Biophysical studies of EGFR conformation and interaction

Dissertation

zur

Erlangung des Doktorgrades (Dr. rer. nat.)

der Mathematisch-Naturwissenschaftlichen Fakultät

der Rheinischen Friedrich-Wilhelms-Universität Bonn

vorgelegt von

Dongsheng Yin

aus

Dongying, China

Bonn, 2018

Angefertigt mit Genehmigung der Mathematisch-Naturwissenschaftlichen Fakultät
der Rheinischen Friedrich-Wilhelms-Universität Bonn

Part of this thesis is published in:

Yin, D.M., Hannam, J.S., Schmitz, A., Schiemann, O., Hagelueken, G., Famulok, M., 2017. Studying the Conformation of a Receptor Tyrosine Kinase in Solution by Inhibitor-Based Spin Labeling. Angew. Chem. Int. Ed. 56, 8417–8421. doi:10.1002/ange.201703154

Yin, D.M., Hammler, D., Peter, M.F., Marx, A., Schmitz, A., Hagelueken, G., 2018. Inhibitor-Directed Spin Labelling-A High Precision and Minimally Invasive Technique to Study the Conformation of Proteins in Solution. Chem. Eur. J. 24, 6665–6671. doi:10.1002/chem.201706047

Contents

Contents

List of Abbreviations

AP-2	Adaptor protein 2
APS	Ammonium peroxydisulphate
Arf	ADP-ribosylation factor
ARNO	Arf nucleotide-binding site opener
ATP	Adenosine 5'-triphosphate
BFA	Brefeldin A
BSA	Bovine serum albumin
CDK2	Cyclin-dependent kinase 2
Cryo-EM	Cryogenic electron microscopy
CT	C-terminal tail
CW-EPR	Continuous-wave electron paramagnetic resonance
DMSO	Dimethyl sulfoxide
DOL	Degree of labeling
DPBS	Dulbecco's phosphate-buffered saline
DTT	1,4-Dithiothreitol
ECD	Extracellular domain
E.coli	*Escherichia coli*
EDTA	Ethylenediaminetetraacetic acid
EGF	Epidermal growth factor
EGFR	Epidermal growth factor receptor
EGTA	Ethylene glycol-bis(β-aminoethyl ether)-N,N,N',N'-tetraacetic acid
EM	Electron microscopy
EPR	Electron paramagnetic resonance
ER	Endoplasmic reticulum
FRET	Fluorescence resonance energy transfer
GDP	Guanosine 5'-diphosphate
GEF	Guanine nucleotide exchange factor
GTP	Guanosine 5'-triphosphate
HEPES	1-[4-(2-Hydroxyethyl)-1-piperazinyl]ethane-2-sulfonic acid
HPLC	High-performance liquid chromatography
HSQC	Heteronuclear single quantum coherence spectroscopy
ICD	Intracellular domain
IDSL	Inhibitor-directed spin labeling
IP4	Inositol 1,3,4,5-tetrakisphosphate
IPTG	Isopropyl β-D-1-thiogalactopyranoside
JM	Juxtamembrane
LB	Lysogeny broth

MAP kinase	Mitogen-activated protein kinase
MBP	Maltose-binding protein
MD	Molecular dynamics
MST	Microscale thermophoresis
MTSL	Methanethiosulfonate spin label
MW	Molecular weight
NBD	Nitrobenzoxadiazole
NHS	N-hydroxysuccinimide
NMR	Nuclear magnetic resonance
NTA	Nitrilotriacetic acid
PAGE	Polyacrylamide gel electrophoresis
PBR	Polybasic region
PEG	Polyethylene glycol
PELDOR	Pulsed electron-electron double resonance
PH	Pleckstrin homology
pI	Isoelectric point
PIP2	Phosphatidylinositol 4,5-bisphosphate
PIP3	Phosphatidylinositol 3,4,5-trisphosphate
PLCγ	Phospholipase Cγ
POPC	1-palmitoyl-2-oleoyl-sn-glycero-3-phosphocholine
POPS	1-palmitoyl-2-oleoyl-sn-glycero-3-phospho-L-serine
PROXYL	(2,2,5,5-tetramethylpyrrolidin-1-yl)oxidanyl
RT	Room temperature
RTK	Receptor tyrosine kinase
SBP	Streptavidin-binding peptide
SDS	Sodium dodecyl sulfate
SEC	Size exclusion chromatography
TBS	Tris-buffered saline
TEMED	Tetramethylethylenediamine
TES	N-tris(hydroxymethyl)methyl-2-aminoethanesulfonic acid
TEV	Tobacco etch virus
TM	Transmembrane
UAA	Unnatural amino acid
WT	Wildtype

List of Figures

List of Tables

Abstract

Epidermal growth factor receptor (EGFR) plays such a crucial role in cell signaling that its activity is intricately regulated. According to the prevailing activation model, EGF binding to EGFR initiates a series of structural rearrangements, leading to allosteric activation of the kinase domain. During this process the oligomeric state of the kinase is promoted from monomer or inactive (symmetric) dimer to active (asymmetric) dimer. Due to the direct correlation with kinase activity, asymmetric/symmetric dimers are key to this activation model, and therefore have been extensively studied via crystallography and negative-stain electron microscopy. However, high-resolution structural description of these entities in solution is still lacking. Here, via pulsed electron-electron double resonance (PELDOR) spectroscopy, a number of different EGFR constructs are shown to exist in solution in a single oligomeric form that reflects asymmetric dimerization as reported in crystallization studies. Mutagenesis and analytical gel filtration studies corroborate this dimeric arrangement. Notably, a new approach to label proteins for PELDOR studies, *i.e.* inhibitor-directed spin labeling, has been established in this pilot study of EGFR, which exemplifies its potential for a more general application in PELDOR-based protein conformation studies. Beyond the well-established model of receptor-intrinsic control of the allosteric activation mechanism, evidence is accumulating for an additional layer of regulation of EGFR activity by less well-understood cytoplasmic modulators. Therefore, the second part of this thesis is devoted to investigating interaction of EGFR with Arf nucleotide-binding site opener (ARNO), a potential activator, via microscale thermophoresis (MST) and nuclear magnetic resonance (NMR) studies. MST assays not only reveal EGFR:ARNO interaction, but also identify the recognition elements as the Sec7 domain of ARNO and the juxtamembrane (JM) segment of EGFR. Competition experiments demonstrate that calmodulin, another EGFR activity modulator, targets the same or a similar binding site in JM yet with stronger affinity compared to Sec7. NMR studies characterize the amino acids of Sec7 and JM that are involved in their interaction and narrow down the binding sites of Sec7 and calmodulin to the same N-terminal region in JM. To integrate the results of these studies into the current view of EGFR signaling an "endocytosis hypothesis" and a "priming hypothesis" are presented.

1 Introduction

1.1 EGFR: introduction and signaling pathways

Epidermal growth factor receptor (EGFR) is a single-pass transmembrane protein from HER/ErbB family (HER for human EGF receptor and ErbB named after related avian erythroblastosis oncogene B), which also includes another three members, *i.e.* HER2/ErbB2/Neu, catalytically impaired HER3/ErbB3 and HER4/ErbB4. EGFR plays such a key role in cell growth, proliferation, migration and differentiation, that its deregulation is often related to carcinogenesis (reviewed by Yarden and Sliwkowski, 2001). As a type I transmembrane protein, EGFR is expressed in its nascent form containing 24-residue signal peptide, which is cleaved off in endoplasmic reticulum (ER), leading to mature EGFR (Ullrich et al., 1984). In this thesis, sequence numbering for EGFR will exclude the signal peptide. As a receptor tyrosine kinase (RTK), EGFR contains EGF-binding extracellular domain (ECD) and kinase-harboring intracellular domain (ICD) which are separated by the helical transmembrane domain (TM). ICD is subdivided into juxtamembrane domain (JM), kinase domain and autophosphorylation site-containing C-terminal tail (CT), with first and second halves of JM being named JM-A and JM-B, respectively (Figure 1.1). Notably, a segment dubbed as "kinase core", which incorporates parts of JM-B and CT and spans residue 672-998, is extensively studied as the integrated kinase unit (Jura et al., 2009; Park et al., 2015; Stamos et al., 2002; Yun et al., 2007; Zhang et al., 2006, 2007).
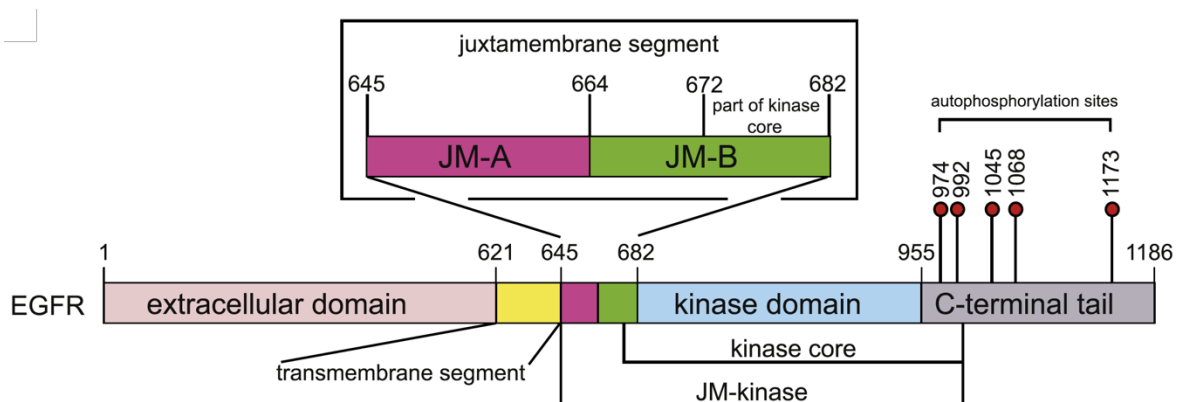
Figure 1.1 Domain sketch of full length EGFR. Sequence numbering is based on human EGFR without signal peptide. Note that depicted autophosphorylation sites are representative yet incomplete. Figure is taken from Jura et al., 2009.

Ligand-induced dimerization is the fundamental idea of canonical model for EGFR activation, in which binding of EGF to ECD mediates EGFR dimerization, and thereby

activates its kinase domain and leads to autophosphorylation at CT tyrosines (Yarden and Schlessinger, 1987b, 1987a). Right at the hub of signal transduction, activated EGFR triggers multitudes of signaling pathways, among which the major ones are Ras-Raf-MEK-ERK, PLCγ-PKC and PI3K-Akt-mTOR pathways. In general, these pathways are initiated through recognition of phosphotyrosines on EGFR by SH2 (Src Homology 2) or PTB (phosphotyrosine binding) domains of effector proteins (reviewed by Pawson, 2004). In Ras-Raf-MEK-ERK pathway, such effector is Grb2 (growth factor receptor-bound protein 2), which docks onto certain phosphotyrosines via its SH2 domain. Grb2 in turn recruits SOS (son of sevenless) through its SH3 (Src Homology 3) domain. SOS functions as guanine nucleotide exchange factor (GEF) to activate GTPase Ras (rat sarcoma), which relays the activation signal further downstream through a chain of phosphorylation reactions, leading to sequential activations of Raf (rapidly accelerated fibrosarcoma), MEK (MAPK/ERK kinase) and ERK (extracellular-signal regulated kinase). Similarly, PLCγ (phospholipase C γ) binds to specific EGFR phosphotyrosines via its SH2 domain. Recruited active PLCγ hydrolyzes plasma membrane-embedded PIP2 (phosphatidylinositol 4,5-bisphosphate) into DAG (diacylglycerol) and IP3 (inositol 1,4,5-triphosphate). IP3, upon binding to Ca^{2+} channel on ER, releases its Ca^{2+} store into cytosol. Ca^{2+} and DAG activate PKC (protein kinase C) together. PI3K-Akt-mTOR cascade is triggered in a slightly different way. Instead of direct docking, PI3K (phosphatidylinositol-3-kinase) is recruited to activated EGFR via Grb2 and GAB1 (Grb2-associated-binding protein 1). Activated PI3K catalyzes phosphorylation of PIP2 into PIP3 (phosphatidylinositol 3,4,5-trisphosphate), and PIP3 facilitates recruitment of Akt (also known as protein kinase B, PKB) to plasma membrane. Thereafter, Akt and mTOR (mammalian target of rapamycin) are sequentially phosphorylated and activated. ERK, PKC and mTOR mediate various cellular processes in response to extracellular stimulation of growth factor (reviewed by Wee and Wang, 2017; reviewed by Yarden and Sliwkowski, 2001).

1.2 Catalytic mechanism of EGFR kinase

As a receptor tyrosine kinase, EGFR harbors in its cytoplasmic department a kinase domain, which upon activation catalyzes phosphoryl transfer reaction between adenosine 5'-triphosphate (ATP) and specific tyrosines on EGFR. The kinase core is composed of N-lobe and C-lobe. During the reaction, ATP is sandwiched between these two lobes and surrounded by key regulatory and catalytic elements from both sides. According to previous structural studies, N-lobe is mainly composed of β-strands while C lobe is highly helical (Blair et al., 2007; Jura et al., 2009; Stamos et al., 2002; Zhang et al., 2006). To

appreciate the catalytic mechanism of EGFR, a crystal structure from one of those studies is presented in Figure 1.2 (A), which displays EGFR kinase in active conformation, complexed with an ATP analog-peptide conjugate (PDB ID:2GS6; Zhang et al., 2006). An inactive kinase from the same study, bearing V924R mutation and complexed with ATP analog AMP-PNP, is presented in Figure 1.2 (B) for comparison (PDB ID: 2GS7). As common practice, secondary structure of EGFR kinase core follows the nomenclature of protein kinase A (Taylor and Radzio-Andzelm, 1994; Zhang et al., 2006).
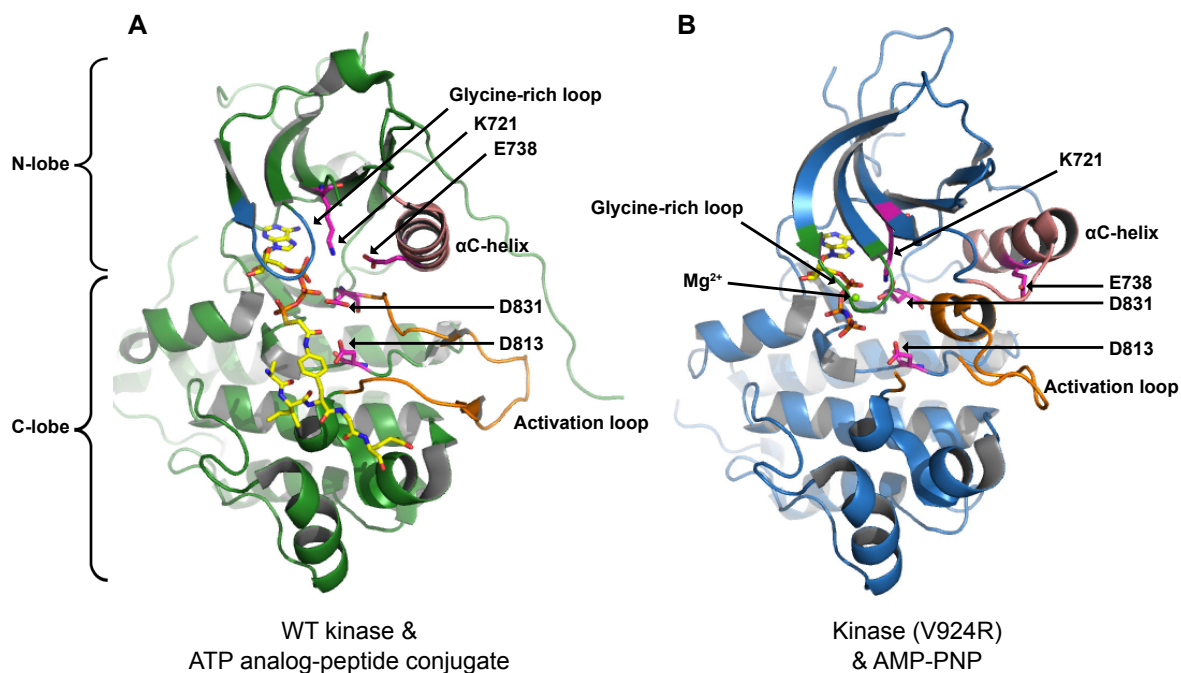
Figure 1.2 Crystal structures of EGFR kinase in active (A) or inactive conformation (B). Wildtype kinase core, *i.e.* ICD(672-998), in (A) and its inactivating mutant V924R in (B) are depicted as dark green and dark blue cartoons, respectively. Corresponding ligands on ATP-binding sites are labeled below the figure, and shown as sticks with carbons colored yellow. Relevant residues are labeled and shown as sticks with carbons colored magenta. Mg^{2+} ion is depicted as a green sphere in (B), and it is missing in (A) due to technical reason. Activation loops are colored orange and αC-helices are shown as pink. Glycine-rich loops are colored dark blue (A) or dark green (B). Note the different orientations of αC-helices in (A) and (B), and the distinct activation loop structure in (B) compared to (A). AMP-PNP: Adenosine 5′-(β,γ-imido)triphosphate. Figure was prepared from PDB ID: 2GS6 (A) and 2GS7 (B) with PyMOL Molecular Graphics System.

For N-lobe, five β-strands (β1-β5) are organized into an antiparallel β-sheet. β1- and β2-strands are connected by a flexible glycine-rich loop (also known as P-loop), which approaches triphosphate moiety to facilitate ATP orientation, while V702 and A719 from β2- and β3-strands, together with L820 from C-lobe, accommodate adenine moiety via hydrophobic contacts. Connecting β3- and β4-strands on the sequence, αC-helix plays a central role in EGFR kinase activation. E738 in the middle of αC-helix forms a salt bridge

with K721 on β3-strand, which in turn interacts with α- and β-phosphates to facilitate ATP binding. The K721-E738 salt bridge maintains "αC-helix in" conformation, an essential configuration of active kinase.

C-lobe of kinase core is primarily composed of six helices (αD-αI). Based on the sequence, catalytic loop and activation loop (also known as A-loop) lie between αE- and αF-helices. On catalytic loop (residue 811-818), a region three amino acids downstream of αE-helix, catalytic residue D813 of HRD motif acts as a base to deprotonate hydroxyl group on tyrosine, promoting its nucleophilic attack towards γ-phosphate of ATP. Activation loop (residue 831-852) begins with D831 of DFG motif, which coordinates an Mg^{2+} ion together with α- and β-phosphates and thereby promotes ATP binding. Altogether, K721, E738, D813 and D831 are considered as the most critical components of catalytic and regulatory machinery of EGFR kinase (reviewed by Roskoski, 2014), and it will be explained in next section concerning how these residues come into position for phosphorylation, in response to the signal beyond plasma membrane.

