Z. Mi, M. J. Grey, N. Nishida, T. Walz, C. Lu, and T. A. Springer, "Functional and structural stability of the epidermal growth factor receptor in detergent micelles and phospholipid nanodiscs," *Biochemistry*, vol. 47, no. 39, pp. 10314–23, 2008. eng. 19, 47, 71

[82] L.-Z. Mi, C. Lu, Z. Li, N. Nishida, T. Walz, and T. A. Springer, "Simultaneous visualization of the extracellular and cytoplasmic domains of the epidermal growth factor receptor," *Nat Struct Mol Biol*, vol. 18, no. 9, pp. 984–9, 2011. eng. 19, 26, 71

[83] T. C. Elleman, T. Domagala, N. M. McKern, M. Nerrie, B. Lnnqvist, T. E. Adams, J. Lewis, G. O. Lovrecz, P. A. Hoyne, K. M. Richards, G. J. Howlett, J. Rothacker, R. N. Jorissen, M. Lou, T. P. Garrett, A. W. Burgess, E. C. Nice, and C. W. Ward, "Identification of a determinant of epidermal growth factor receptor ligand-binding specificity

using a truncated, high-affinity form of the ectodomain," *Biochemistry*, vol. 40, no. 30, pp. 8930–9, 2001. eng. 21

[84] J. P. Dawson, M. B. Berger, C.-C. Lin, J. Schlessinger, M. A. Lemmon, and K. M. Ferguson, "Epidermal growth factor receptor dimerization and activation require ligand-induced conformational changes in the dimer interface," *Mol Cell Biol*, vol. 25, no. 17, pp. 7734–42, 2005. eng. 21

[85] D. Mattoon, P. Klein, M. A. Lemmon, I. Lax, and J. Schlessinger, "The tethered configuration of the egf receptor extracellular domain exerts only a limited control of receptor function," *Proc Natl Acad Sci USA*, vol. 101, no. 4, pp. 923–8, 2004. eng. 21, 27

[86] F. Walker, S. G. Orchard, R. N. Jorissen, N. E. Hall, H.-H. Zhang, P. A. Hoyne, T. E. Adams, T. G. Johns, C. Ward, T. P. J. Garrett, H.-J. Zhu, M. Nerrie, A. M. Scott, E. C. Nice, and A. W. Burgess, "Cr1/cr2 interactions modulate the functions of the cell surface epidermal growth factor receptor," *J Biol Chem*, vol. 279, no. 21, pp. 22387–98, 2004. eng. 21, 27, 81, 93

[87] F. Ozcan, P. Klein, M. A. Lemmon, I. Lax, and J. Schlessinger, "On the nature of low- and high-affinity egf receptors on living cells," *Proc Natl Acad Sci USA*, vol. 103, no. 15, pp. 5735–40, 2006. eng. 21, 27, 29, 81, 89

[88] R. Bose and X. Zhang, "The erbb kinase domain: Structural perspectives into kinase activation and inhibition," *Experimental Cell Research*, vol. 315, no. 4, pp. 649–658, 2009. eng. 21

[89] N. Jura, X. Zhang, N. F. Endres, M. A. Seeliger, T. Schindler, and J. Kuriyan, "Catalytic control in the egf receptor and its connection to general kinase regulatory mechanisms," *Molecular Cell*, vol. 42, no. 1, pp. 9–22, 2011. eng. 21, 22, 24, 25, 26

[90] N. Gotoh, A. Tojo, M. Hino, Y. Yazaki, and M. Shibuya, "A highly conserved tyrosine residue at codon 845 within the kinase domain is not required for the transforming

activity of human epidermal growth factor receptor," *Biochem Biophys Res Commun*, vol. 186, no. 2, pp. 768–74, 1992. eng. 22, 89

[91] X. Zhang, K. A. Pickin, R. Bose, N. Jura, P. A. Cole, and J. Kuriyan, "Inhibition of the egf receptor by binding of mig6 to an activating kinase domain interface," *Nature*, vol. 450, no. 7170, pp. 741–744, 2007. eng. 22

[92] H. L. De Bondt, J. Rosenblatt, J. Jancarik, H. D. Jones, D. O. Morgan, and S. H. Kim, "Crystal structure of cyclin-dependent kinase 2," *Nature*, vol. 363, no. 6430, pp. 595–602, 1993. eng. 22