1.3 Allosteric activation and autoinhibition of EGFR

Comparing (A) and (B) in Figure 1.2 reveals that the major structural differences can be narrowed down to changes in two regions, *i.e.* αC-helix and activation loop. In active kinase structure, αC-helix is leaning towards ATP-binding site and forming K721-E738 salt bridge ("αC-helix in" conformation), while activation loop adopts an extended and open conformation suited for substrate docking. Yet in inactive conformation, αC-helix rotates away with K721-E738 salt bridge disrupted ("αC-helix out" conformation), while activation loop is coiled and closed, typically with a newly-formed short helix (residue 833-838) at the beginning of the loop, packing against αC-helix to stabilize its position in "αC-helix out" form. Taken together, active EGFR kinase adopts "αC-helix in" and "activation loop open" form, whereas inactive one displays "αC-helix out" and "activation loop closed" form. Additionally, in both kinase structures, D831 of DFG motif face towards ATP-binding site, a conformation known as "DFG-in". In general, inactive conformations across different protein kinases are quite diverse, in contrast to active kinases, which share the highly conserved conformation due to structural restraints required for catalysis of common reaction, *i.e.* protein phosphorylation (reviewed by Jura et al., 2011). As for EGFR kinase, its inactive structure resembles that of Src kinase (PDB ID: 2SRC; Xu et al., 1999) or cyclin-dependent kinase (CDK for short; PDB ID: 1HCL; De Bondt et al., 1993), hence the name "Src/CDK-like inactive conformation" for this structure (reviewed by Jura et al., 2011; Zhang et al., 2006).
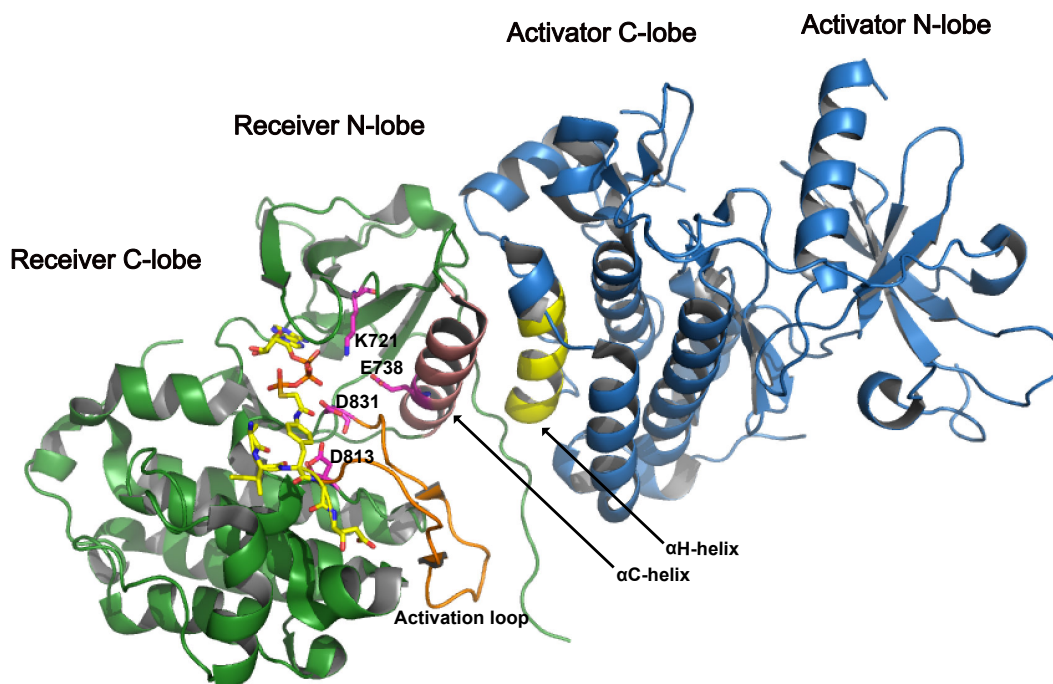
Figure 1.3 Crystal structures of EGFR kinase in asymmetric (active) dimer conformation. Kinase core, *i.e.* ICD(672-998), is depicted as dark green or dark blue cartoon in either monomer. ATP analog-peptide conjugate in receiver is shown as sticks with carbons colored yellow. Relevant residues are labeled and shown as sticks with carbons colored magenta. Activation loop of receiver is colored orange, receiver αC-helix pink and activator αH-helix yellow. Figure was prepared from PDB ID: 2GS6 with PyMOL Molecular Graphics System.

Kinase activation is a process of structural transition from inactive to active conformation, during which most other RTKs require tyrosine phosphorylation on the activation loop to relieve its occlusion at the ATP- or substrate-binding site. Untypically, this step is not necessary for EGFR kinase activation, which instead involves an allosteric mechanism (reviewed by Hubbard and Miller, 2007; reviewed by Lemmon and Schlessinger, 2010; Gotoh et al., 1992; Stamos et al., 2002; Zhang et al., 2006). According to this mechanism, one kinase molecule ("activator") allosterically activates another ("receiver") by forming an active dimer, in which C-lobe of the activator packs against N-lobe of the receiver. Due to the head-to-tail orientation of molecules, active dimer is also referred to as asymmetric dimer. Upon dimerization, αH-helix of activator contacts αC-helix of the receiver and shifts it to "αC-helix in" position, restoring the critical K721-E738 salt bridge for ATP binding. Rotation of αC-helix also disengages itself from interaction with the short helix on activation loop, which "relaxes" the loop to open conformation for substrate docking (Figure 1.3). The way that activator triggers structural movement of receiver via its impact on αC-helix, is very similar to the activation of CDK2 by cyclin A (Jeffrey et al., 1995). Moreover, further studies indicate that JM segment stabilizes the asymmetric dimer and contributes greatly to the kinase activity. Part of receiver JM-B (residue 664-672), known

as "juxtamembrane latch", "cradles" the C-lobe of activator and thereby maintains the dimer in its asymmetric form (Figure 1.4; Jura et al., 2009; Red Brewer et al., 2009). Additionally, JM-A segment forms antiparallel helical dimer in active kinase, further stabilizing the asymmetric dimer (Endres et al., 2013; Jura et al., 2009). Therefore, a construct containing only EGFR kinase and JM is intrinsically active, with activity in solution increased by almost 70-fold due to JM (Jura et al., 2009).
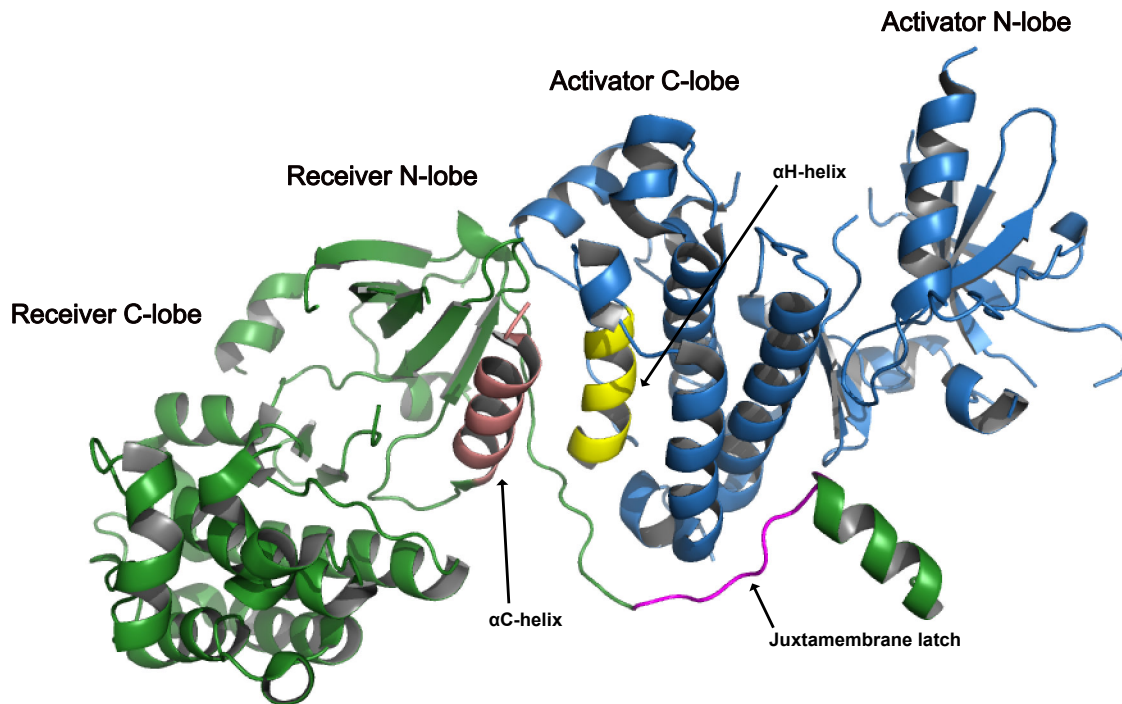
Figure 1.4 Crystal structure of EGFR kinase to reveal juxtamembrane latch. ICD(645-998) bearing K721M mutation is depicted as dark green or dark blue cartoon in either monomer. In activator, fragment spanning residue 645-671 is omitted. αC-helix of receiver is colored pink, αH-helix of activator yellow, and juxtamembrane latch magenta. Figure was prepared from PDB ID: 3GOP (Red Brewer et al., 2009) with PyMOL Molecular Graphics System.

In contrast, activity of full length EGFR, which is more than a mere kinase protein, is normally kept at bay in resting phase and only boosted upon stimulation by growth factors. Regulation of EGFR activity is very crucial, as evidenced by constitutive activation of EGFR often observed in various cancer conditions (reviewed by Masuda et al., 2012; reviewed by Roskoski, 2014; reviewed by Soria et al., 2012). As will be introduced, multiple layers of delicate autoinhibition mechanisms in EGFR have been evolved by nature and coupled together to regulate its activity, so EGFR activation is a chain of events to loosen up these inhibitory restraints so that kinase domain can dimerize in active form and initiate signal transduction by autophosphorylation. Activation and inhibition are like two sides of a coin. To appreciate the activation process, autoinhibition

should be clarified first. Three major inhibitory elements for EGFR activity are present, *i.e.* ECD obstruction, electrostatic attraction to anionic lipids and CT hindrance.

ECD of EGFR functions like a scout, posted beyond the plasma membrane, ready to collect the signal from rapidly changing cell exterior and relay it to the interior. To achieve this, binding of ligand such as EGF induces ECD dimerization as the first step. Among RTKs, dimer interface of ECD can be composed of ligand alone as in the case of nerve growth factor receptor (Urferl et al., 1995; Wlesmann et al., 1999), ligand together with receptor as in the case of stem cell factor receptor c-Kit (Blechman et al., 1995; Lev et al., 1993) or ECD is constitutively dimerized as in insulin receptor (reviewed by De Meyts, 2008; reviewed by Ward et al., 2007). In contrast, EGFR-ECD dimerization interface solely consists of receptor contacts (reviewed by Endres et al., 2014; reviewed by Lemmon et al., 2014). ECD of EGFR is subdivided into domain I (residue 1-165), domain II (residue 166-310), domain III (residue 311-480) and domain IV (residue 481-620). Domain I and III are both leucine-rich and share a rigid β-helix structure, while domain II and domain IV adopt extended structures and are rich in cysteines, which form disulfide bridges (reviewed by Roskoski, 2014). Upon stimulation, EGF binds simultaneously to domain I and III of the same EGFR molecule. The binding event brings those two domains closer and thereby promotes a global rotation of domain I and II to expose the dimerization arm, a β-hairpin in domain II that mediates ECD dimerization (PDB ID: 3NJP; Lu et al., 2010). In the autoinhibited state, however, ECD exists in a "tethered" form with this β-hairpin making intramolecular contact with domain IV, which obstructs ECD dimerization (PDB ID: 1NQL; reviewed by Burgess et al., 2003; Ferguson et al., 2003). Since EGF is not required as part of dimer interface, ECD obstruction is crucial in preventing ECD from ligand-independent dimerization (reviewed by Endres et al., 2014).

The second autoinhibition mechanism involves electrostatic attraction between JM and anionic lipids. Previous studies indicate that anionic lipids on the inner leaflet of plasma membrane interacts with positively charged JM, especially JM-A segment (Aifa et al., 2002; Sato et al., 2006; Sengupta et al., 2009). This electrostatic attraction tethers JM to the plasma membrane, hindering its activating functions such as JM-A antiparallel helical dimerization and juxtamembrane latching. Correspondingly, a TM-JM structural coupling mechanism has been proposed to relieve JM from its tethered state (Arkhipov et al., 2013; Endres et al., 2013). In the cell exterior, EGF-induced dimerization rearranges domain IV of ECD into a V shape, which leads to contact between C-termini of two ECDs and close proximity of N-termini of TM helical dimer. This in turn pulls JM-A off plasma membrane, promoting its antiparallel helices formation, reorientation of kinases into asymmetric dimer and allosteric activation, probably with the help of a cellular modulator, *i.e.* calmodulin, via

competitive binding to JM (McLaughlin et al., 2005). In contrast, during resting phase, N-termini of TM helices are kept at distance due to ECD obstruction, which renders TM in either monomeric state, or inactive dimeric state with C-termini of TM in proximity and JM tethered. This concerted movement of TM and JM delicately connects ECD dimerization to kinase activation, so that extracellular signal effectively passes through membrane and elicits cellular responses (Arkhipov et al., 2013; Endres et al., 2013).
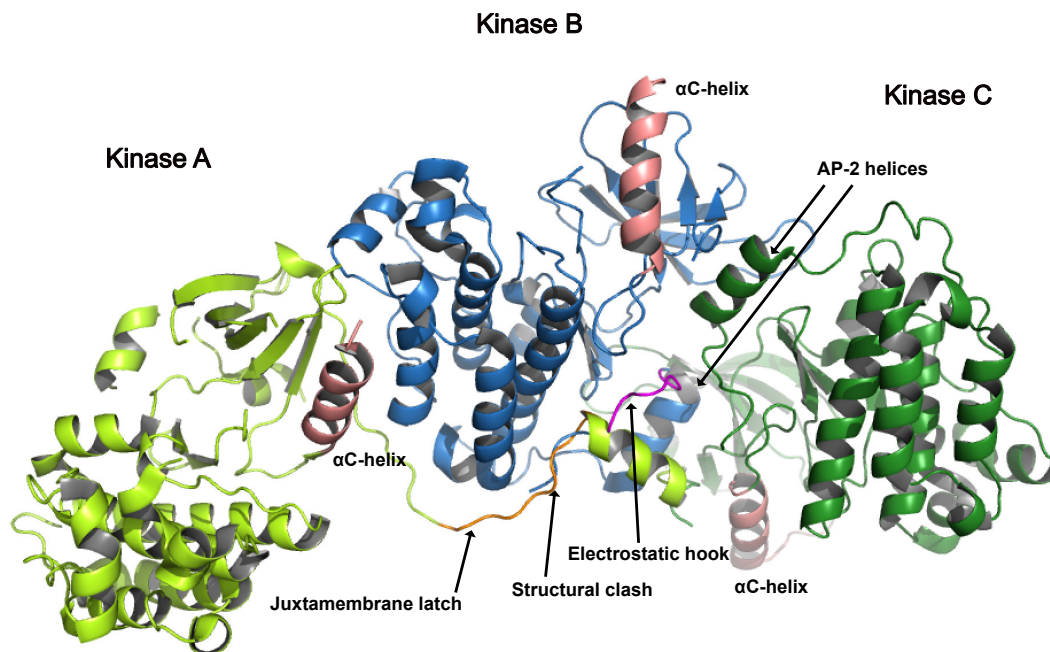
Figure 1.5 Crystal structure of EGFR kinase to reveal CT autoinhibition. Kinase B and C (PDB ID: 3GT8), containing kinase core with inactivating V924R mutation, are oriented as symmetric dimer and depicted as dark blue and dark green cartoons, respectively. Kinase A (PDB ID: 3GOP), containing ICD(645-998) with K721M mutation, is positioned relative to Kinase B following asymmetric dimer orientation and depicted as lemon green cartoon. Ligands on ATP-binding sites are omitted. αC-helices are colored pink, electrostatic hook from Kinase B magenta, and juxtamembrane latch from Kinase A orange. AP-2 helices and the structural clash between JM and CT are labeled. Figure was prepared from PDB ID: 3GT8 and 3GOP with PyMOL Molecular Graphics System.

CT hindrance represents the third EGFR autoinhibition mechanism, in which CT blocks the JM-mediated active dimerization. In a previous study, EGFR kinase bearing an deactivating mutation V924R was crystallized, which contains the key structural elements of inactive kinase described before (PDB ID: 3GT8; Jura et al., 2009). As shown in Figure 1.5, CT of kinase B folds back to kinase domain with AP-2 helix (residue 967-978; AP-2 for adaptor protein 2) approaching hinge region behind the ATP-binding site. Seven residues following AP-2-helix are dubbed as "electrostatic hook", because they are highly acidic, containing four aspartates and two glutamates, and attracted to basic residues at

the kinase surface. Residues further downstream are thereby "pinned" by electrostatic hook onto the C-lobe of kinase B, with their binding region overlapping with that of juxtamembrane latch. This structural clash with JM can hinder asymmetric dimer formation and explain the autoinhibitory effect of CT. When charge-reversing mutations were introduced in electrostatic hook, the basal phosphorylation of EGFR was enhanced (Jura et al., 2009). Meanwhile, other studies also corroborate the autoinhibitory role of CT (Bublil et al., 2010; Frederick et al., 2000; Pines et al., 2010; Wood et al., 2004). Furthermore, two inactive kinases (B and C) in Figure 1.5 are oriented in head-to-head fashion, and hence named "symmetric dimer", with AP-2 helices contributing greatly to the dimeric interface. Such a symmetric arrangement is also observed in negative-stain EM studies (Lu et al., 2012; Mi et al., 2011), and proposed as the favored dimeric form under ligand-free conditions (Arkhipov et al., 2013). Considering that preformed EGFR dimer on the cell surface has been reported before (Sako et al., 2000; Yu et al., 2002), this symmetric autoinhibited dimer could be one of the inactive forms for EGFR prior to stimulation, especially in cancer conditions, where EGFR is often overexpressed (Masuda et al., 2012; Soria et al., 2012).
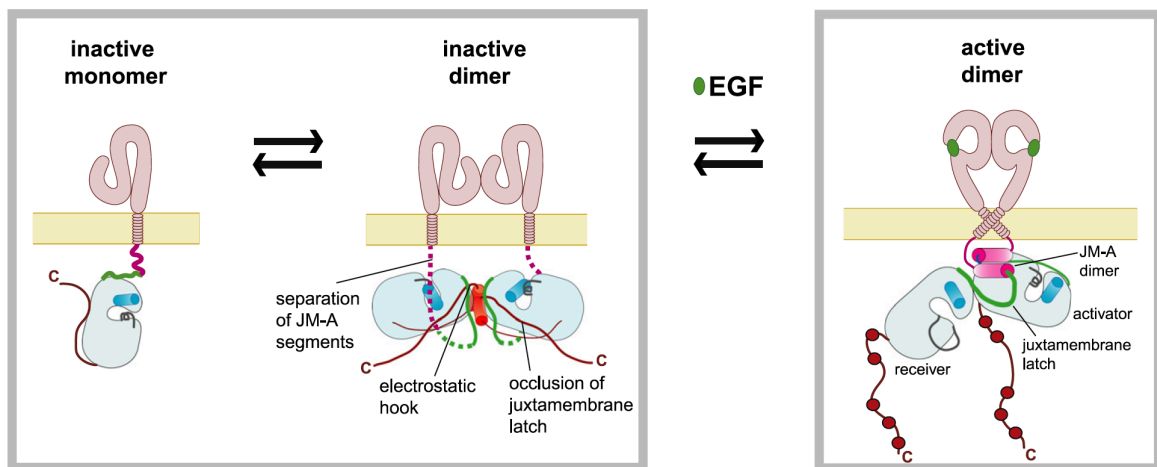
Figure 1.6 Activation scheme of EGFR. JM tethering and C-termini dimerization of TM in inactive states are not visualized. Figure is taken from Jura et al., 2009.

Combining aforementioned autoinhibition mechanisms leads to a full picture of EGFR activation. To begin with, extracellular binding of EGF gives rise to rotation of domain I and II, exposing dimerization arms and promoting ECD dimerization. Consequently, the V-shape orientation of dimerized domain IV results in N-termini dimerization of TM segments, which is coupled to JM dissociation from plasma membrane. Untethered JM facilitates rearrangement of kinase domain from either monomer or symmetric (inactive) dimer to asymmetric (active) dimer, relieving the CT autoinhibition and promoting

allosteric activation of EGFR kinase via antiparallel helical dimerization of JM-A and juxtamembrane latching of JM-B (Figure 1.6).