[93] W. Xu, A. Doshi, M. Lei, M. J. Eck, and S. C. Harrison, "Crystal structures of c-src reveal features of its autoinhibitory mechanism," *Molecular Cell*, vol. 3, no. 5, pp. 629–38, 1999. eng. 22

[94] J. Monsey, W. Shen, P. Schlesinger, and R. Bose, "Her4 and her2/neu tyrosine kinase domains dimerize and activate in a reconstituted in vitro system," *J Biol Chem*, vol. 285, no. 10, pp. 7035–44, 2010. eng. 24

[95] J. L. Macdonald-Obermann 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*, vol. 284, no. 20, pp. 13570–6, 2009. eng. 24, 27, 81, 92

[96] K. W. Thiel and G. Carpenter, "Epidermal growth factor receptor juxtamembrane region regulates allosteric tyrosine kinase activation," *Proc Natl Acad Sci USA*, vol. 104, no. 49, pp. 19238–43, 2007. eng. 24, 92

[97] M. Red Brewer, S. H. Choi, D. Alvarado, K. Moravcevic, A. Pozzi, M. A. Lemmon, and G. Carpenter, "The juxtamembrane region of the egf receptor functions as an activation domain," *Molecular Cell*, vol. 34, no. 6, pp. 641–51, 2009. eng. 24, 26, 47, 92

[98] E. R. Wood, L. M. Shewchuk, B. Ellis, P. Brignola, R. L. Brashear, T. R. Caferro, S. H. Dickerson, H. D. Dickson, K. H. Donaldson, M. Gaul, R. J. Griffin, A. M. Hassell, B. Keith, R. Mullin, K. G. Petrov, M. J. Reno, D. W. Rusnak, S. M. Tadepalli, J. C. Ulrich, C. D. Wagner, D. E. Vanderwall, A. G. Waterson, J. D. Williams, W. L. White, and D. E. Uehling, "6-ethynylthieno[3,2-d]- and 6-ethynylthieno[2,3-d]pyrimidin-4-anilines as tunable covalent modifiers of erbb kinases," *Proc Natl Acad Sci USA*, vol. 105, no. 8, pp. 2773–8, 2008. eng. 24, 26

[99] E. V. Bocharov, K. S. Mineev, P. E. Volynsky, Y. S. Ermolyuk, E. N. Tkach, A. G. Sobol, V. V. Chupin, M. P. Kirpichnikov, R. G. Efremov, and A. S. Arseniev, "Spatial structure of the dimeric transmembrane domain of the growth factor receptor erbb2 presumably corresponding to the receptor active state," *J Biol Chem*, vol. 283, no. 11, pp. 6950–6, 2008. eng. 26

[100] L. Chen, M. Merzlyakov, T. Cohen, Y. Shai, and K. Hristova, "Energetics of erbb1 transmembrane domain dimerization in lipid bilayers," *Biophys J*, vol. 96, no. 11, pp. 4622–30, 2009. eng. 26

[101] J.-P. Duneau, A. P. Vegh, and J. N. Sturgis, "A dimerization hierarchy in the transmembrane domains of the her receptor family," *Biochemistry*, vol. 46, no. 7, pp. 2010–9, 2007. eng. 26

[102] J. M. Mendrola, M. B. Berger, M. C. King, and M. A. Lemmon, "The single transmembrane domains of erbb receptors self-associate in cell membranes," *J Biol Chem*, vol. 277, no. 7, pp. 4704–12, 2002. eng. 26

[103] K. S. Mineev, E. V. Bocharov, Y. E. Pustovalova, O. V. Bocharova, V. V. Chupin, and A. S. Arseniev, "Spatial structure of the transmembrane domain heterodimer of erbb1 and erbb2 receptor tyrosine kinases," *J Mol Biol*, vol. 400, no. 2, pp. 231–43, 2010. eng. 26

[104] B. E. Magun, L. M. Matrisian, and G. T. Bowden, "Epidermal growth factor. ability of tumor promoter to alter its degradation, receptor affinity and receptor number," *J Biol Chem*, vol. 255, no. 13, pp. 6373–81, 1980. eng. 27, 81, 89

[105] M. Shoyab, J. E. De Larco, and G. J. Todaro, "Biologically active phorbol esters specifically alter affinity of epidermal growth factor membrane receptors," *Nature*, vol. 279, no. 5712, pp. 387–91, 1979. eng. 27, 81, 89