1.4 EGFR downregulation

Once EGFR is stimulated and activated, it initiates various signaling pathways as described in section 1.1. However, the signal transduction needs to be contained to avoid overstimulation. To accomplish this, cells have established multiple safeguarding systems to internalize EGFR, quench its signal and in the end recycle or degrade the receptor. This process, also known as EGFR downregulation, is mediated by several endocytic pathways.

First endocytosis route is suited for EGFR species sporadically activated in EGF-free conditions. In this case, GTPase Rab11 mediates EGFR uptake and returns it back to the plasma membrane via recycling endosome. This endocytic route effectively collects and recycles aberrantly activated EGFR to maintain EGFR abundance on the cell surface. (reviewed by Bakker et al., 2017; Ullrich et al., 1996). Additionally, as activated EGFR travels towards cell interior, it gradually becomes dephosphorylated due to the activity of ER-located PTP1B (protein-tyrosine phosphatase 1B; Eden et al., 2010). Thereafter, recycled EGFR resumes unphosphorylated, inactive form, ready for the next activating event. Second endocytosis route is applied to the situation where EGFR is stimulated by low concentrations of EGF and triggers clathrin-dependent endocytosis (Sigismund et al., 2005). Upon EGFR activation, AP-2 docks onto its recognition motif, *i.e.* AP-2 helix in EGFR-CT, which is sequestered before in inactive forms of EGFR (Jura et al., 2009; Sorkin and Carpenter, 1993; Sorkin et al., 1996). AP-2 attracts clathrin and thereby recruits EGFR to clathrin-coated pits, which later buds into cytosol as clathrin-coated vesicles (reviewed by Robinson, 2015). Interestingly, this budding process enriches EGFR locally and thus potentiates its autophosphorylation (Ibach et al., 2015). Furthermore, these vesicles are fused to Rab5-regulated early endosome, where EGFR is still able to maintain its signaling to downstream effectors (reviewed by Murphy et al., 2009; reviewed by Villaseñor et al., 2016; Tomshine et al., 2009; Wang et al., 2002). Therefore, both mechanisms enhance the sensitivity of EGFR in response to low-dose stimulation. EGFR, activated as such, is later dephosphorylated and returned back to cell surface through recycle endosome (reviewed by Bakker et al., 2017). Third endocytosis route takes place when high levels of EGF stimulates the receptor and leads to its degradation. In contrast to second pathway, EGFR endocytosis via third route is typically clathrin-independent and involves ubiquitination (Sigismund et al., 2005). EGFR is ubiquitinated by Cbl, which efficiently binds to highly phosphorylated EGFR via Grb2 and functions as ubiquitin-

protein ligase E3 to transfer ubiquitin from ubiquitin-conjugating enzyme E2 to EGFR (reviewed by Roskoski, 2014; Jiang, 2003; Levkowitz, 1996). Modified EGFR is transported via early endosome to Rab7-regulated late endosome, with the help of ESCRTs (endosomal sorting complexes required for transport). Thereafter, late endosome engulfs EGFR into ILV (intraluminal vesicles) to form MVB (multivesicular body). Contrary to the orientation in early endosome, EGFR-ICD faces towards the inside of ILVs, which spatially prohibits any contact with cytoplasmic effectors for further signaling. As the final step, fusion of MVB and protease-rich lysosome brings EGFR to its doom (reviewed by Bakker et al., 2017).

In addition to three endocytic pathways, other mechanisms are also involved in EGFR downregulation. Proteasome has been reported to facilitate degradation via de-ubiquitination prior to proteolysis in lysosome (Alwan et al., 2003). Furthermore, Mig6 (mitogen-induced gene 6), a negative modulator of EGFR, occupies C-lobe side of the asymmetric dimer interface and thereby inhibits kinase activity (Hackel et al., 2001; Zhang et al., 2007). In addition, it has been indicated that another segment of Mig6 inhibits EGFR activation by recognizing the binding site of juxtamembrane latch on kinase C-lobe (Jura et al., 2009).

1.5 Introduction of ARNO

ARNO (Arf nucleotide-binding site opener) belongs to cytohesin family, which includes four homologous members, *i.e.* cytohesin-1, cytohesin-2 (ARNO), cytohesin-3 (also known as Grp1 for general receptor of phosphoinositides 1) and cytohesin-4. The family name "cytohesin" is derived from its founding member cytohesin-1, a cytoplasmic activator of $\alpha L\beta 2$ integrin which mediates cell adhesion (Kolanus et al., 1996). All four family members share a common domain layout, which features a central Sec7 domain flanked by coiled-coil domain upstream and PH (pleckstrin homology) domain downstream. Sec7 domain is homologous to yeast Sec7p and it functions as GEF for Arf (ADP-ribosylation factor) GTPases (Cox et al., 2004; Meacci et al., 1997). Coiled-coil domain induces protein dimerization and interacts with other proteins for either cytohesin regulation or downstream functions (reviewed by Kolanus, 2007; DiNitto et al., 2007), while PH domain is in charge of membrane recruitment via its affinity towards phosphoinositides (Klarlund et al., 2000). Additionally, PH domain is adjacent to Sec7-PH linker at the N-terminus and a helix at the C-terminus, with C-terminal helix overlapping with the polybasic region (PBR), an autoinhibitory motif in GEF regulation (DiNitto et al., 2007).

Unlike cytohesin-1 and cytohesin-4, which are mainly distributed in leukocytes, ARNO and Grp1 are ubiquitously expressed (reviewed by Casanova, 2007), indicating their relevance to more general cellular functions. Due to their central role in Arf activation via guanine nucleotide exchange, cytohesins participate in Arf-regulated processes such as vesicular traffic and cytoskeleton remodeling (reviewed by D'Souza-Schorey and Chavrier, 2006; reviewed by Kolanus, 2007). Furthermore, according to a previous study, an aptamer recognizing coiled-coil and Sec7 domain of ARNO effectively inhibits the serum-induced transcriptional activation and ERK1/2 activity (Theis et al., 2004). In another study, SecinH3, an Sec7-targeting inhibitor developed from aptamer displacement screen, induces insulin resistance in human liver cells (Hafner et al., 2006). Therefore, these studies point to the involvement of ARNO in both ERK activation and insulin signaling. Additionally, ARNO has been reported to be involved in G protein-independent signaling of GPCR (G protein-coupled receptor) via direct binding (Gsandtner et al., 2005), regulation of vesicle transport via interaction with V-ATPase (Hurtado-Lorenzo et al., 2006), and crosstalk to receptor signaling via binding to adaptor proteins such as cytohesin interacting protein (Cytip; Boehm et al., 2003). Altogether, these examples from various perspectives prove the versatile role played by ARNO in cells.

1.6 Guanine nucleotide exchange mechanism of ARNO

As the core of ARNO, Sec7 domain possesses guanine nucleotide exchange activity for Arf, a group of GTPases from Ras superfamily. Arf is renowned for its role in vesicular traffic via recruiting coating complex onto the intracellular vesicles or activating lipid-regulating enzymes (reviewed by Casanova, 2007), and it functions as a molecular switch, cycling between GDP-bound inactive form and GTP-bound active form. GEF accelerates substitution of GTP for GDP to turn on the Arf activity, while GAP (GTPase-activating protein) stimulates the intrinsic GTPase activity of Arf, promoting hydrolysis of bound GTP and turning off its activity. Arf activation can be divided into GDP-releasing phase and GTP-binding phase. ARNO stimulates GDP release in the first phase so that new nucleotide can bind thereafter. In the second phase, GTP is preferred to enter the empty nucleotide-binding site, since cellular concentration of GTP is much higher than that of GDP (reviewed by Traut, 1994).

ARNO accelerates the first phase by employing a conserved glutamate (dubbed as "glutamic finger") from Sec7 domain to approach and destabilize Mg^{2+}:GDP via electrostatic force (Béraud-Dufour et al., 1998). Meanwhile, ARNO also induces a structural movement named "interswitch toggle" in Arf, preparing it for GTP binding (Renault et al., 2003). Notably, "interswitch toggle" requires the proximity of cellular

membranes as well, due to the amphipathic N-terminal helix of Arf and the myristoylated glycine residue on it (reviewed by Jackson and Casanova, 2000). In the inactive form, cytosolic Arf has N-terminal helix folded back with hydrophobic residues and myristoyl group buried, whereas in the presence of ARNO, Arf extends and inserts N-terminal helix into the lipid bilayer, rendering Arf tethered to the membrane, which in turn stabilizes this conformation by sequestering those hydrophobic moieties on N-terminal helix (Béraud-Dufour et al., 1999; Goldberg, 1998). Therefore, this structural transition of "interswitch toggle" requires synergistic efforts from both ARNO and lipid bilayer.

As the product of GDP-releasing phase, the nucleotide-free Sec7:Arf complex is relatively stable, in contrast to the Sec7:Arf:Mg^{2+}:GDP intermediate. However, the latter is more informative and therefore motivates crystallization endeavor for capturing this key structure (Mossessova et al., 2003). To achieve this, brefeldin A (BFA), a fungal macrolide targeting Sec7:Arf complex, was introduced to structural studies. To be exact, BFA binds to the intermolecular interface between Sec7 and Arf to prohibit GDP release (Cherfils and Melancon, 2005). This abortive complex is like a snapshot of the dynamic exchange process, which allows structural basis of nucleotide release to be inspected. However, members of cytohesin family are BFA-insensitive, unlike other Arf GEFs such as BIG1 (Brefeldin A-inhibited guanine nucleotide-exchange protein 1; reviewed by Kolanus, 2007). Therefore, in a crystallization study of ARNO-Sec7:Arf1 complex, BFA-sensitizing mutations were introduced into ARNO-Sec7 to enable BFA binding (Renault et al., 2003).

As can be seen in Figure 1.7, Sec7 consists of 10 α-helices (αA-J according to Mossessova et al., 1998), with αF-H forming a hydrophobic groove. The major Sec7-contacting regions in Arf1 are switch 1 (residue 47-53) and switch 2 (residue 73-81), connected by a interswitch region spanning β2- and β3-strands (nomenclature following Goldberg, 1998). Rest of Arf1 structure is the core region, which is rigid compared to switch and interswitch regions. In this structure trapped by BFA, Arf1 is "straddling" Sec7, with one leg of switch 1 reaching into hydrophobic groove and the other leg of switch 2 contacting αH-helix near the protein C-terminus. Hydrophobic residues I49, F51, I74 and L77 from both switch regions contribute to the docking of Arf1, so does an electrostatic attraction between K73 of Arf1 and D183 of Sec7 (Béraud-Dufour et al., 1998; Renault et al., 2003). Meanwhile, BFA resides at Sec7:Arf1 interface, caged in a hydrophobic cavity composed of residues from both proteins. Interestingly, glutamic finger (E156 at the beginning of αG-helix) is not close to Mg^{2+}:GDP, with a distance of 8.2Å between terminal carboxyl carbon and Mg^{2+}, which explains the inhibitory effect of BFA on nucleotide release. Notably, this BFA-captured structure shows merely the initial docking form of Arf1:Sec7. Another crystal of Arf1 in complex with ARNO-Sec7(E156K) in the absence of

BFA reveals the pre-dissociation form, in which "interswitch toggle" has already taken place (PDB ID: 1R8S; Renault et al., 2003). β2-/β3-strands and the connecting λ3 loop extrude from Arf1 core, impairing the binding site for N-terminal helix, so that it will shift away from the Arf1 and be ready for membrane insertion (Goldberg, 1998). More importantly, the hydrophobic cavity harboring BFA collapses in pre-dissociation form, which results in Arf1 core leaning towards Sec7 and thereby brings Mg^{2+}:GDP closer to E156. Thus, docking form and pre-dissociation form, together with nucleotide-free form, represent three key steps in GDP-releasing procedure, in which Sec7 mediates structural rearrangement of Arf1 not just to expose the N-terminal helix for membrane anchoring, but also to drive GDP into closer proximity of glutamic finger before GDP being repelled and pushed off Arf1 (Renault et al., 2003).
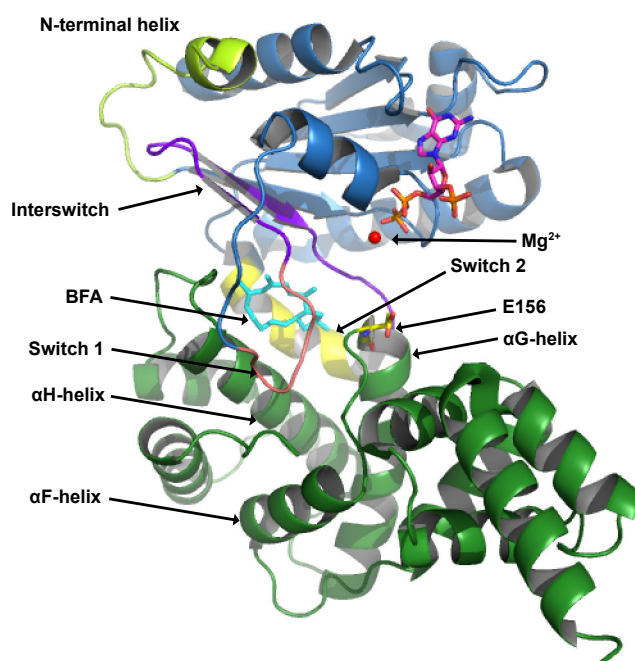
Figure 1.7 Crystal structure of ARNO-Sec7:Arf1:BFA complex. Full length Arf1 and ARNO-Sec7(50-252) bearing BFA-sensitizing mutations (F190Y/A191S/S198D/P208M) are depicted as dark blue and dark green cartoons, respectively. Mg^{2+} ion is depicted as a red sphere, while glutamic finger and G3D are labeled and shown as sticks with carbons colored yellow and magenta, respectively. BFA is labeled and shown as sticks colored cyan, while αF-H are labeled. N-terminal helix, switch 1, interswitch and switch 2 are colored lemon green, orange, purple and yellow, respectively. G3D stands for guanosine-3'-monophosphate-5'-diphosphate, which is from *E.coli* overexpression of Arf1. Figure was prepared from PDB ID: 1R8Q (Renault et al., 2003) with PyMOL Molecular Graphics System.

1.7 Activation of cytohesin

Autoinhibition mechanism of cytohesins has been established, based on previous structural study of Grp1 (cytohesin-3), in which Sec7-PH linker and C-terminal helix are located in the proximity of hydrophobic groove of Sec7 and interfere with Arf binding. Additionally, deletion of PBR restores the cytohesin activities to different extents, suggesting its involvement in autoinhibition as well (DiNitto et al., 2007).

Concerning the activation mechanism to relieve cytohesin autoinhibition, Arf6 has been unexpectedly reported as one candidate activator for cytohesin (Cohen et al., 2007). In contrast to Golgi-located Arf1, Arf6 is mainly distributed on plasma membrane and endoplasmic vesicle network, similar to cytohesins (D'Souza-Schorey and Chavrier, 2006). This colocalization of Arf6 and cytohesins inspired the aforementioned cellular study, which proposes an activation mechanism that active GTP-bound Arf6 recruits ARNO or Grp1 to the plasma membrane and activates these GEFs in the presence of phosphoinositides. Instead of just being a substrate of cytohesins, Arf6 could also activate them by interacting with PH domain (Cohen et al., 2007). This allosteric activation mechanism was also confirmed *in vitro* via PIP2-containing liposomes (Stalder et al., 2011).

To unveil the structural basis of PH:Arf6 interaction, structure of Grp1-PH in complex with Arf6 was determined (Malaby et al., 2013). In this structure, Mg^{2+}:GTP is bound to NΔ13Arf6(Q67L) which is constitutively active and lacking the N-terminal helix, while the head group of PIP3, *i.e.* IP4 (inositol 1,3,4,5-tetrakisphosphate), is bound to Grp1(247-399) spanning Sec7-PH linker, PH domain and C-terminal helix. As can be seen in Figure 1.8, interaction is primarily mediated by switch and interswitch regions from active Arf6 and β3-/β4-/βi1-/βi2-strands from PH domain. Meanwhile, Sec7-PH linker and C-terminal helix flank the core interface by contacting switch 2 and switch1, respectively. Notably, β3-/β4-/βi1-/βi2-strands of PH domain also interact with IP4. Therefore, structural changes in these regions, due to binding of highly charged IP4 (or PIP3), could affect Arf6 docking, which may explain the inositide-dependence of Arf6 allosteric activation (DiNitto et al., 2007; Lietzke et al., 2000; Malaby et al., 2013). Interestingly, IP4 is on the same side of PH:Arf6 complex as the N-terminus of active Arf6, as shown in Figure 1.8 (B). Since PH is typically anchored to the membrane via phosphoinositides, while Arf GTPases via myristoylated N-terminal helix, this observation could explain how PH:Arf6 complex is oriented on the plasma membrane and thereby add more credit to the crystal structure.
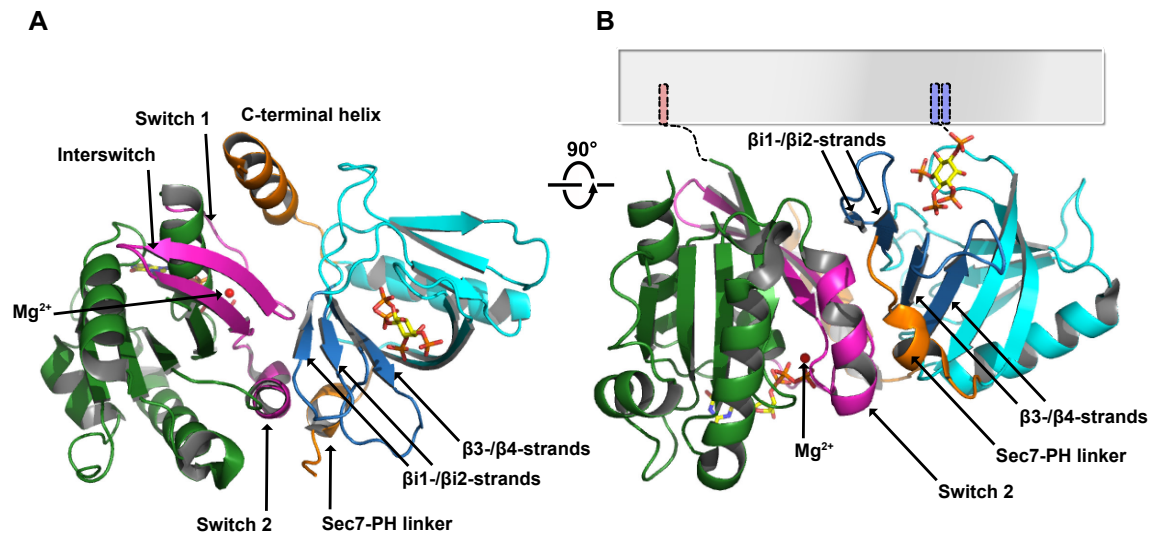
Figure 1.8 Crystal structure of Arf6:PH complex. (B) is acquired from 90° rotation of (A) as indicated. NΔ13Arf6(Q67L) and PH-containing Grp1(247-399) are depicted as dark green and cyan cartoons, respectively. Mg^{2+} ion is depicted as a red sphere, while GTP and IP4 are shown as sticks with carbons colored yellow. Switch 1, interswitch and switch 2 of Arf6 are colored magenta; β3-/β4-/βi1-/βi2-strands of PH are colored dark blue; Sec7-PH linker and C-terminal helix are colored orange. In (B), plasma membrane is schematized as a grey box, while PIP3 and N-terminal helix are completed with dashed lines, with myristoyl group and acyl groups of PIP3 highlighted as light red and light purple sticks, respectively. Nomenclature of secondary structures for both proteins follows Malaby et al., 2013. Figure was prepared from PDB ID: 4KAX (Malaby et al., 2013) with PyMOL Molecular Graphics System.

Comparison of Arf6-bound PH with PH from autoinhibited Grp1 sheds light on the cytohesin activation mechanism (Malaby et al., 2013). As shown in Figure 1.9 (A), PH domains in both cases are essentially identical, whereas Sec7-PH linker and C-terminal helix shift drastically upon binding of active Arf6. With both autoinhibitory elements swung away from hydrophobic groove and catalytic glutamic finger, the activating role of GTP-bound Arf6 could be partially explained. Yet with the bulky PH domain nearby, it is still unclear how Sec7 domain is reoriented to fully clear the hydrophobic groove for Arf substrates docking. One recent study, combining small-angle X-ray scattering and negative-stain EM, as well as rigid-body and ensemble analyses, reveals the structural flexibility of Grp1 hinge regions in solution, which are the residues connecting PH domain to Sec7-PH linker and C-terminal helix. According to this study, the swinging movement of C-terminal helix due to hinge flexibility brings it from autoinhibitory position to the proximity of PH domain, which leads to the free sampling of Sec7-PH linker in configuration space. Due to this linker flexibility, Sec7 is repositioned so that its catalytic site is entirely exposed and ready for accommodating substrate Arf (Malaby et al., 2018), as shown in Figure 1.9 (B).
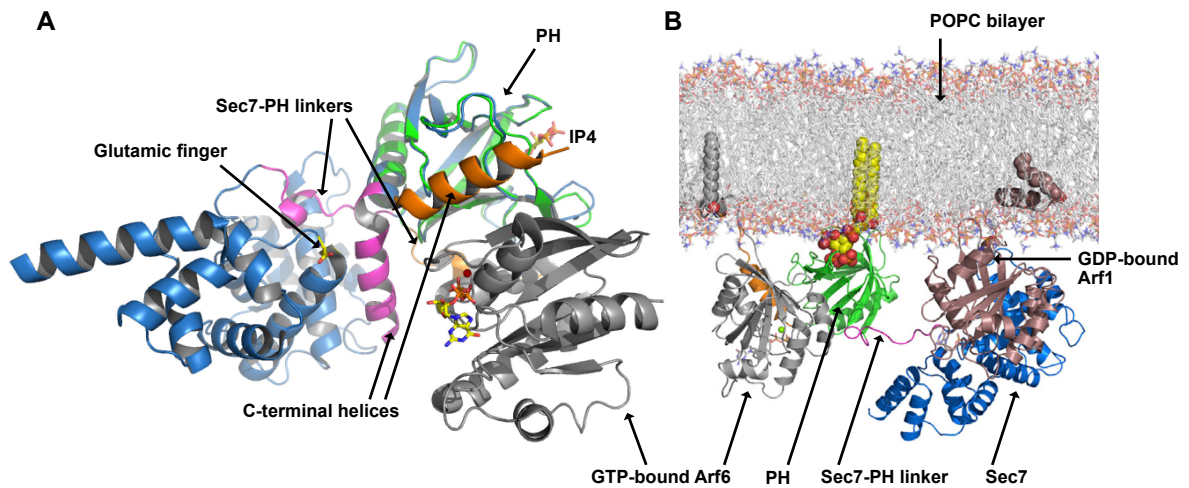
Figure 1.9 Activation mechanism of cytohesins. (A) Arf6:PH complex from Figure 1.8 and Grp1(63-399) are overlaid based on superimposed PH domains. Grp1(63-399) carries K68A/H260Y mutations to achieve better diffraction without impairing exchange activity. Grp1(63-399) is depicted as dark blue cartoon with His-tag omitted, while Arf6 and PH from Arf6:PH complex are depicted as grey and green cartoons, respectively. Glutamic finger, IP4 and GTP are shown as sticks with carbons colored yellow. Sec7-PH linker and C-terminal helix are colored magenta in autoinhibited state, and orange in Arf6-bound state. Note the swinging movements of both elements. (B) MultiFoXS model of Arf6-activated Grp1 in complex with Arf1 substrate. Orientation of GDP-bound Arf1 is based on Sec7:Arf1 complex in PDB ID: 1R8S. Proteins are depicted as cartoons, while PIP3 on PH and myristoyl groups on Arf1/Arf6 are depicted as space-filling models. Figure (A) was prepared from PDB ID: 2R09 (DiNitto et al., 2007) and 4KAX with PyMOL Molecular Graphics System. Figure (B) was adapted from Malaby et al., 2018.

2 Research Aims

According to the prevalent activation model of EGFR established by John Kuriyan lab, EGF binding triggers a chain of structural rearrangements which leads to allosteric activation of kinase domain (Arkhipov et al., 2013; Endres et al., 2013; Jura et al., 2009; Zhang et al., 2006). In this process, conformation equilibrium of EGFR kinase is shifted from inactive monomer/dimer towards active dimer (Jura et al., 2009). Yet the structural studies, which this activation model is based upon and supported by, have their own limitations. Firstly, the active (asymmetric) dimer structure was originally determined in crystallization studies (Stamos et al., 2002; Zhang et al., 2006), using proteins lacking key activating elements such as juxtamembrane latch and JM-A antiparallel helices. The same asymmetric dimer was also reported in catalytically dead EGFR kinase (Red Brewer et al., 2009). This counter-intuitive observation could be caused by crystal packing effect due to high protein concentrations. Furthermore, the inactive (symmetric) dimer is derived from crystal structure of an EGFR mutant that purposefully impairs active dimer interface to force protein into its inactive form (Jura et al., 2009). Considering the central roles of asymmetric/symmetric dimers in EGFR activation model, it should be addressed first whether these two dimeric forms are actually present in dilute solution using unmutated EGFR, and if so, how frequently they are sampled by EGFR molecules. Secondly, both dimeric forms have been independently visualized via negative-stain electron microscopy (Lu et al., 2012; Mi et al., 2011). However, these images are "coarse-grained" and more evidence at molecular level is still urgently needed to support the activation model. As one of the emerging structural biology tools, electron paramagnetic resonance (EPR) promises protein sample analysis in solution or in detergent-/lipid-solubilized forms. Especially, pulsed electron-electron double resonance (PELDOR), one of the EPR methods, enables high-precision structural analysis of spin-labeled proteins by measuring interspin distance and deducing conformational dynamics from distance distribution. Therefore, the first part of this thesis aims at analyzing conformation distribution of dimeric EGFR kinases in solution and detergent/lipid micelles through collaboration with an EPR lab headed by Prof. Olav Schiemann.

Another important aspect of EGFR activation is regulation by cytoplasmic modulators. EGFR activation is so important for cellular functions that apart from self-regulation via autoinhibition mechanism, it is also regulated by several cellular modulator proteins, which either interact directly with JM (calmodulin and TRAF4; Aifa et al., 2002; Cai et al., 2018; Martín-Nieto and Villalobo, 1998) or mimic binding mode of JM to the kinase core (Mig6;

Jura et al., 2009), underscoring the critical role of JM in EGFR regulation. Meanwhile, emerging evidence suggests ARNO (cytohesin-2) as a potential EGFR activator. One previous study has shown the correlation between ARNO overexpression and EGFR signaling in human colorectal cancer, with chemical inhibition of ARNO by SecinH3 interfering with cell migration and proliferation (Pan et al., 2014). Another study shows that SecinH3 treatment effectively reduces EGFR signaling and proliferation of lung cancer cells (Bill et al., 2012). Connection between ARNO and EGFR in functional studies leads to the question of whether these two proteins directly interact, and if so, where the binding sites would be. Therefore, the second part of this thesis aims to investigate interaction between EGFR-ICD and ARNO, using established method of microscale thermophoresis (MST). Once interaction is confirmed, further MST assays will narrow down recognition motifs to protein domains or segments. Next, through collaboration with an NMR lab led by Dr. Manuel Etzkorn, EGFR-ICD:ARNO interaction shall be studied on amino acid level, using the domains identified by MST to characterize residues that are involved in this interaction.

3 Results

3.1 Conformation study of EGFR

3.1.1 Purification and labeling of constructs for conformation study

All the constructs investigated in this study are illustrated in Figure 3.1, together with a domain sketch of full length EGFR on the top. Most of the constructs were designed with MBP-tag at the N-terminus as a stabilizer during purification and labeling. Four among them contain EGFR-ICD truncated at the C-terminal border of kinase core (residue 998), which is the common practice for many structural and functional studies (Jura et al., 2009; Lu et al., 2012; Mi et al., 2008, 2011; Stamos et al., 2002; Zhang et al., 2006). Only one MBP-tagged construct, MBPT-ICD, includes complete intracellular domain of EGFR. Additionally, the SBP-tagged construct EGFRΔC contains entire extracellular and transmembrane domains, as well as intracellular domain truncated at residue 998, with SBP-tag fused to the C-terminus.
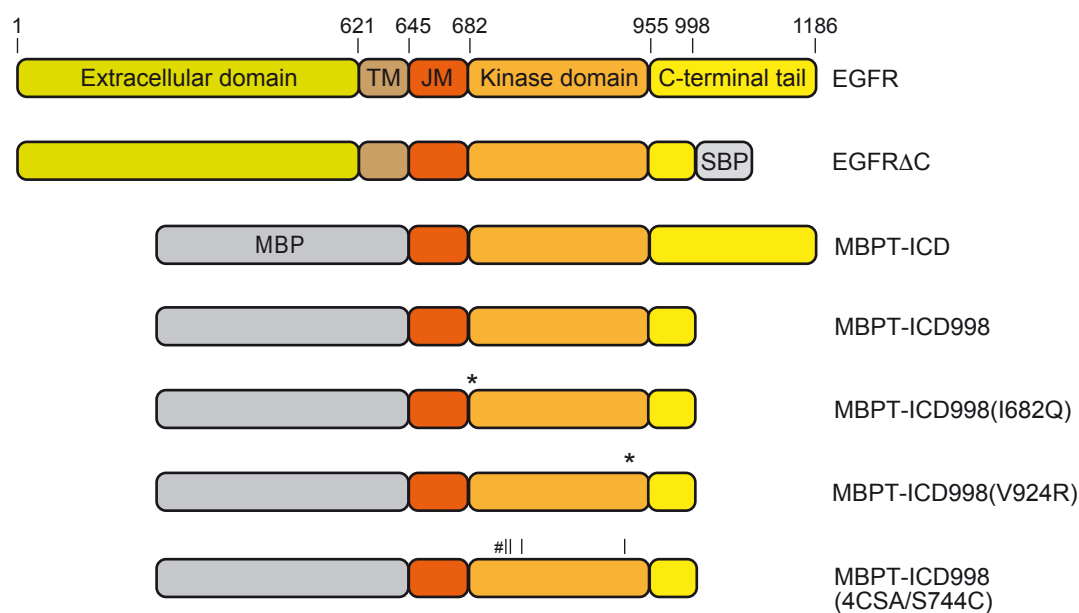
Figure 3.1 Domain sketch of full length EGFR and all the constructs used in conformation study. Sequence numbering is based on human EGFR without signal peptide. Mutations in MBPT-ICD998(I682Q) and MBPT-ICD998(V924R) are denoted by asterisks. In MBPT-ICD998(4CSA/S744C), 4CSA mutations are indicated by vertical bars while S744C by hash sign.

Proteins were expressed in Sf9 cells and large-scale purification coupled with labeling procedure was performed as described in Methods section. Briefly, MBP-tagged constructs were purified via amylose affinity chromatography, before labeling and size

exclusion chromatography (SEC). Collected fractions were pooled and concentrated, followed by buffer exchange into D_2O-containing buffer in centrifugal filters. As for membrane protein EGFRΔC, all purification buffers were supplemented with Triton® X-100. Strep-Tactin® affinity chromatography was performed first, with sample eluted in D_2O-containing buffer, after which sample was concentrated and labeled. Labeling strategy for each construct will be introduced in following sections.

Purified final products were visualized via SDS-PAGE and Coomassie staining. As shown in Figure 3.2, protein samples are relatively pure, except that MBPT-ICD is contaminated by ladder-patterned bands, which are probably protein degradation products.
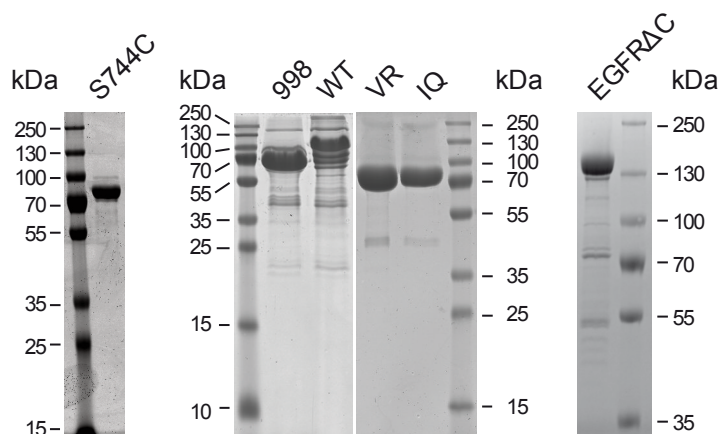
Figure 3.2 Final products of labeled protein for conformation study. 50 pmol of concentrated samples were resolved in SDS-PAGE. From left to right, the separating gels are 10%, 12.5%, 10% and 8%, respectively. S744C: MBPT-ICD998(4CSA/S744C), 998: MBPT-ICD998, WT: MBPT-ICD, VR: MBPT-ICD998(V924R), IQ: MBPT-ICD998(I682Q).

3.1.2 Design and characterization of an MTSL-labeled construct

In order to investigate EGFR-ICD conformation in solution, PELDOR spectroscopy was employed, which is one of the EPR spectroscopy methods. Briefly, PELDOR is a convenient, sensitive, low sample-consuming spectroscopic approach, in which diamagnetic protein is labeled by paramagnetic chemicals (also known as spin labels) and PELDOR acts as a molecular ruler to measure distance between paramagnetic centers. Acquired distance distribution is used to deduce information about protein conformation and structural dynamics. Depending on research interest, both intramolecular and intermolecular distances are measurable.

Conventional protein PELDOR study utilizes nitroxide spin label such as MTSL (*S*-(1-oxyl-2,2,5,5-tetramethyl-2,5-dihydro-1H-pyrrol-3-yl)methyl methanesulfonothioate), and the

labeling process relies upon reactivity of thiosulfate ester towards sulfhydryl group in the side chain of cysteine (Figure 3.3).
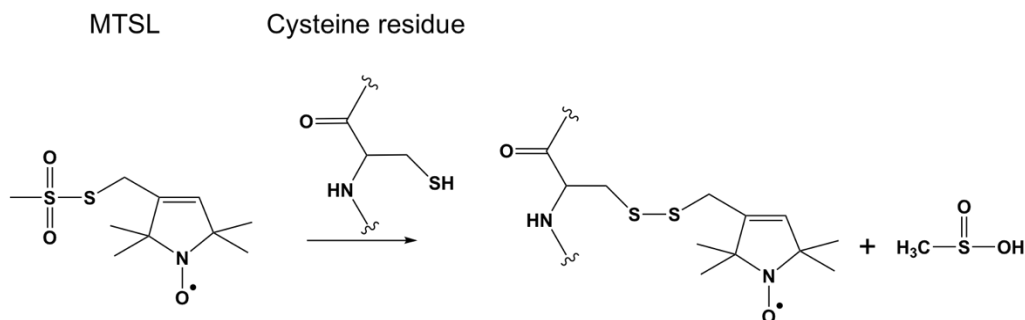
Figure 3.3 Reaction of MTSL with a cysteine residue of the protein

However, native cysteines of EGFR are unsuited for the labeling, because they do not meet the candidate criteria described below. By site-directed mutagenesis, solvent-exposed endogenous cysteines were substituted with either serine or alanine depending on hydrophobicity around the residue, after which new cysteine was introduced at proper amino acid site for MTSL labeling. In the first round of mutation, 4CSA combination (C751S/C757A/C773S/C915A) was designed. According to crystal structures of active and inactive kinase domain of EGFR (PDB ID: 2GS6 and 3GT8), native cysteines C794 and C926 face inside of the helix bundle in C-lobe, and therefore remain unmutated since they are not accessible to MTSL. In the second round of mutation, only one cysteine was introduced into ICD aimed at measuring intermolecular distance, considering that research goal is to study EGFR-ICD conformation, or to be exact, the distribution of the different possible conformations of dimeric EGFR in solution. Proper amino acid site for the new cysteine was selected based on several criteria. First, candidate locations have to be solvent-exposed in both asymmetric and symmetric ICD dimers. In this thesis project, structural models of both dimers are taken from a previous study (Arkhipov et al., 2013), in which dimeric near-full-length EGFR in lipid bilayer was simulated, with both kinase dimer models being based on aforementioned structures (PDB ID: 2GS6 and 3GT8). Second, in both asymmetric and symmetric dimers, distances between candidate locations should be within optimal PELDOR measurement range, which is approximately 18 to $80\,\text{Å}$ (reviewed by Jeschke and Polyhach, 2007; reviewed by Schiemann and Prisner, 2007; reviewed by Tsvetkov et al., 2008). Third, aforementioned distances should be different between asymmetric and symmetric dimers, so that these two conformations are distinguishable via PELDOR.

Candidate construct MBPT-ICD998(4CSA/S744C), which meets all the selection criteria, was expressed, purified and labeled as documented in Methods section, with final sample

visualized via SDS-PAGE (Figure 3.2). Before starting PELDOR measurement, autophosphorylation assay was also performed to examine protein activity. The intracellular domain (ICD) of EGFR harbors at its core a kinase domain which, upon activation, transfers the γ-phosphate group of ATP to one of the tyrosines in C-terminal tail. Thus, phosphorylation of tyrosine can be utilized as indicator of kinase activity. The EGFR autophosphorylation has such physiological significance that plenty of antibodies have been commercialized for immunodetection, targeting generic or specific phosphotyrosine. Here an antibody targeting phosphotyrosine 992 was utilized in western blot. As shown in Figure 3.4, MBPT-ICD998(4CSA/S744C) is unfortunately inactive, with no increased phosphorylation at 1 or 3min. The loss of activity could be due to multiple mutations or MTSL labeling. Without evidence to prove that protein is still functional, this study cannot proceed to PELDOR measurement. New labeling strategy is needed.
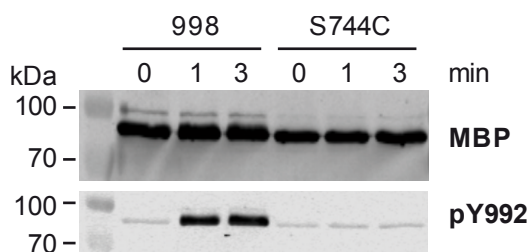
Figure 3.4 Autophosphorylation assay of MTSL-labeled MBPT-ICD998(4CSA/S744C). Detailed assay procedure is provided in Methods section. 1.4 pmol of phosphorylated samples stopped at indicated time points were visualized via western blot. Total EGFR intensities were detected by anti-MBP antibody on MBP blot, and phosphorylated Y992 by anti-phosphotyrosine antibody on pY992 blot. 998: MBPT-ICD998, S744C: MBPT-ICD998(4CSA/S744C) labeled by MTSL.

3.1.3 Design and characterization of PDsl

3.1.3.1 Introduction of IDSL and PDsl

The following sections will focus on a new protein labeling strategy, named inhibitor-directed spin labeling (IDSL). This strategy takes advantage of the excellent binding affinity of certain protein inhibitors, which are covalently attached to spin label moieties to suit measurement purposes. Once protein is bound to such modified inhibitor, it becomes spin-labeled. One major benefit of this method is that no mutation is required, which means protein structure information is extracted under more natural conditions. Some inhibitors even covalently bind to the target protein, so purification steps after labeling will not wash them off. One such inhibitor, PD168393 developed by Parke, Davis & Co. (now part of Pfizer) was used in this study and it targets EGFR kinase domain. After occupying

ATP-binding cleft, PD168393 forms a covalent bond with EGFR via Michael addition between its acrylamide warhead and sulfhydryl group of C773 nearby (Figure 3.5).