[106] A. C. King and P. Cuatrecasas, "Resolution of high and low affinity epidermal growth factor receptors. inhibition of high affinity component by low temperature, cycloheximide, and phorbol esters," *J Biol Chem*, vol. 257, no. 6, pp. 3053–60, 1982. eng. 27, 81

[107] J. L. Macdonald and L. J. Pike, "Heterogeneity in egf-binding affinities arises from negative cooperativity in an aggregating system," *Proceedings of the National Academy of Sciences*, vol. 105, no. 1, pp. 112–7, 2008. eng. 27, 28, 81, 90, 92, 93, 95

[108] J. L. Macdonald-Obermann and L. J. Pike, "Palmitoylation of the egf receptor impairs signal transduction and abolishes high-affinity ligand binding," *Biochemistry*, vol. 48, no. 11, pp. 2505–13, 2009. eng. 27, 81, 92

[109] A. R. Rees, M. Gregoriou, P. Johnson, and P. B. Garland, "High affinity epidermal growth factor receptors on the surface of a431 cells have restricted lateral diffusion," *EMBO J*, vol. 3, no. 8, pp. 1843–7, 1984. eng. 27, 81

[110] J. Schlessinger, "Ligand-induced, receptor-mediated dimerization and activation of egf receptor," *Cell*, vol. 110, no. 6, pp. 669–72, 2002. eng. 27, 28, 81

[111] I. Chung, R. Akita, R. Vandlen, D. Toomre, J. Schlessinger, and I. Mellman, "Spatial control of egf receptor activation by reversible dimerization on living cells," *Nature*, pp. 1–6, 2010. eng. 27, 30, 54, 81, 95

[112] K. Mayawala, 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*, vol. 579, no. 14, pp. 3043–7, 2005. eng. 27, 89

[113] P. Klein, D. Mattoon, M. A. Lemmon, and J. Schlessinger, "A structure-based model for ligand binding and dimerization of egf receptors," *Proc Natl Acad Sci USA*, vol. 101, no. 4, pp. 929–34, 2004. eng. 27, 29, 81

[114] J. Macdonald, Z. Li, W. Su, and L. J. Pike, "The membrane proximal disulfides of the egf receptor extracellular domain are required for high affinity binding and signal transduction but do not play a role in the localization of the receptor to lipid rafts," *Biochimica et Biophysica Acta (BBA) - Molecular Cell Research*, vol. 1763, no. 8, pp. 870–878, 2006. eng. 27

[115] C. Wofsy, B. Goldstein, K. Lund, and H. S. Wiley, "Implications of epidermal growth factor (egf) induced egf receptor aggregation," *Biophys J*, vol. 63, no. 1, pp. 98–110, 1992. eng. 27

[116] D. Alvarado, D. E. Klein, and M. A. Lemmon, "Structural basis for negative cooperativity in growth factor binding to an egf receptor," *Cell*, vol. 142, no. 4, pp. 568–79, 2010. eng. 27, 28, 29, 90

[117] P. De Meyts, "[insulin receptors and mechanism of action of insulin and of insulin-like growth factors]," *Bull Mem Acad R Med Belg*, vol. 149, no. 3-4, pp. 181–90; discussion 190–4, 1994. fre. 27, 30, 90

[118] P. De Meyts, "The insulin receptor: a prototype for dimeric, allosteric membrane receptors?," *Trends Biochem Sci*, vol. 33, no. 8, pp. 376–84, 2008. eng. 27, 30, 90

[119] H. Riedel, T. J. Dull, J. Schlessinger, and A. Ullrich, "A chimaeric receptor allows insulin to stimulate tyrosine kinase activity of epidermal growth factor receptor," *Nature*, vol. 324, no. 6092, pp. 68–70, 1986. eng. 27

[120] U. Coskun, M. Grzybek, D. Drechsel, and K. Simons, "Regulation of human egf receptor by lipids," *Proc Natl Acad Sci USA*, vol. 108, no. 22, pp. 9044–8, 2011. eng. 28, 30, 47, 79, 81

[121] J. A. Berkers, P. M. van Bergen en Henegouwen, A. J. Verkleij, and J. Boonstra, "Membrane vesicles of a431 cells contain one class of epidermal growth factor binding sites," *Biochim Biophys Acta*, vol. 1052, no. 3, pp. 453–60, 1990. eng. 28

[122] A. H. A. Clayton, "Ligand-induced dimer-tetramer transition during the activation of the cell surface epidermal growth factor receptor-a multidimensional microscopy analysis," *Journal of Biological Chemistry*, vol. 280, no. 34, pp. 30392–30399, 2005. eng. 30, 54

[123] T. W. Gadella and T. M. Jovin, "Oligomerization of epidermal growth factor receptors on a431 cells studied by time-resolved fluorescence imaging microscopy. a stereochemical model for tyrosine kinase receptor activation," *The Journal of Cell Biology*, vol. 129, no. 6, pp. 1543–58, 1995. eng. 30, 54

[124] P. Liu, T. Sudhaharan, R. Koh, L. Hwang, S. Ahmed, I. Maruyama, and T. Wohland, "Investigation of the dimerization of proteins from the epidermal growth factor receptor family by single wavelength fluorescence cross-correlation spectroscopy," *Biophysical Journal*, vol. 93, no. 2, pp. 684–698, 2007. eng. 30, 54

[125] M. Martin-Fernandez, D. T. Clarke, M. J. Tobin, S. V. Jones, and G. R. Jones, "Preformed oligomeric epidermal growth factor receptors undergo an ectodomain structure change during signaling," *Biophysical Journal*, vol. 82, no. 5, pp. 2415–2427, 2002. eng. 30, 54

[126] T. Moriki, H. Maruyama, and I. N. Maruyama, "Activation of preformed egf receptor dimers by ligand-induced rotation of the transmembrane domain," *J Mol Biol*, vol. 311, no. 5, pp. 1011–26, 2001. eng. 30, 54

[127] P. Nagy, J. Claus, T. M. Jovin, and D. J. Arndt-Jovin, "Distribution of resting and ligand-bound erbb1 and erbb2 receptor tyrosine kinases in living cells using number and brightness analysis," *Proc Natl Acad Sci USA*, vol. 107, no. 38, pp. 16524–9, 2010. eng. 30, 54

[128] S. Saffarian, Y. Li, E. L. Elson, and L. J. Pike, "Oligomerization of the egf receptor investigated by live cell fluorescence intensity distribution analysis," *Biophys J*, vol. 93, no. 3, pp. 1021–31, 2007. eng. 30, 54

[129] Y. Teramura, J. Ichinose, H. Takagi, K. Nishida, T. Yanagida, and Y. Sako, "Single-molecule analysis of epidermal growth factor binding on the surface of living cells," *EMBO J*, vol. 25, no. 18, pp. 4215–4222, 2006. eng. 30, 54

[130] S. E. D. Webb, S. K. Roberts, S. R. Needham, C. J. Tynan, D. J. Rolfe, M. D. Winn, D. T. Clarke, R. Barraclough, and M. L. Martin-Fernandez, "Single-molecule imaging and fluorescence lifetime imaging microscopy show different structures for high- and low-affinity epidermal growth factor receptors in a431 cells," *Biophysical Journal*, vol. 94, no. 3, pp. 803–819, 2008. eng. 30, 54, 94

[131] X. Yu, K. D. Sharma, T. Takahashi, R. Iwamoto, and E. Mekada, "Ligand-independent dimer formation of epidermal growth factor receptor (egfr) is a step separable from ligand-induced egfr signaling," *Mol Biol Cell*, vol. 13, no. 7, pp. 2547–57, 2002. eng. 30, 54

[132] J. Liu and B. Rost, "Comparing function and structure between entire proteomes," *Protein Sci*, vol. 10, no. 10, pp. 1970–9, 2001. eng. 31

[133] W. D. Kohn, C. M. Kay, and R. S. Hodges, "Orientation, positional, additivity, and oligomerization-state effects of interhelical ion pairs in alpha-helical coiled-coils," *J Mol Biol*, vol. 283, no. 5, pp. 993–1012, 1998. eng. 31

[134] A. Lupas, "Coiled coils: new structures and new functions," *Trends Biochem Sci*, vol. 21, no. 10, pp. 375–82, 1996. eng. 31

[135] F. Crick, "The packing of alpha-helices: simple coiled coils," *Acta Crystallographica*, vol. 6, pp. 689–698, 1953. eng. 31

[136] P. Burkhard, J. Stetefeld, and S. V. Strelkov, "Coiled coils: a highly versatile protein folding motif," *Trends Cell Biol*, vol. 11, no. 2, pp. 82–8, 2001. eng. 31