IC_{50} of PD168393 from publications is concluded as either < 1nM (Blair et al., 2007) or 0.08nM (Michalczyk et al., 2008), which is convenient for the labeling process. In a previous study, PD168393 was coupled to NBD fluorophore via PEG at C7 of the quinazoline backbone, and was successfully applied to cellular study as a fluorescence probe. Additionally, ICD(672-998) was already cocrystallized with PD168393, which provides useful binding detail of this inhibitor on ATP-binding site (PDB ID: 2J5F; Blair et al., 2007). Furthermore, one molecule of EGFR-ICD binds to only one copy of PD168393, which is especially suitable for conformation studies on the EGFR dimer.
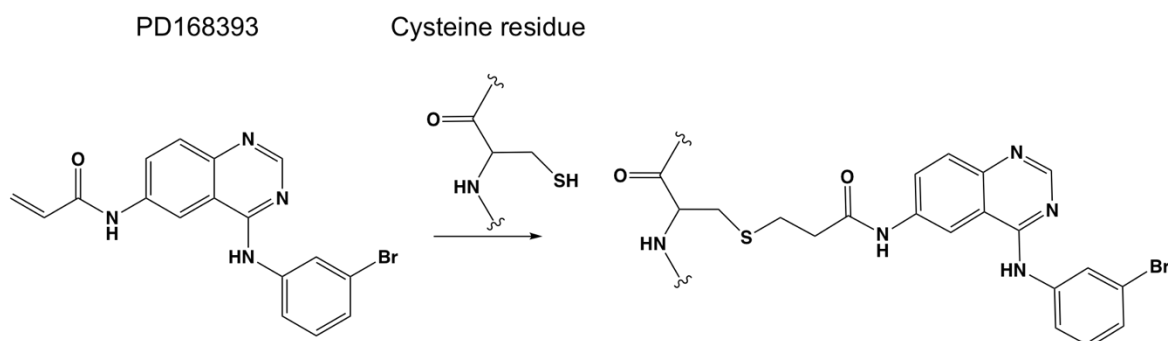
Figure 3.5 Reaction of PD168393 with cysteine residue in the protein
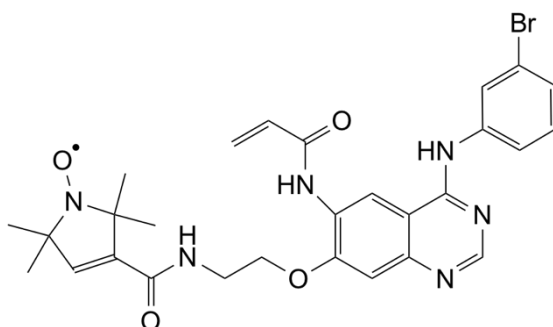
Figure 3.6 Structure of PD168393sl

In the current PELDOR study, an analogous strategy was adopted for probe design, in which PD168393 was linked to nitroxide spin label at C7 of the quinazoline backbone. A linker shorter than PEG, ethanolamine, was used since it increases steric rigidity between the spin label moiety and inhibitor moiety, thus leading to narrower peaks in distance distribution graph. This EPR probe was named PD168393sl (PDsl for short), with structure shown in Figure 3.6. Based on ICD-PD168393 cocrystal structure (PDB ID: 2J5F) and other simulation models (Arkhipov et al., 2013), spin-spin distances in both asymmetric

and symmetric EGFR dimers were predicted using a PyMOL plugin mtsslWizard (Hagelueken et al., 2012), and they lie within PELDOR measurement range. More details about the calculation will be provided in section 3.1.4. The probe was kindly synthesized and characterized in detail by Jeffrey Hannam (Yin et al., 2017).

3.1.3.2 Activity assay comparing PD168393 & PDsl

Since PD168393 is an ATP-competitive kinase inhibitor, a natural approach to confirm protein labeling is to visualize inhibited autophosphorylation of EGFR via western blot. Also, it is unclear whether attached spin label on PDsl influences its inhibitory effect, and this issue can be resolved via the same method.

Autophosphorylation assay and western blot were performed as described in Methods section. For each sample at indicated time point, the ratio between phosphorylation and EGFR intensities was calculated, after which these ratios were normalized by the value at 0 min, giving relative phosphorylation (rel. phos.). As shown in Figure 3.7, PD168393 and PDsl inhibit ICD autophosphorylation to the same extent, with approximately 6.7% relative phosphorylation intensity left, compared to unlabeled protein at both 1 and 3min. To conclude, PDsl effectively binds to EGFR-ICD, and attached spin label has no influence over its binding.
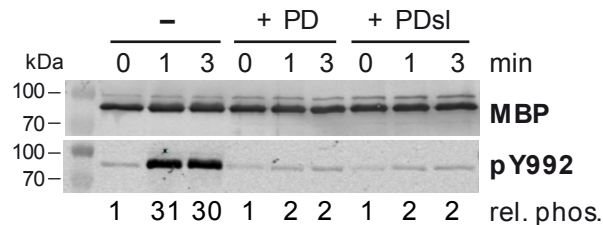
Figure 3.7 Autophosphorylation assay comparing PD168393- and PDsl-labeled MBPT-ICD998. Detailed assay procedure is provided in Methods section. 1.4 pmol of phosphorylated samples stopped at indicated time points were visualized via western blot. Total EGFR intensities were detected by anti-MBP antibody on MBP blot, and phosphorylated Y992 by anti-phosphotyrosine antibody on pY992 blot. "-": unlabeled protein, "+PD": PD168393-labeled, "+PDsl": PDsl-labeled, rel.phos.: relative phosphorylation.

3.1.3.3 Activity assay of PDsl-labeled constructs

The purpose of this section is to fully examine PDsl labeling of all the constructs in EPR project, using the same methodology as in previous section, *i.e.* autophosphorylation assay.
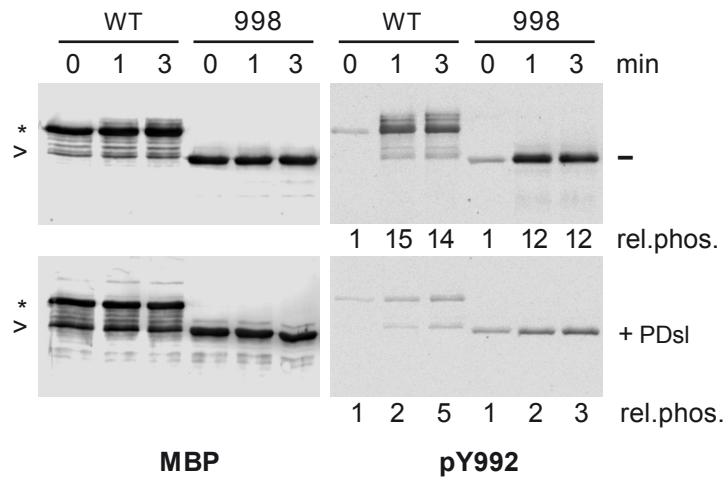
Figure 3.8 Autophosphorylation assay of PDsl-labeled constructs. Detailed assay procedure is provided in Methods section. 1.5 pmol of phosphorylated samples stopped at indicated time points were visualized via western blot. Total EGFR intensities were detected by anti-MBP antibody on MBP blot, and phosphorylated Y992 by anti-phosphotyrosine antibody on pY992 blot. "-": unlabeled protein, "+PDsl": PDsl-labeled protein, WT and *: MBPT-ICD, 998 and >: MBPT-ICD998, rel.phos.: relative phosphorylation.
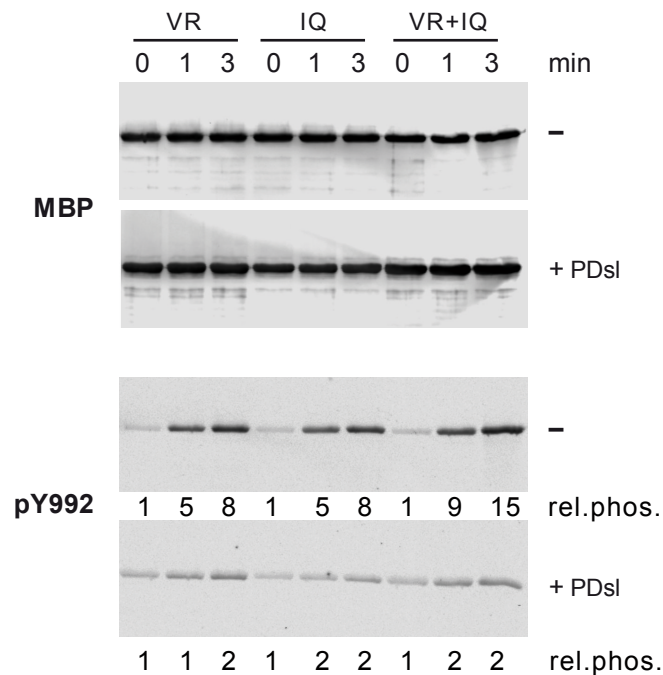
Figure 3.9 Autophosphorylation assay of PDsl-labeled constructs. Detailed assay procedure is provided in Methods section. 1.5 pmol of phosphorylated samples stopped at indicated time points were visualized via western blot. Total EGFR intensities were detected by anti-MBP antibody on MBP blot, and phosphorylated Y992 by anti-phosphotyrosine antibody on pY992 blot. "-": unlabeled protein, "+PDsl": PDsl-labeled protein, VR: MBPT-ICD998(V924R), IQ: MBPT-ICD998(I682Q), rel.phos.: relative phosphorylation.

In Figure 3.8, assay results for MBPT-ICD and MBPT-ICD998 were visualized via western blot, with relative phosphorylation intensities calculated and denoted under sample lanes. One can clearly see that PDsl inhibits autophosphorylation of these two constructs. Notably, contaminations in MBPT-ICD, visualized by SDS-PAGE in Figure 3.2, are recognized by anti-MBP antibody. In conjunction with the fact that they were copurified based on the binding of N-terminal MBP-tag to amylose resin, these additional bands probably represent degraded proteins losing parts of EGFR C-terminal tail.

In previous studies, V924R and I682Q mutations were introduced at asymmetric ICD dimer interface on the C-lobe and N-lobe side, respectively (Jura et al., 2009; Zhang et al., 2006). Since asymmetric dimer is associated with active kinase conformation (reviewed by Bose and Zhang, 2009), either mutation alone impairs phosphorylation activity of EGFR by hindering asymmetric dimer formation. However, V924R still retains an intact N-lobe and I682Q an intact C-lobe. When these two combine to form a complete asymmetric dimer interface, kinase activity is partially restored. In the current study, MBPT-ICD998(V924R) and MBPT-ICD998(I682Q) were designed for EPR investigation. Judged by relative phosphorylation in Figure 3.9, both constructs show certain level of kinase activity, and activity increases after combining them in "VR+IQ". However, the contrast between activities of "VR+IQ" and "VR" or "IQ" is less pronounced, compared to previous studies (Jura et al., 2009; Zhang et al., 2006). This could be explained by the different methodologies adopted in activity assays, *i.e.* western blot of autophosphorylation at Y992 here *vs.* enzyme-coupled spectrophotometric assay of substrate phosphorylation in aforementioned publications. Nevertheless, inhibited autophosphorylation is observed in PDsl-treated group (Figure 3.9), which indicates proteins being labeled. Noteworthy, autophosphorylation assay is not intended to quantitatively measure labeling efficiency of PDsl. To achieve that, a more direct approach named spin counting was adopted (see section 3.1.4).

Another construct included in EPR study is EGFRΔC, which contains extracellular, transmembrane and intracellular domain truncated at G998, as well as SBP-tag fused at C-terminus. Large-scale purification and labeling process is documented in Methods section. For autophosphorylation assay, unlabeled EGFRΔC was pre-incubated with PDsl and EGF in 1: 1: 1.2 molar ratio on ice for 10min. As controls, EGF was replaced by equal volume of water (-EGF), or PDsl was replaced by equal volume of DMSO (-). Reaction was initiated by adding 1mM ATP and 5mM $MgCl_2$, and terminated by adding SDS-PAGE loading buffer with 25mM EDTA. Samples were analyzed via western blot, which shows clearly that PDsl blocks EGF-stimulated as well as basal EGF-independent kinase activity (Figure 3.10). Thus EGFRΔC is also effectively labeled by PDsl.
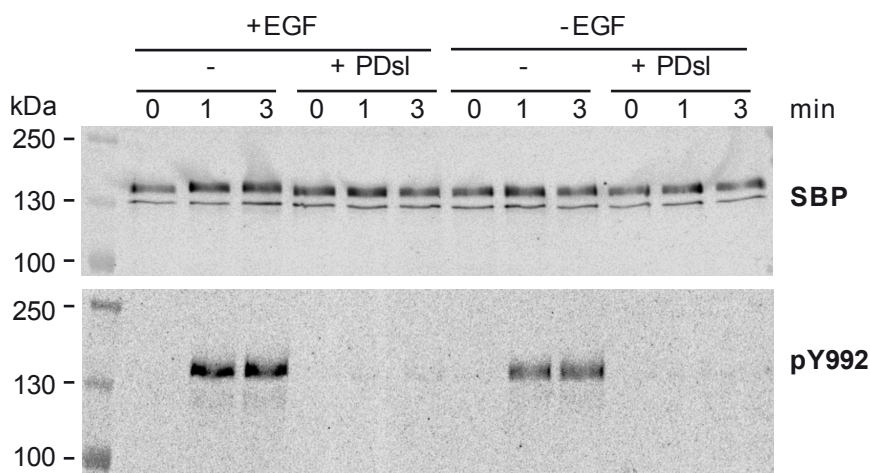
Figure 3.10 Autophosphorylation assay of PDsl-labeled EGFRΔC. Detailed assay procedure is provided in Methods section. 0.375 pmol of phosphorylated samples stopped at indicated time points were visualized via western blot. Total EGFR intensities were detected by Streptavidin on SBP blot, and phosphorylated Y992 by anti-phosphotyrosine antibody on pY992 blot. "-": unlabeled protein, "+PDsl": PDsl-labeled protein.
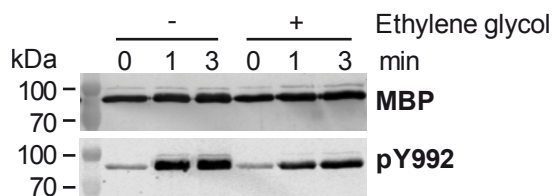
Figure 3.11 Autophosphorylation assay studying effect of deuterated ethylene glycol. Detailed assay procedure is provided in Methods section. 1.4 pmol of phosphorylated samples stopped at indicated time points were visualized via western blot. Total EGFR intensities were detected by anti-MBP antibody on MBP blot, and phosphorylated Y992 by anti-phosphotyrosine antibody on pY992 blot.

Since final PELDOR samples should be heavily supplemented with deuterated ethylene glycol (50%), kinase activity under such conditions needs to be investigated. (The reason for the supplementation will be explained in Section 3.1.4). As shown in Figure 3.11, sample with 50% deuterated ethylene glycol (+) shows slightly less phosphorylation than sample without deuterated ethylene glycol (-), which could be explained by increased viscosity and reduced phosphorylation rate by ethylene glycol. As kinase activity is not drastically reduced in 50% ethylene glycol, this study can proceed without major concern regarding a potential denaturing effect of ethylene glycol.

3.1.4 EPR measurement of PDsl-labeled constructs

3.1.4.1 CW-EPR for spin counting

The mission of this section is to determine labeling efficiency, which is defined as the ratio of protein-bound spin concentration to protein concentration. Labeling efficiency reflects the extent to which protein is spin-labeled. Since underlabeling hinders subsequent PELDOR data interpretation, it is necessary to confirm the sufficient labeling of all samples before PELDOR measurement.
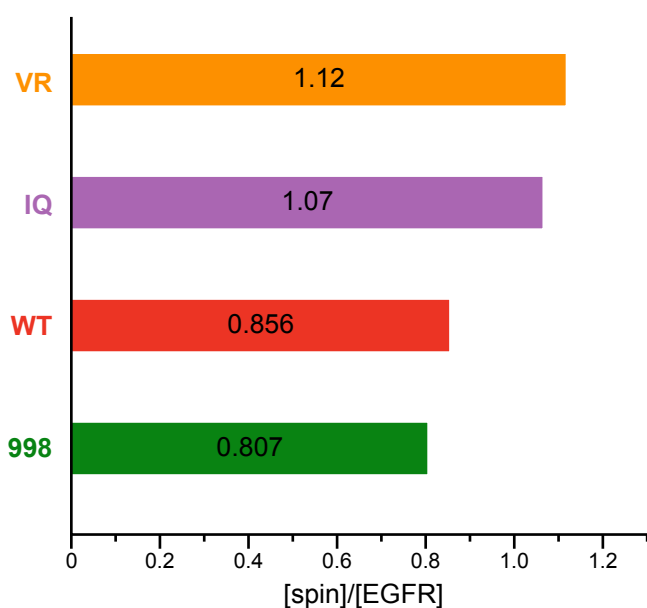
Figure 3.12 Labeling efficiencies of MBP-tagged constructs by PDsl. Horizontal axis represents the labeling efficiency as the ratio between spin and protein concentrations. Each sample is presented as a color-coded horizontal bar with length equal to its marked labeling efficiency. All samples are spin-counted in the absence of ethylene glycol. WT: MBPT-ICD, 998: MBPT-ICD998, VR: MBPT-ICD998(V924R), IQ: MBPT-ICD998(I682Q).

Spin concentrations were determined via spin counting, while protein concentrations were measured via NanoDrop beforehand. To count the spin, X band continuous wave EPR (CW-EPR) and subsequent analysis were performed by Gregor Hagelüken. Double integrals of the field sweep spectra of labeled samples were compared with an internal reference of the spectrometer in order to derive the spin concentrations. Due to distinct labeling procedure EGFRΔC sample contains free spin, which means counted spin is not exclusively protein-bound spin. Therefore, labeling efficiency of EGFRΔC was not determined. For other constructs, ratios between spin and protein concentrations are illustrated in Figure 3.12 as labeling efficiencies. As can be seen, labeling efficiencies of "VR" and "IQ" are close to 1.1, yet for "WT" and "998" they are close to 0.85. Considering

the +/- 10-20% accuracy of spin counting, the variation between values is acceptable. Therefore, samples are considered sufficiently labeled, and ready for PELDOR measurement.

3.1.4.2 PELDOR for conformation study

After evaluating the labeling efficiency of PDsl, Q-band PELDOR measurements of spin-labeled proteins were performed by Gregor Hagelüken, in order to explore the dimer distribution of EGFR. Samples were prepared as described in Methods section, and measured in PELDOR at the same protein concentration. Noteworthy, all samples were exchanged into D_2O-containing buffer to reduce proton spin diffusion and achieve longer T_2 relaxation time. Furthermore, labeled samples were mixed with deuterated ethylene glycol in 1:1 volume ratio before being flash-frozen in Q band PELDOR tubes. Without ethylene glycol, water would form ice crystal during freezing process and expel out labeled protein, which causes spin clustering and shorter T_2 relaxation times. In PELDOR, longer T_2 relaxation time is generally desired for measuring long spin-spin distance and achieving high signal noise ratio.
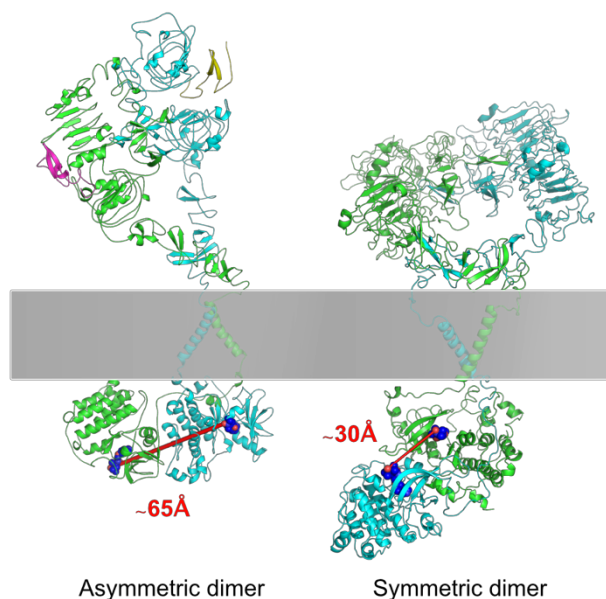
Figure 3.13 Simulation models of asymmetric and symmetric EGFR dimers (Arkhipov et al., 2013). Both dimer conformations are depicted as cartoon, with either monomer colored green or cyan. The grey horizontal bar represents plasma membrane, while magenta and yellow cartoons represent EGF peptides bound to extracellular domains. Modeled spin labels are denoted by dark blue spheres, and interspin distances are presented as red lines with labeled values. This figure was prepared by Gregor Hagelüken using PyMOL Molecular Graphics System.

At first, based on ICD-PD168393 cocrystal structure (PDB ID: 2J5F) and simulation models (Arkhipov et al., 2013), spin-spin distances in symmetric and asymmetric dimers were calculated by mtsslWizard to be $30\,\text{Å}$ and $65\,\text{Å}$, respectively (Figure 3.13). Thereafter, MBP-tagged constructs containing EGFR intracellular domain were measured in PELDOR. The data were processed using DeerAnalysis2016 Software (Jeschke et al., 2006). Briefly, the background was removed by dividing signal by a stretched exponential function and the spin-spin dipolar distances were extracted by Tikhonov regularization.

Surprisingly, both MBPT-ICD and MBPT-ICD998 show single well-shaped peak around $63\,\text{Å}$ in distance distribution, close to the value for asymmetric dimer conformation. No discernible peak can be found at other distances, especially the one corresponding to symmetric dimer (Figure 3.15). To identify the nature of this single peak, V924R and I682Q mutations were introduced individually, so that if $63\,\text{Å}$ peak disappears in their separate PELDOR measurements and reappears when two mutants are combined, it would confirm that the peak indeed results from asymmetric dimer. However, "VR" and "IQ" still result in the same PELDOR distance peak centered around $63\,\text{Å}$ (Figure 3.15). Also, "VR" contains three bumps at smaller distances, probably due to lack of pronounced oscillation pattern in its time trace (Figure 3.14 & 3.15).
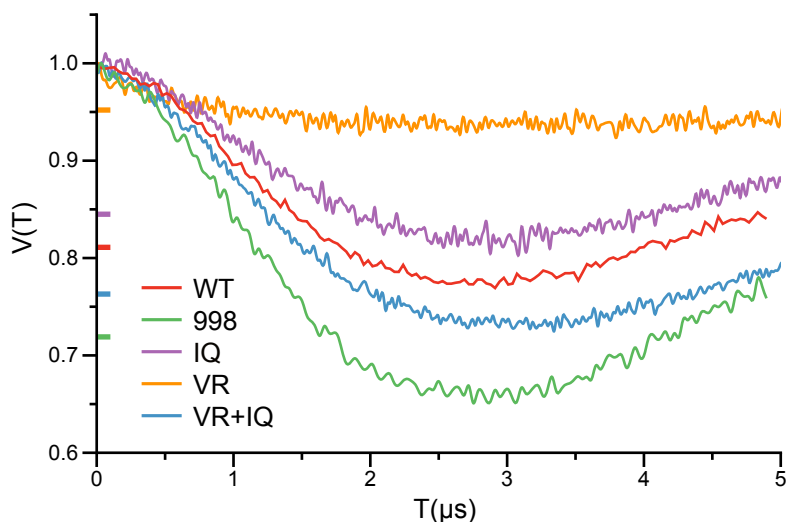
Figure 3.14 Corrected PELDOR time traces of MBP-tagged constructs. Each sample is at 13µM concentration, except "VR + IQ" (7µM+ 7µM). Original time traces were divided by background functions (in the form of exponential decay) to acquire corrected time trace. V(T) is integral echo intensity normalized at time zero and horizontal axis is spin-spin dipolar evolution time. Modulation depth Δ in each measurement is marked as a short bar on the vertical axis at value of 1- Δ. Time trace and modulation depth bar for each construct are color-coded. Uncorrected time traces and background functions are plotted in Appendix Figure S1. WT: MBPT-ICD, 998: MBPT-ICD998, VR: MBPT-ICD998(V924R), IQ: MBPT-ICD998(I682Q).

31
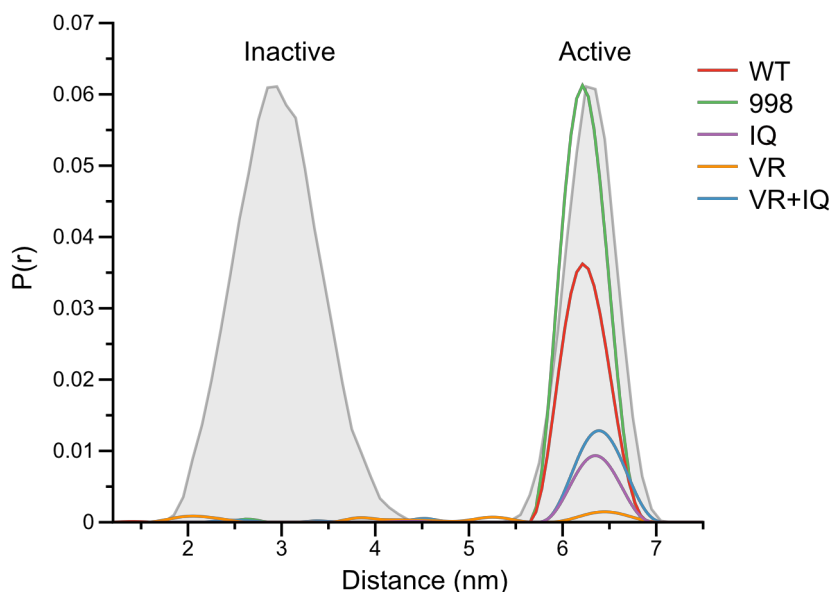
Figure 3.15 Distance distributions of MBP-tagged constructs. Horizontal axis represents spin-spin distance in nanometer, while P(r) on the vertical axis is distance distribution function. Curves of all constructs are color-coded and superimposed with simulated asymmetric or symmetric dimer distance distribution (grey peaks). WT: MBPT-ICD, 998: MBPT-ICD998, VR: MBPT-ICD998(V924R), IQ: MBPT-ICD998(I682Q).
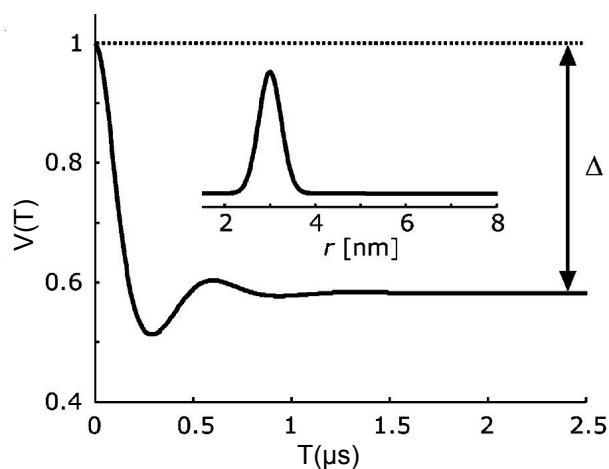
Figure 3.16 Illustration of modulation depth of a corrected PELDOR time trace from simulation. Vertical axis V(T) represents normalized integral echo intensity, while horizontal axis represents dipolar evolution time. Vertical double-headed arrow denotes the modulation depth Δ of the dipolar oscillation pattern. Corresponding distance distribution is shown in the inset. Figure is adapted from Jeschke and Polyhach, 2007.

Another perspective for data interpretation in Figure 3.14 is modulation depth comparison. Modulation depth Δ of a PELDOR time trace is described as difference between $V(T)_{max}$ (1.0 at T = 0) and V(T) at the time point when spin-spin dipolar oscillation almost dies out, as illustrated in Figure 3.16. Modulation depth is correlated with the number of coupled spins. The more spins are coupled, the more deeply time traces are modulated. In Figure

3.14, modulation depth for each individual construct is marked on vertical axis at the value of 1- Δ. Although "VR" and "IQ" both show peaks centered around the same distance as unmutated MBPT-ICD998 does (63Å), they have smaller modulation depths (4.8% and 15.5%, respectively) compared to MBPT-ICD998 (28.1%). Since data in previous section indicate that labeling efficiencies of "VR" and "IQ" are close to each other and slightly higher than MBPT-ICD998, it rules out the possibility that shallower modulations in "VR" and "IQ" are caused by underlabeling. Instead, data can be explained by disruption of asymmetric dimer interface and lack of symmetric dimerization in V924R and I682Q samples. These two mutations perhaps push EGFR conformation equilibrium from asymmetric dimer only towards monomeric form, not towards symmetric dimer. Thus fewer spins are coupled, resulting in reduced modulation depth of the time traces. This explanation is corroborated by PELDOR measurement of "VR+IQ", where mixture of the two mutants shows deeper modulation (23.7%) than either of them alone. As elaborated in 3.1.3.3, V924R and I682Q mixture could form asymmetric "heterodimers", and thereby partially regains the deep modulation. Yet still the modulation is not fully restored to the depth of unmutated MBPT-ICD998, because there is only 50% chance that a V924R molecule encounters an I682Q molecule with which it can form an asymmetric dimer, whereas every two MBPT-ICD998 molecules are able to pair up as an asymmetric dimer. Additionally, since orientation selectivity of the spin label could also cause differences in modulation depths, PELDOR time traces at different frequency offsets were recorded to exclude this possibility (Yin et al., 2017).
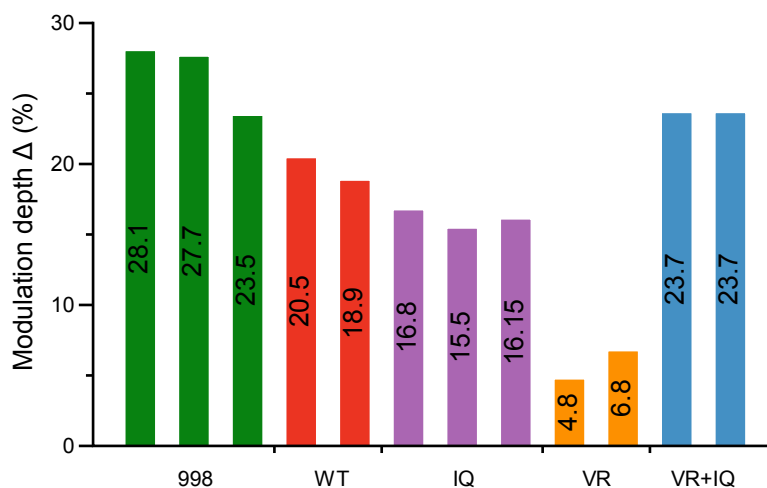
Figure 3.17 PELDOR modulation depths of MBP-tagged constructs. Each bar is color-coded same as in Figure 3.14, with its height equal to modulation depth. Values of modulation depth Δ are labeled as percentage on the bars. WT: MBPT-ICD, 998: MBPT-ICD998, VR: MBPT-ICD998(V924R), IQ: MBPT-ICD998(I682Q). Each bar represents a single measurement.

Another interesting comparison lies between MBPT-ICD998 and MBPT-ICD, where the latter shows reduced modulation depth (18.9%) compared to the former (28.1%). It is also

observed that labeling efficiencies of these two constructs are similar, with MBPT-ICD slightly higher than MBPT-ICD998 (see previous section), thus excluding underlabeling as the reason for shallower modulation. According to other studies, C-terminal tail exerts autoinhibitory influence over the EGFR kinase activity by pushing conformation equilibrium away from asymmetric dimer (Bublil et al., 2010; Jura et al., 2009). Therefore, reduced modulation depth of MBPT-ICD could be explained by less asymmetric dimer formation because of C-terminal tail autoinhibition, although this explanation requires further experimental confirmation. Additionally, each construct was purified, spin-labeled and measured in PELDOR at least one more time, with consistent modulation depth data summarized in Figure 3.17.

So far, all of the MBP-tagged constructs have been measured in PELDOR, mostly with distance distribution showing a single peak centered around $63\,\text{Å}$, which is assigned to ICD asymmetric dimer after modulation depth analysis. However, it is unclear whether the large MBP-tag (40kDa) influences spin-spin distance, or whether measurement of intracellular domain alone reflects the conformation distribution of EGFR as a transmembrane protein. To answer these questions, EGFRΔC construct with two distinct features was designed. First, MBP-tag is replaced by a small SBP-tag (4.3kDa) fused at C-terminus. Second, as mentioned in section 3.1.1, this construct contains almost full length EGFR (residue 1-998) with only C-terminal tail partly missing. During purification, EGFRΔC was solubilized by Triton$^{®}$ X-100 micelles. Since Triton$^{®}$ X-100 interferes with protein concentration measurement at 280nm on NanoDrop, EGFRΔC concentration was determined via western blot, by comparing its band intensity with those of serially diluted EGFR-ICD998(K721M), of which the concentration can be measured via 280nm absorption. Thereafter, EGFRΔC was labeled by mixing it with PDsl at molar ratio of 1:1.1, immediately followed by ethylene glycol supplementation and flash freezing. Since there was no gel filtration step to remove unbound PDsl, labeling efficiency could not be correctly calculated from measured spin concentration. Also, free spins reduce modulation depth of PELDOR trace, so no conclusion could be drawn from modulation depth comparison. Therefore, data interpretation for EGFRΔC is confined to analyzing extracted distance distribution.

As can be seen in Figure 3.18, PELDOR trace of EGFRΔC shares a similar oscillation pattern with MBPT-ICD998 trace. Accordingly, its interspin distance distribution features the same pattern as MBPT-ICD998 does, *i.e.* the peak around $63\,\text{Å}$ (Figure 3.19). Therefore, labeled SBP-tagged EGFRΔC also stays in asymmetric dimer conformation under the conditions of Triton$^{®}$ X-100 micelle, so the conclusion on conformational

distribution of MBP-tagged EGFR-ICD constructs can be extended to larger transmembrane EGFR construct.
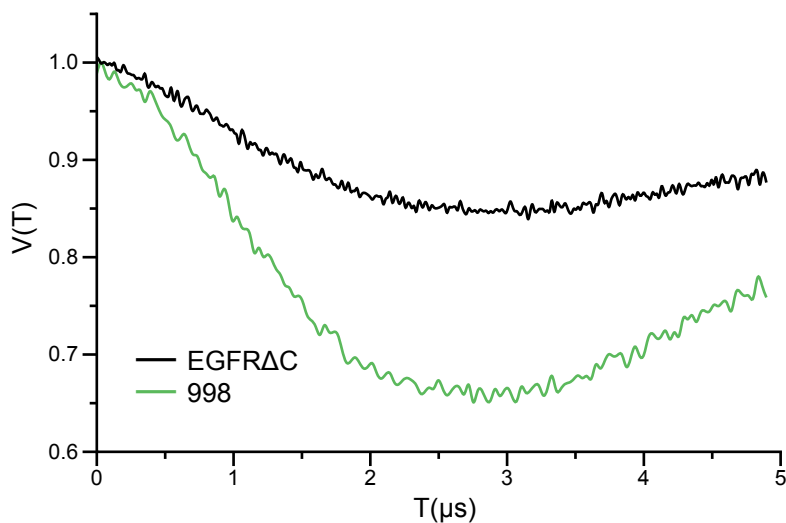
Figure 3.18 Corrected PELDOR time traces of EGFRΔC and MBPT-ICD998. Sample concentration is 13μM. "998" trace is taken from Figure 3.14 as a reference. V(T) is integral echo intensity normalized at time zero and horizontal axis is spin-spin dipolar evolution time. Time trace for each construct is color-coded. Uncorrected time traces and background functions are plotted in Appendix Figure S1. 998: MBPT-ICD998.
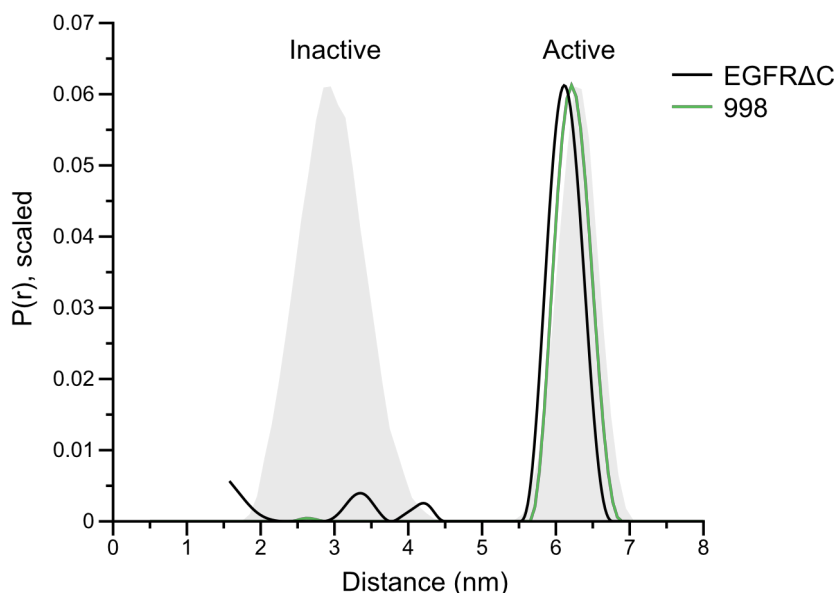
Figure 3.19 Distance distributions of EGFRΔC and MBPT-ICD998. Horizontal axis represents spin-spin distance in nanometer, while P(r) on the vertical axis is distance distribution function. Peaks centered around $63\,\text{Å}$ are scaled to the same height for comparison. Curves of both constructs are color-coded and superimposed with simulated asymmetric or symmetric dimer distance distribution (grey peaks). 998: MBPT-ICD998.

3.1.5 Analytical gel filtration

As mentioned in previous section, the shallower modulation of "VR" and "IQ" PELDOR traces can be explained by the shift of conformational equilibrium from asymmetric dimer to monomer. The goal of this section is to compare oligomeric states cross different PELDOR samples via analytical gel filtration, thereby further corroborating that explanation.
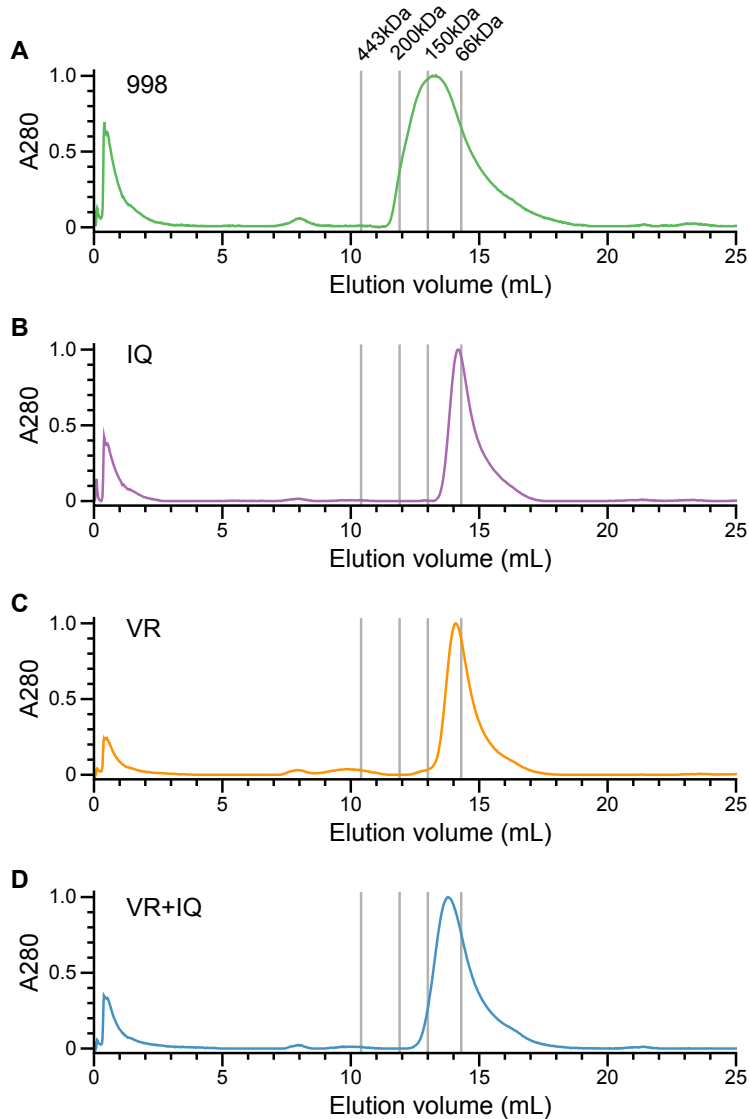
Figure 3.20 Analytical gel filtration of MBP-tagged constructs. Vertical axis represents relative absorbance at 280nm. PELDOR samples were retrieved and injected into Superdex™ 200 10/300 GL column. The chromatogram of each construct is rescaled with maximum absorbance at 1.0, and color-coded in the same way as in Figure 3.14. Protein standards from Gel Filtration Marker Kit MWGF1000 were injected separately as references, and the elution volume for maximum absorbance of each standard is denoted in the plot as a grey vertical line, with corresponding molecular weight labeled on the top. 998: MBPT-ICD998, VR: MBPT-ICD998(V924R), IQ: MBPT-ICD998(I682Q).