[137] D. Krylov, I. Mikhailenko, and C. Vinson, "A thermodynamic scale for leucine zipper stability and dimerization specificity: e and g interhelical interactions," *EMBO J*, vol. 13, no. 12, pp. 2849–61, 1994. eng. 31

[138] B. Catimel, M. C. Faux, M. Nerrie, J. Rothacker, L. J. Otvos, J. D. Wade, E. C. Nice, and A. W. Burgess, "The use of coiled-coil interactions for the analysis and micropreparative isolation of adenomatous polyposis coli protein complexes," *J Pept Res*, vol. 58, no. 6, pp. 493–503, 2001. eng. 31, 87

[139] H. Chao, D. L. Bautista, J. Litowski, R. T. Irvin, and R. S. Hodges, "Use of a heterodimeric coiled-coil system for biosensor application and affinity purification," *J Chromatogr B Biomed Sci Appl*, vol. 715, no. 1, pp. 307–29, 1998. eng. 31, 87

[140] W. D. Kohn, O. D. Monera, C. M. Kay, and R. S. Hodges, "The effects of interhelical electrostatic repulsions between glutamic acid residues in controlling the dimerization and stability of two-stranded alpha-helical coiled-coils," *J Biol Chem*, vol. 270, no. 43, pp. 25495–506, 1995. eng. 31

[141] J. R. Litowski, "Designing heterodimeric two-stranded alpha -helical coiled-coils. effects of hydrophobicity and alpha -helical propensity on protein folding, stability, and specificity," *Journal of Biological Chemistry*, vol. 277, no. 40, pp. 37272–37279, 2002. eng. 31, 32, 36, 39, 86, 87

[142] J. R. Litowski and R. S. Hodges, "Designing heterodimeric two-stranded alpha-helical coiled-coils: the effect of chain length on protein folding, stability and specificity," *J Pept Res*, vol. 58, no. 6, pp. 477–92, 2001. eng. 31, 36

[143] B. Tripet, L. Yu, D. L. Bautista, W. Y. Wong, R. T. Irvin, and R. S. Hodges, "Engineering a de novo-designed coiled-coil heterodimerization domain off the rapid detection, purification and characterization of recombinantly expressed peptides and proteins," *Protein Eng*, vol. 9, no. 11, pp. 1029–42, 1996. eng. 31, 43, 87

[144] J. Liu, Y. Deng, Q. Zheng, C.-S. Cheng, N. R. Kallenbach, and M. Lu, "A parallel coiled-coil tetramer with offset helices," *Biochemistry*, vol. 45, no. 51, pp. 15224–31, 2006. eng. 31

[145] J. Liu, Q. Zheng, Y. Deng, C.-S. Cheng, N. R. Kallenbach, and M. Lu, "A seven-helix coiled coil," *Proc Natl Acad Sci USA*, vol. 103, no. 42, pp. 15457–62, 2006. eng. 31, 87

[146] D. A. Lindhout, J. R. Litowski, P. Mercier, R. S. Hodges, and B. D. Sykes, "Nmr solution structure of a highly stablede novo heterodimeric coiled-coil," *Biopolymers*, vol. 75, no. 5, pp. 367–375, 2004. eng. 32, 39, 86

[147] B. Apostolovic and H.-A. Klok, "ph-sensitivity of the e3/k3 heterodimeric coiled coil," *Biomacromolecules*, vol. 9, no. 11, pp. 3173–80, 2008. eng. 36, 39, 45, 46, 86

[148] N. C. Shaner, P. A. Steinbach, and R. Y. Tsien, "A guide to choosing fluorescent proteins," *Nat Meth*, vol. 2, no. 12, pp. 905–909, 2005. eng. 38

[149] E. M. Bublil, G. Pines, G. Patel, G. Fruhwirth, T. Ng, and Y. Yarden, "Kinase-mediated quasi-dimers of egfr," *The FASEB Journal*, vol. 24, no. 12, pp. 4744–4755, 2010. eng. 47, 93

[150] K. S. Yang, J. L. Macdonald-Obermann, D. Piwnica-Worms, and L. J. Pike, "Asp-960/glu-961 controls the movement of the c-terminal tail of the epidermal growth fac-

tor receptor to regulate asymmetric dimer formation," *J Biol Chem*, vol. 285, no. 31, pp. 24014–22, 2010. eng. 47

[151] M. Landau, S. J. Fleishman, and N. Ben-Tal, "A putative mechanism for downregulation of the catalytic activity of the egf receptor via direct contact between its kinase and c-terminal domains," *Structure*, vol. 12, no. 12, pp. 2265–75, 2004. eng. 47

[152] M. Mustafa, A. Mirza, and N. Kannan, "Conformational regulation of the egfr kinase core by the juxtamembrane and c-terminal tail: a molecular dynamics study," *Proteins*, vol. 79, no. 1, pp. 99–114, 2011. eng. 47

[153] C. Qiu, M. K. Tarrant, T. Boronina, P. A. Longo, J. M. Kavran, R. N. Cole, P. A. Cole, and D. J. Leahy, "In vitro enzymatic characterization of near full length egfr in activated and inhibited states," *Biochemistry*, vol. 48, no. 28, pp. 6624–32, 2009. eng. 47, 50, 76, 88

[154] A. Citri, J. Gan, Y. Mosesson, G. Vereb, J. Szollosi, and Y. Yarden, "Hsp90 restrains erbb-2/her2 signalling by limiting heterodimer formation," *EMBO Rep*, vol. 5, no. 12, pp. 1165–1170, 2004. eng. 54

[155] B. D. Adair, J.-P. Xiong, C. Maddock, S. L. Goodman, M. A. Arnaout, and M. Yeager, "Three-dimensional em structure of the ectodomain of integrin alphavbeta3 in a complex with fibronectin," *The Journal of Cell Biology*, vol. 168, no. 7, pp. 1109–18, 2005. eng. 56, 58

[156] S. H. W. Scheres, R. Nez-Ramrez, C. O. S. Sorzano, J. M. Carazo, and R. Marabini, "Image processing for electron microscopy single-particle analysis using xmipp.," *Nat Protoc*, vol. 3, no. 6, pp. 977–990, 2008. 58, 70

[157] S. H. W. Scheres, M. Valle, and J.-M. Carazo, "Fast maximum-likelihood refinement of electron microscopy images.," *Bioinformatics*, vol. 21 Suppl 2, pp. ii243–ii244, Sep 2005. 58, 70, 113

[158] S. H. W. Scheres, M. Valle, R. Nuez, C. O. S. Sorzano, R. Marabini, G. T. Herman, and J.-M. Carazo, "Maximum-likelihood multi-reference refinement for electron microscopy images.," *J Mol Biol*, vol. 348, pp. 139–149, Apr 2005. 58, 70, 113

[159] J. J. Lammerts van Bueren, W. K. Bleeker, A. Brnnstrm, A. von Euler, M. Jansson, M. Peipp, T. Schneider-Merck, T. Valerius, J. G. J. van de Winkel, and P. W. H. I. Parren, "The antibody zalutumumab inhibits epidermal growth factor receptor signaling by limiting intra- and intermolecular flexibility," *Proc Natl Acad Sci USA*, vol. 105, no. 16, pp. 6109–14, 2008. eng. 63, 92

[160] S. H. W. Scheres, H. Gao, M. Valle, G. T. Herman, P. P. B. Eggermont, J. Frank, and J.-M. Carazo, "Disentangling conformational states of macromolecules in 3d-em through likelihood optimization.," *Nat Methods*, vol. 4, pp. 27–29, Jan 2007. 67, 114

[161] V. Wolber, K. Maeda, R. Schumann, B. Brandmeier, L. Wiesmller, and A. Wittinghofer, "A universal expression-purification system based on the coiled-coil interaction of myosin heavy chain," *Biotechnology (NY)*, vol. 10, no. 8, pp. 900–4, 1992. eng. 87

[162] Y. Yano, A. Yano, S. Oishi, Y. Sugimoto, G. Tsujimoto, N. Fujii, and K. Matsuzaki, "Coiled-coil tag–probe system for quick labeling of membrane receptors in living cell," *ACS Chem Biol*, vol. 3, no. 6, pp. 341–5, 2008. eng. 87

[163] K. S. Yang, M. X. G. Ilagan, D. Piwnica-Worms, and L. J. Pike, "Luciferase fragment complementation imaging of conformational changes in the epidermal growth factor receptor," *J Biol Chem*, vol. 284, no. 12, pp. 7474–82, 2009. eng. 87