Analytical gel filtration of four PELDOR samples, *i.e.* MBPT-ICD998, "VR", "IQ" and "VR+IQ", was performed by Fraser Duthie. First, four gel filtration markers were injected successively into pre-equilibrated Superdex^{TM} 200 10/300 GL column. Thereafter, PELDOR samples were retrieved from Q-band tubes and injected separately into the same column. Chromatograms were then compared between PELDOR samples and protein markers.

In Figure 3.20, four protein markers are presented as grey vertical lines in each graph. The chromatogram of every PELDOR sample contains single major peak, with maximum absorbance passing between "150kDa" and "66kDa" markers. Since theoretical mass differences among these constructs are negligible, it is safe to assume that monomeric state is 82kDa and dimeric state is 164kDa for all constructs. The peak of "998" is broad with a long tail after the maximum absorbance around "150kDa", suggesting a majority of dimeric state mixed with monomers. On the other hand, "VR" and "IQ" show more focused peaks close to "66kDa", whereas "VR+IQ" lies between "998" and "VR"/ "IQ". Therefore, both "VR" and "IQ" stay more in the monomeric state compared to "998", while "VR+IQ" partially restores the dimeric state of "998", which fits well with the conclusion from PELDOR. Since the elution volume gap between "150kDa" and "66kDa" is narrow (1.3mL), monomeric and dimeric states cannot be resolved as separate peaks in chromatograms, which explains why "VR+IQ" shows only one peak.

To conclude the EGFR conformation study, a new spin labeling strategy, *i.e.* inhibitor-directed spin labeling (IDSL) has been developed, and it was successfully applied to EGFR conformation study by designing one such inhibitor-based spin label PDsl for interspin distance determination. PELDOR measurements of MBP-tagged ICD and SBP-tagged EGFRΔC show single peak in distance distribution which corresponds to asymmetric dimer of kinase domain. V924R and I682Q mutations shift the conformational equilibrium from asymmetric dimer towards monomer while no symmetric dimer is observed. Analytical gel filtration further corroborates the equilibrium shift by examining the oligomeric states of PELDOR samples.

3.2 Protein interaction study of EGFR

3.2.1 Purification of constructs for interaction study

Shown below is an overlook of constructs involved in the protein interaction study (Figure 3.21). All constructs are based on human sequences. As protein interactions of EGFR intracellular domain are studied, all EGFR constructs contain only the intracellular domain or parts thereof.
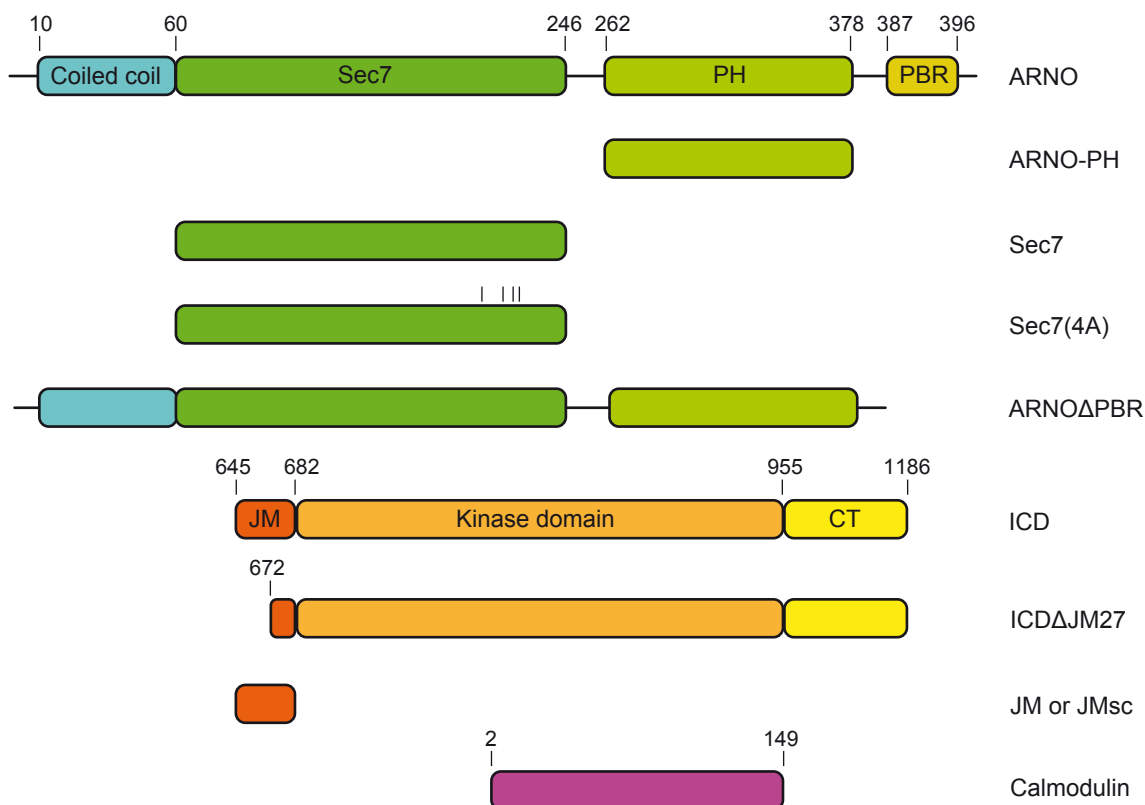
Figure 3.21 Domain sketch of all the constructs used in interaction study. Sequence numbering is based on human proteins. Mutations in Sec7(4A) are indicated by vertical bars. JMsc: scrambled JM.

Details about plasmids, recombinant protein expression and purification procedures can be found in Materials and methods. Regarding purification, briefly, proteins were purified at first via different affinity chromatography methods depending on the attached tags. Eluted samples were subsequently buffer-exchanged into desalting buffer, followed by TEV protease cleavage. TEV-treated samples were incubated with affinity resin again to remove the free affinity tags and uncleaved protein, after which size exclusion chromatography was performed to remove impurities and protein aggregate. Samples from target peak were collected and if necessary, concentrated with centrifugal filters to reach the concentration as required by interaction study.

It is worth mentioning that ICDΔJM27 construct could only be obtained in sufficient amounts, when the first 27 amino acids of JM were still present during expression. Therefore, a TEV cleavage site was inserted between amino acids 27 and 28 of the JM segment and the fragment containing first 27 amino acids was cut off by TEV during purification, leading to final product of ICDΔJM27.

Purified protein samples were visualized via SDS-PAGE and Coomassie staining. As can be seen in Figure 3.22, final products are very pure except that ARNOΔPBR is slightly contaminated by multiple residual protein bands. Moreover, JM and scrambled JM (noted as JMsc) show less focused bands in SDS-PAGE, probably due to smaller sizes. To sum up, acquired protein samples are fit for the downstream analyses.
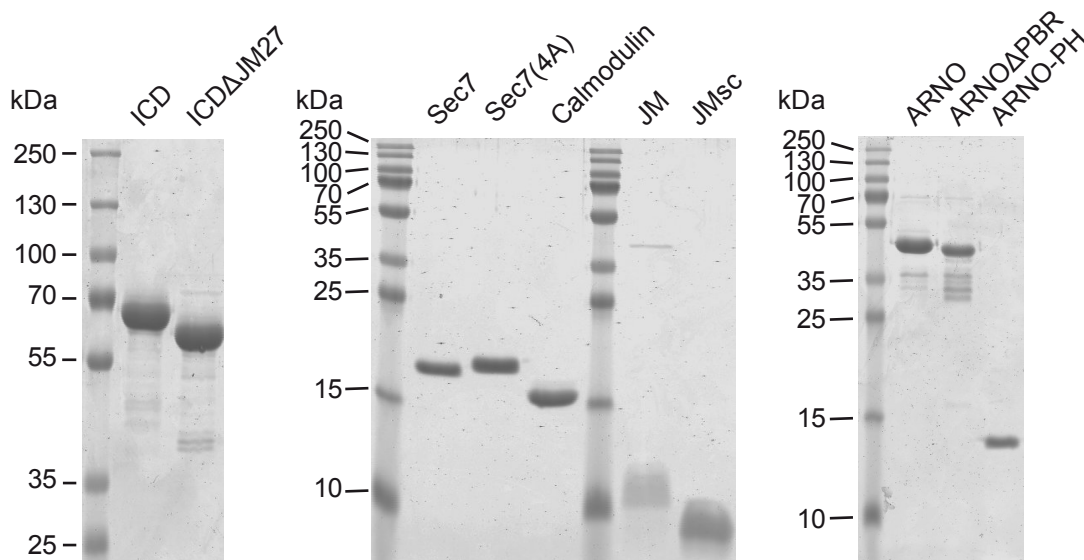
Figure 3.22 Final products of EGFR, ARNO and calmodulin constructs. 50 pmol of concentrated samples were visualized via SDS-PAGE (200 pmol for JM and JMsc). From left to right, the separating gels are 8%, 15% and 12.5%, respectively. Purification detail is in Methods section.

3.2.2 Quality control of purified constructs

Once protein purification is completed, it would be necessary to perform certain activity assay to examine whether protein functions are still intact. For EGFR, autophosphorylation assay was applied to check kinase activity. As for ARNO-Sec7, it functions as a guanine nucleotide exchange factor on ARF GTPases, so various ARNO-related constructs were tested in guanine nucleotide exchange assay to compare aforementioned activities between constructs.

3.2.2.1 Autophosphorylation assay

Autophosphorylation assay of ICD and ICDΔJM27 was conducted as described in Methods section, followed by western blot with antibody against generic phosphotyrosine and antibody recognizing cytoplasmic domain of EGFR. Protein bands were visualized by corresponding secondary antibodies conjugated to fluorophores, before being scanned and quantified by Image Studio™ software. Phosphorylation intensities were normalized against total EGFR intensities for each time-point. Such normalized phosphorylation was

used to further calculate relative phosphorylation (rel. phos.) at 1min and 3min, with values at 0 min set as 1. In Figure 3.23, ICDΔJM27, lacking major part of JM, shows much less phosphorylation than ICD. At 1min, only 12% of relative phosphorylation was left in ICDΔJM27 compared to ICD, while at 3min, 15%. The dramatic drop of activity can be attributed to the deletion of first 27 amino acids of JM segment, which plays a critical role in dimerization and activation of EGFR kinase domain. So far the phosphorylation result is well in line with canonical EGFR activation model, supporting the usage of these two constructs for further interaction studies.
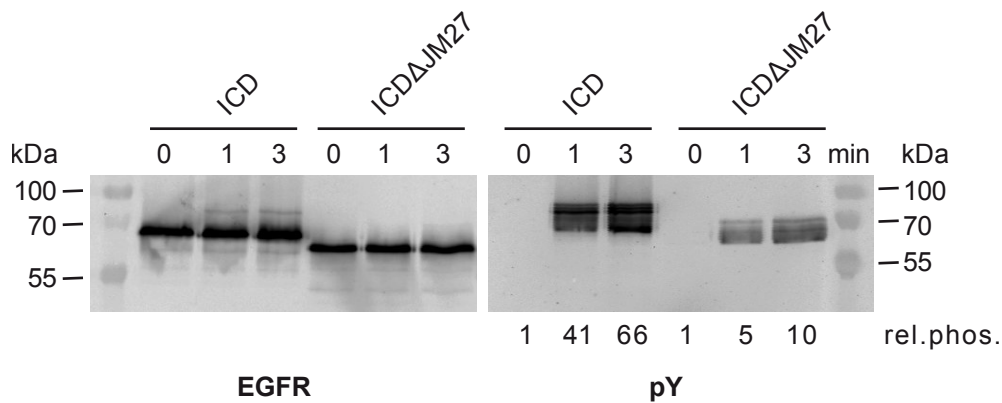
Figure 3.23 Autophosphorylation assay of ICD constructs analyzed by western blot. Assay was performed with 100nM ICD and ICDΔJM27. Total EGFR intensities were detected by anti-EGFR antibody on the left (EGFR blot) while phosphorylated EGFR by anti-phosphotyrosine antibody on the right (pY blot). Relative phosphorylation (rel. phos.) is marked under each lane of the pY blot.

3.2.2.2 Guanine nucleotide exchange assay

Sec7 domain functions as GEF on Arf GTPase such as Arf1, which is exploited in activity assay of Sec7-containing constructs. The common guanine nucleotide exchange assays use radioactive, fluorescent-labeled nucleotide, non-hydrolysable nucleotide analog or just plain GTP/GDP pair (Eberth and Ahmadian, 2009; Randazzo et al., 2013; Stalder et al., 2011). For Sec7:Arf1 nucleotide exchange, an assay system with GTP/GDP pair has been well established in the lab, which monitors intrinsic tryptophan fluorescence instead of radioactivity or external fluorescence. In addition, it is unnecessary to use non-hydrolysable nucleotide such as GppNHp due to the low inherent GTPase activity of His-NΔ17Arf1, a variant of Arf1 truncated for better solubility.

The basis of the tryptophan fluorescence assay is that the structural switch of GTPase from GDP-bound state to GTP-bound state, and the concomitant microenvironment change around tryptophan residue(s) on GTPase leads to tryptophan fluorescence change. Here four Sec7-containing ARNO constructs were tested as GEFs in exchange

assay: Sec7, Sec7(4A), ARNO and ARNOΔPBR. Assay was performed at room temperature as described in Methods section. Briefly, His-NΔ17Arf1 was preloaded with GDP, before being combined with different ARNO constructs or assay buffer as negative control. GTP addition started exchange reaction, in which tryptophan fluorescence change was monitored every 10 second for 15min, with excitation wavelength of 280nm and emission wavelength of 340nm. One can clearly see that Sec7 has the highest exchange activity, showing fastest tryptophan fluorescence increase (Figure 3.24). 4A mutation effectively destroys exchange activity of Sec7, since its time trace is as flat as buffer control. Details about 4A mutation will be elaborated in section 3.2.3.4. The time trace of ARNO looks flat as well, probably due to the autoinhibition by Sec7-PH linker and C-terminal helix, which fold back to Sec7 and obstruct docking of Arf1. On the other hand, deletion of PBR in ARNOΔPBR partially restores its exchange activity towards Arf1. Qualitatively speaking, the data obtained here fit well with the published catalytic efficiency data (DiNitto et al., 2007). All constructs are suited to further interaction study.
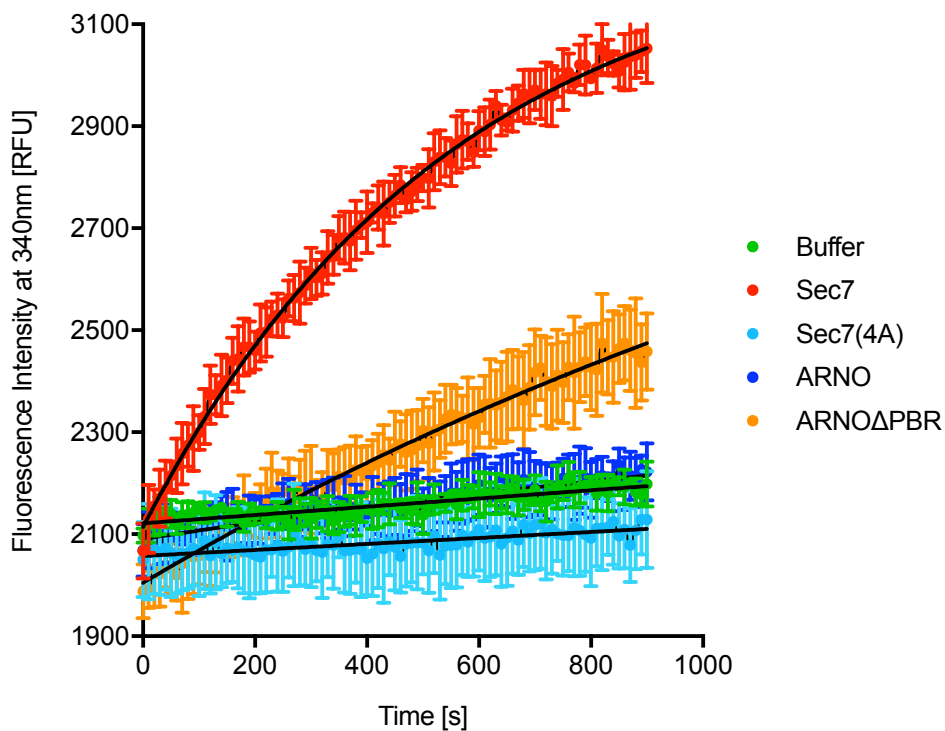
Figure 3.24 Guanine nucleotide exchange assay of various ARNO constructs with Arf1. Indicated ARNO proteins or buffer was combined with GTPase His-NΔ17Arf1 for the assay. Tryptophan fluorescence increase upon nucleotide exchange was monitored by Tecan plate reader at 340nm for 15min. Triplicates in each dataset are displayed as dots and error bars, representing mean and standard deviation, respectively. RFU: relative fluorescence units.

3.2.3 Investigation of interaction between EGFR-ICD and ARNO

In order to investigate interaction between EGFR-ICD and ARNO, MST assay was performed, because it is free of immobilization compared to surface plasmon resonance (SPR), and with minimal protein consumption compared to isothermal titration calorimetry (ITC). In addition, compared to other protein-protein interaction analyses such as crosslinking and label transfer, MST is much faster, taking only less than 30min for one complete assay, whereas the former two need hours up to one day. Standard MST assay, which has been applied to this thesis project, is introduced in detail concerning both principle and practice in Methods section. On the other hand, label-free MST, which exploits intrinsic tryptophan fluorescence of one protein for detection but requires the other to be tryptophan-free, was not performed here due to the presence of tryptophans in both EGFR-ICD and ARNO.