[164] M. C. Wehr, R. Laage, U. Bolz, T. M. Fischer, S. Grnewald, S. Scheek, A. Bach, K.-A. Nave, and M. J. Rossner, "Monitoring regulated protein-protein interactions using split tev," *Nat Methods*, vol. 3, no. 12, pp. 985–93, 2006. eng. 87

[165] E. G. Hofman, A. N. Bader, J. Voortman, D. J. Van Den Heuvel, S. Sigismund, A. J. Verkleij, H. C. Gerritsen, and P. M. P. Van Bergen En Henegouwen, "Ligand-induced egf receptor oligomerization is kinase-dependent and enhances internalization," *Journal of Biological Chemistry*, vol. 285, no. 50, pp. 39481–39489, 2010. eng. 87, 88

[166] S. K. Muthuswamy, M. Gilman, and J. S. Brugge, "Controlled dimerization of erbb receptors provides evidence for differential signaling by homo- and heterodimers," *Mol Cell Biol*, vol. 19, no. 10, pp. 6845–57, 1999. eng. 87

[167] Q. Wang, G. Villeneuve, and Z. Wang, "Control of epidermal growth factor receptor endocytosis by receptor dimerization, rather than receptor kinase activation," *EMBO Rep*, vol. 6, no. 10, pp. 942–8, 2005. eng. 87

[168] C. Ruch, G. Skiniotis, M. O. Steinmetz, T. Walz, and K. Ballmer-Hofer, "Structure of a vegf-vegf receptor complex determined by electron microscopy," *Nat Struct Mol Biol*, vol. 14, no. 3, pp. 249–50, 2007. eng. 87

[169] J. Amann, S. Kalyankrishna, P. P. Massion, J. E. Ohm, L. Girard, H. Shigematsu, M. Peyton, D. Juroske, Y. Huang, J. Stuart Salmon, Y. H. Kim, J. R. Pollack, K. Yanagisawa, A. Gazdar, J. D. Minna, J. M. Kurie, and D. P. Carbone, "Aberrant epidermal growth factor receptor signaling and enhanced sensitivity to egfr inhibitors in lung cancer," *Cancer Research*, vol. 65, no. 1, pp. 226–35, 2005. eng. 89

[170] S. Adak, D. DeAndrade, and L. J. Pike, "The tethering arm of the egf receptor is required for negative cooperativity and signal transduction," *J Biol Chem*, vol. 286, no. 2, pp. 1545–55, 2011. eng. 90, 93

[171] J. Kästner, H. H. Loeffler, S. K. Roberts, M. L. Martin-Fernandez, and M. D. Winn, "Ectodomain orientation, conformational plasticity and oligomerization of erbb1 receptors investigated by molecular dynamics," *Journal of Structural Biology*, vol. 167, no. 2, pp. 117–128, 2009. file://localhost/Users/fernandez/Desktop/Papers/2009/K94

[172] C. J. Tynan, S. K. Roberts, D. J. Rolfe, D. T. Clarke, H. H. Loeffler, J. Kastner, M. D. Winn, P. J. Parker, and M. L. Martin-Fernandez, "Human epidermal growth factor receptor (egfr) aligned on the plasma membrane adopts key features of drosophila egfr asymmetry," *Mol Cell Biol*, vol. 31, no. 11, pp. 2241–2252, 2011. eng. 94

[173] J. A. Krall, E. M. Beyer, and G. Macbeath, "High- and low-affinity epidermal growth factor receptor-ligand interactions activate distinct signaling pathways," *PLoS ONE*, vol. 6, no. 1, p. e15945, 2011. eng. 95

[174] E. A. Jares-Erijman and T. M. Jovin, "Fret imaging," *Nat Biotechnol*, vol. 21, no. 11, pp. 1387–95, 2003. eng. 104

[175] C. Suloway, J. Pulokas, D. Fellmann, A. Cheng, F. Guerra, J. Quispe, S. Stagg, C. S. Potter, and B. Carragher, "Automated molecular microscopy: the new leginon system.," *J Struct Biol*, vol. 151, pp. 41–60, Jul 2005.

[176] J. A. Mindell and N. Grigorieff, "Accurate determination of local defocus and specimen tilt in electron microscopy.," *J Struct Biol*, vol. 142, pp. 334–347, Jun 2003.

[177] G. Tang, L. Peng, P. R. Baldwin, D. S. Mann, W. Jiang, I. Rees, and S. J. Ludtke, "Eman2: an extensible image processing suite for electron microscopy.," *J Struct Biol*, vol. 157, pp. 38–46, Jan 2007.

[178] C. O. S. Sorzano, R. Marabini, J. Velzquez-Muriel, J. R. Bilbao-Castro, S. H. W. Scheres, J. M. Carazo, and A. Pascual-Montano, "Xmipp: a new generation of an open-source image processing package for electron microscopy.," *J Struct Biol*, vol. 148, pp. 194–204, Nov 2004.

[179] E. F. Pettersen, T. D. Goddard, C. C. Huang, G. S. Couch, D. M. Greenblatt, E. C. Meng, and T. E. Ferrin, "Ucsf chimera–a visualization system for exploratory research and analysis.," *J Comput Chem*, vol. 25, pp. 1605–1612, Oct 2004.

7

Acknowledgements

I want to thank all members of the Marlovits lab, for the great atmosphere, and of course to Thomas, who gave me the opportunity of working on this exciting project. I have to give special thanks to Matthias, who was of incredible help with the three-dimensional reconstruction.

I also want to thank the Cowan group, for all the good times I had with them, and especially Carrie, who helped me many, many times.

I want to thank Alex Dammermann and his group for letting me work with them during the months that it took me to write this thesis, and to my examination committee, that revised it.

I have got so much help from the groups at IMBA and IMP that I would just like to thank everyone in the institutes for sharing reagents and especially knowledge, that allowed me to give this project a very nice shape.

In the personal aspect, I have many people to thank, all those that have shared weekends and evenings with me, with an especial mention to Sabina, Javi, Pablo, Sara and Carmen. And of course, Martin and my little Pepe, who have been supporting me during all these years. If I got to this point, it is thanks to all of you!

8

Curriculum vitae

Jesús Fernández Rodríguez

Personal Information

Name: Jesús Fernández Rodríguez
Nationality: Spanish
Date and Place of Birth: 18th August 1982, Madrid, Spain
Email: jesus.fernandez@imba.oeaw.ac.at

Education

- **Institute of Molecular Biotechnology**, Vienna, Austria, *Thomas Marlovits Group.*
  - Ph.D. Candidate, 2006–2011
  - Dissertation Topic: *Biochemical and Structural Characterization of EGF receptor preformed dimers*
- **University of Vienna**, Vienna, Austria, *Erasmus grant.*
  - MSc. Molecular Biology, 2005–2006
- **Universidad Autónoma de Madrid**, Madrid, Spain, *Faculty of Sciences.*
  - MSc, Biochemistry, 2000–2006.
  - Diploma Thesis: *Mechanical Unfolding of a Cohesin–I Module Using Molecular Dynamics Simulation Techniques. Cajal Institute, CSIC, Madrid, Spain*

Research Experience

- **Max F. Perutz Laboratories**, Vienna, Austria, *Alex Dammermann Group*
  - Research Assistant, 2011
  - Topic: *Centriole Assembly and Function*
- **Institute of Molecular Biotechnology**, Vienna, Austria, *Thomas Marlovits Group*
  - PhD Student in Thomas Marlovits Group, 2006–2011
  - Dissertation Topic: *Biochemical and Structural Characterization of EGF receptor preformed dimers*
- **Cajal Institute, CSIC**, Madrid, Spain, *Mariano Carrión Group*
  - Diploma Thesis student, 2004–2005
  - Diploma Thesis: *Mechanical Unfolding of a Cohesin–I Module Using Molecular Dynamics Simulation Techniques.*

Honors and Awards

Erasmus fellowship at University of Vienna, 2005–06,
Extraordinary Prize from the Spanish Governement for the first year of University (outstanding results in the high–school leaving certificate), 2000.

Publications

Carrión–Vázquez, M., Oberhauser, A. F., Díez, H., Hervás, R., Oroz, **J., Fernández**, J., and Martínez-Martín, D. *Protein nanomechanics – As studied by Single–Molecule Force Spectroscopy AFM.* In: Arrondo, J.L.R., and Alonso, A (eds.) Emerging Techniques in Biophysics. Springer–Verlag, Heidelberg (2005).

Languages

Spanish:	mother tongue
English:	high advanced level (written and spoken)
German:	advanced level (written and spoken)