3.2.3.1 ICD interacts with ARNOΔPBR

The protein-protein interaction study begins with MST measurement of ICD with ARNO, ARNOΔPBR or ARNO-PH, aiming at establishing ICD:ARNO interaction and locating the ICD-binding site on ARNO.
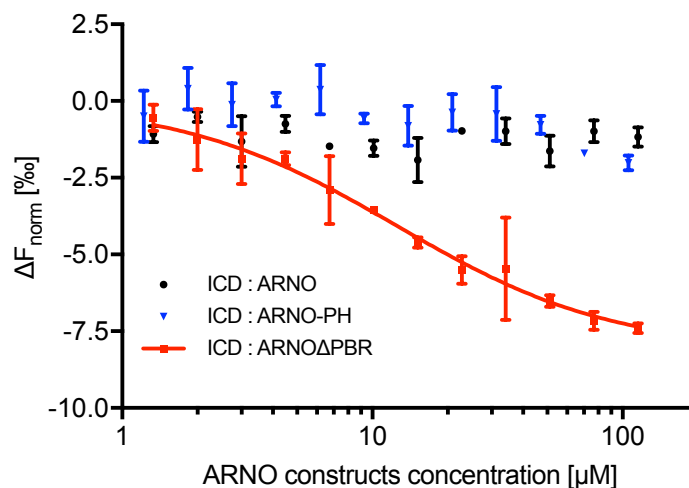
Figure 3.25 MST binding assay of ICD with ARNO, ARNO-PH or ARNOΔPBR. Three independent measurements in each dataset are merged and displayed as points and error bars, which represent means and standard deviations of ΔF_{norm} at each concentration, respectively. Single MST measurements are plotted as scatter graphs in Appendix Figure S2.

Detailed labeling and MST procedures can be found in Methods section. ICD was labeled by fluorescent dye Red-NHS 2nd Generation, with degree of labeling (DOL) determined as 1.41 via NanoDrop. Assay was conducted at 25℃ in MST buffer (20mM HEPES

pH7.8, 150mM NaCl, 0.005% Triton$^{®}$ X-100, 10µM BSA) with 20% excitation power and 40% MST power. Concentration of labeled ICD was held at 100nM, whereas concentration of ARNO, ARNOΔPBR and ARNO-PH varied by 1:3 serial dilution, starting from 115µM for ARNO and ARNOΔPBR, 106µM for ARNO-PH.

As shown in Figure 3.25, ICD recognizes ARNOΔPBR displaying a nice binding curve and Kd of 12.5µM according to dose response fitting of MO.Affinity Analysis. Conversely, ICD:ARNO and ICD:ARNO-PH bear no clear sign of binding, with ΔF_{norm} values all distributed close to zero, and dose response fitting unable to return reasonable values. Altogether, these three datasets reveal that ICD interacts with ARNOΔPBR, but not with ARNO or ARNO-PH. PH domain is clearly ruled out as the binding site for ICD. Furthermore, distinct MST results of ARNO and ARNOΔPBR could be explained by their structural difference due to autoinhibition, with Sec7 domain in ARNO shielded and in ARNOΔPBR more exposed, suggesting Sec7 as the potential binding site for ICD. However, direct measurement of ICD:Sec7 interaction is needed to support this.

3.2.3.2 ICD binds to ARNO-Sec7

The goal of this section is two-fold. First, to investigate direct interaction between ICD and Sec7, as discussed in last section. Second, to identify the Sec7-binding site on ICD, if direct interaction is confirmed. Since the known cytoplasmic modulators of EGFR target either juxtamembrane segment (calmodulin and TRAF4; Aifa et al., 2002; Cai et al., 2018; Martín-Nieto and Villalobo, 1998) or its kinase domain (Mig6; Park et al., 2015; Zhang et al., 2007), ICDΔJM27 was included in this section to first test whether JM could be the binding site for Sec7. In ICDΔJM27, only first 27 amino acids are deleted instead of the whole 38 amino acids of JM, because these extra 11 amino acids are viewed as an integral part of the kinase core and routinely retained in constructs for structural and functional studies (Jura et al., 2009; Stamos et al., 2002; Zhang et al., 2006).

ICDΔJM27 was labeled by fluorescent dye Red-NHS 2nd Generation following the same procedure as ICD, with DOL determined via NanoDrop as 1.13. MST assay was performed with ICD:Sec7 and ICDΔJM27:Sec7 pairs, under the same buffer condition and measurement setup as described in previous section. Concentration of labeled ICD and ICDΔJM27 was held constant at 100nM, while concentration of Sec7 varied by 1:3 serial dilution, starting from 389µM.

One can clearly see that ICD interacts with Sec7 but ICDΔJM27 does not, even though weak interaction at the highest Sec7 concentrations cannot be fullly excluded (Figure 3.26). For ICD:Sec7 pair, dose response fitting of MO.Affinity Analysis returns Kd of

49.2µM. In contrast, ICDΔJM27:Sec7 measurement provides ΔF_{norm} values all around zero, except last two data points with highest Sec7 concentrations. For this scatter plot, software fails to execute dose response fitting. To conclude, ICD does interact with Sec7 directly and JM could be the Sec7-binding site of ICD. However, there is another possibility that JM-mediated ICD dimerization forms a new protein surface that binds to Sec7 and such surface is not present in ICDΔJM27. To resolve this issue, a direct MST measurement of Sec7:JM pair is necessary. Moreover, it is noteworthy that Kd of ICD:Sec7 here is higher than ICD:ARNOΔPBR pair measured in previous section. Since PH domain has been excluded as binding site for ICD, the stronger affinity of ARNOΔPBR towards ICD could be ascribed to N-terminal coiled-coil segment as additional binding site, or ARNO dimerization driven by coiled-coil segment. As shown in Figure 3.23, ICD at 100nM still function as dimers to phosphorylate each other. Therefore, under MST assay conditions, both ARNOΔPBR and ICD exist in dimeric form, and the dimer-dimer interaction may further stabilize ICD:ARNOΔPBR complex, giving stronger apparent affinity. On the other hand, Sec7 exists only as monomer and will not enjoy such additional contribution to binding affinity.
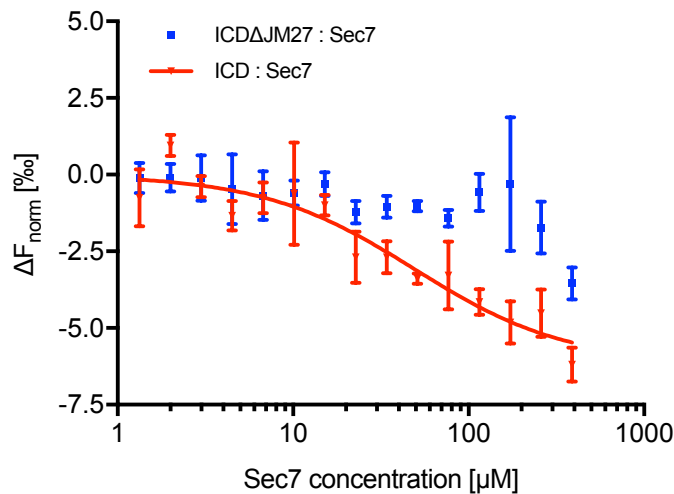
Figure 3.26 MST binding assay of ICD or ICDΔJM27 with Sec7. Three independent measurements in each dataset are merged and displayed as points and error bars, which represent means and standard deviations of ΔF_{norm} at each concentration, respectively. Individual MST measurements are plotted as scatter graphs in Appendix Figure S2.

3.2.3.3 Sec7 interacts with JM peptide directly

As discussed in section 3.2.3.2, Sec7:JM pair should be measured via MST, in order to ascertain that JM segment is the Sec7-binding site on ICD. A scrambled version of JM peptide, *i.e.* the same amino acids but in random sequence, JMsc, was included in control measurement as Sec7:JMsc pair.

Labeling and MST procedures are described in Methods section. Briefly, Sec7 was labeled by Alexa Fluor™ 647 NHS Ester, with DOL determined as 0.746. MST was performed with 60% excitation power and 50% MST power. Concentration of labeled Sec7 was held at 100nM, and concentration of JM and JMsc varied by 1:3 serial dilution, starting from 320µM.

As depicted in Figure 3.27, Sec7 binds directly to JM peptide with Kd of 49.6µM. On the other hand, data from Sec7:JMsc pair could not be fitted by MO.Affinity Analysis. Kd of Sec7:JM agrees surprisingly well with Kd of ICD:Sec7 interaction (49.2µM), considering labeling conditions and MST setup are very different between these two pairs, which supports the argument that Sec7 binds to ICD on its JM segment. Since JM has a basic isoelectric point whereas Sec7 has an acidic pI (11.91 and 6.38, respectively), one could assume unspecific electrostatic interaction of both. However, MST result of Sec7:JMsc excludes the possibility that Sec7 interacts with JM via pure electrostatic attraction, and shows this interaction is sequence-dependent. Further characterization of Sec7:JM interaction on amino acid level shall be elaborated in Discussion section, combined with NMR data from collaborators.
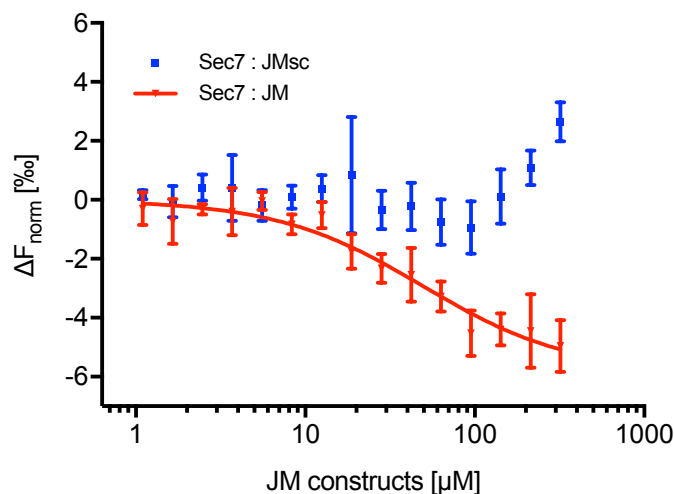
Figure 3.27 MST binding assay of Sec7 with JM or scrambled JM (JMsc). Three independent measurements in each dataset are merged and displayed as points and error bars, which represent means and standard deviations of ΔF_{norm} at each concentration, respectively. Individual MST measurements are plotted as scatter graphs in Appendix Figure S2.

3.2.3.4 Sec7 mutation undermines interaction with ICD and JM

In collaboration with Dr. Manuel Etzkorn and Dr. Aldino Viegas, Sec7:JM interaction was intensively studied via NMR spectroscopy. One of the conclusions from NMR study is that part of the JM-binding site of Sec7 lies in αH-helix (helix nomenclature from Mossessova

et al., 1998), which also participates in Arf1 docking (Cherfils et al., 1998; Mossessova et al., 1998; Renault et al., 2003) and autoinhibition mechanism (DiNitto et al., 2007) via its hydrophobic patch. It would be interesting to find out whether alanine mutations at these hydrophobic residues impair Sec7:JM interaction. In this section, a mutant construct Sec7(4A) was studied via MST, containing Y186A, F190A, I193A and M194A mutations in αH-helix.

Detailed MST protocol is described in Methods section. Sec7(4A) was labeled by Alexa Fluor^{TM} 647 NHS Ester, with DOL of 1.08. For Sec7(4A):JM interaction, MST was performed with 60% excitation power and 50% MST power. Concentration of labeled Sec7 was held at 100nM, and concentration of JM varied by 1:3 serial dilution, starting from 320µM. For ICD:Sec7(4A) interaction, ICD labeled by Red-NHS 2nd Generation (same batch as in previous sections) was measured, with 20% excitation power and 40% MST power. Each capillary contained 100nM labeled ICD and 1:3 serially diluted Sec7(4A) starting at 389µM.

In Figure 3.28 (A), there is no binding event, with all data values fluctuating about zero. In Figure 3.28 (B), labeled-ICD shows some dose response to high concentration of Sec7(4A), but it is nothing like a saturated biding curve, and software fails to execute dose response fitting. Considering the highest Sec7(4A) concentration is 389µM, even there is any possible binding, it would be a really weak one, compared to ICD:Sec7 data from section 3.2.3.2. Therefore, 4A mutation effectively undermines Sec7:JM and ICD:Sec7 interaction, supporting the observation from NMR study.
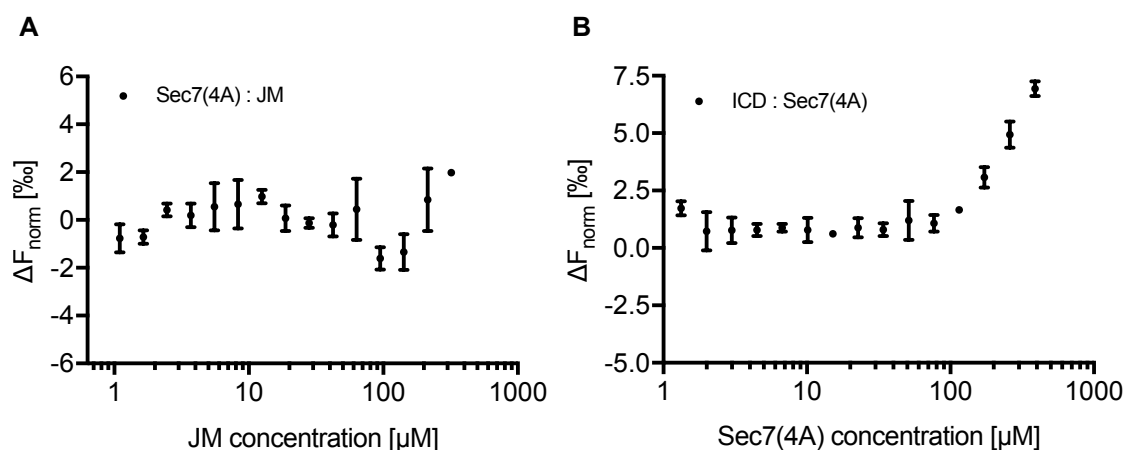
Figure 3.28 MST binding assay of Sec7(4A) with JM and ICD with Sec7(4A). Three independent measurements in each dataset are merged and displayed as points and error bars, which represent means and standard deviations of ΔF_{norm} at each concentration, respectively. Individual MST measurements are plotted as scatter graphs in Appendix Figure S2.

3.2.4 Competition study between calmodulin and Sec7

It has long been known that calmodulin (CaM for short) interacts with JM segment of EGFR (Aifa et al., 2002; Martín-Nieto and Villalobo, 1998). The common binding site on EGFR raises the question of whether calmodulin could compete against Sec7 for binding towards JM. To answer this question, MST binding assay of ICD:CaM pair should be performed first.

3.2.4.1 Calmodulin binds to JM of ICD in a Ca^{2+}-dependent manner

The purpose of the ICD:CaM binding assay is two-fold: to confirm ICD:CaM interaction via MST, and to decide proper dose of CaM for the competition assay.

Description of MST is in Methods section. The same batch of ICD and ICDΔJM27 labeled in previous sections by Red-NHS 2nd Generation was used here. Since ICD:CaM interaction is known to be Ca^{2+}-dependent, one control of ICD:CaM in EGTA-containing buffer was tested in MST, in addition to ICD:CaM and ICDΔJM27:CaM measurements in $CaCl_2$-containing buffer. MST was performed with 20% excitation power and 40% MST power. Concentrations of labeled ICD and ICDΔJM27 were held at 100nM, and concentration of calmodulin varied by 1:2 serial dilution, starting from 65µM.

In the presence of $CaCl_2$, ICD binds nicely to calmodulin with Kd of 0.76µM (Figure 3.29), close to published values (Aifa et al., 2002; Li and Villalobo, 2002; Martín-Nieto and Villalobo, 1998). According to the Kd values acquired so far, CaM has stronger affinity toward ICD than Sec7 does (0.76µM *vs*. 49.2µM), so it is easier to see CaM inhibiting ICD:Sec7 interaction than Sec7 inhibiting ICD:CaM interaction. Therefore, CaM is chosen as the competitor in the following competition assay. In control measurement with 2mM EGTA, ICD:CaM interaction is completely abandoned, and all ΔF_{norm} values fluctuate around zero. Furthermore, ICDΔJM27:CaM in the presence of 2mM $CaCl_2$ also shows no sign of binding. The last two data points on the right are drifted away from zero, which could be attributed to unspecific binding to ICDΔJM27 at high calmodulin concentrations, while the third point to the right with 15µM calmodulin does not show unspecific binding. In ICD:CaM ($CaCl_2$), the data point at the same calmodulin concentration (15µM) lies in the saturation phase of the binding curve. So 15µM is the optimal concentration for competition assay, for it reaches binding saturation without any unspecific interaction. To conclude, the MST assay in this section has proven that ICD recognizes CaM in a JM- and Ca^{2+}-dependent manner, with 15µM calmodulin determined as the optimal competitor for competition assay.
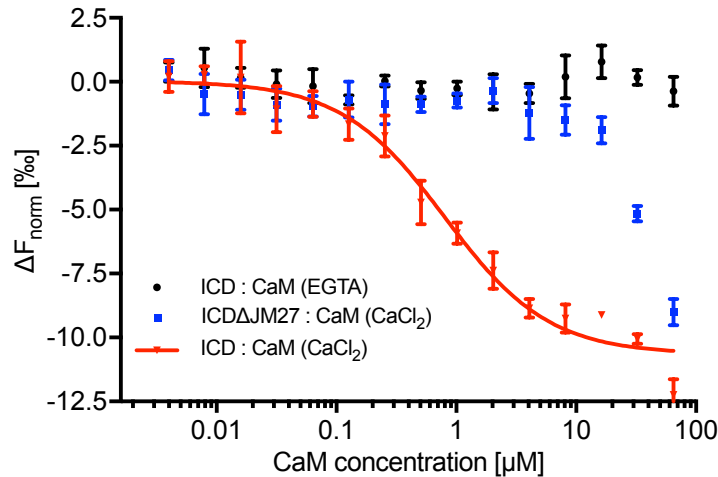
Figure 3.29 MST binding assay of ICD or ICDΔJM27 with calmodulin. ICD:CaM was measured in the presence of 2mM $CaCl_2$ or 2mM EGTA, while ICDΔJM27:CaM was measured in the presence of 2mM $CaCl_2$. Three independent measurements in each dataset are merged and displayed as points and error bars, which represent means and standard deviations of ΔF_{norm} at each concentration, respectively. Individual MST measurements are plotted as scatter graphs in Appendix Figure S2.

3.2.4.2 Calmodulin and Sec7 compete for binding to ICD

MST procedure is described in Methods section. Assay was performed in buffer containing 2mM $CaCl_2$, with 20% excitation power and 40% MST power. The same batch of ICD labeled by Red-NHS 2nd Generation was used here. In both "+ CaM" and "-CaM" measurements, labeled ICD was held at constant concentration of 100nM, while Sec7 concentration varied by 1:3 serial dilution, starting from 389µM. In "+ CaM", 30µM of calmodulin was pre-incubated with labeled ICD at RT for 5min, before being mixed with equal volume of serially diluted Sec7.

ICD:Sec7 interaction gives a nice binding curve with Kd of 68.4µM, not far from 49.2µM determined in section 3.2.3.2. The competition of 15µM calmodulin undermines this interaction, leaving no sign of binding in "+ CaM" measurement, so CaM successfully blocks interaction between ICD and Sec7 (Figure 3.30). However, other than CaM competes against Sec7 by binding to ICD, there is an alternative explanation that CaM binds to Sec7 and prevents it from interacting with ICD, thus leaving ICD unbound. Further data analysis excludes such possibility. As mentioned in Methods section 5.2.6, $F_{norm,unbound}$ and $F_{norm,bound}$ are signatures of two species, *i.e.* labeled protein alone and labeled protein bound to unlabeled protein, respectively. In Figure 3.31, three datasets from previous figures are superimposed, with F_{norm} instead of ΔF_{norm} on the Y axis. Within each dataset, points close to left end of the plot is viewed as asymptotically approaching

$F_{norm,unbound}$, as most of the labeled protein molecules are in unbound state. Meanwhile, points close to right end of the plot is viewed as asymptotically approaching $F_{norm,bound}$, where unlabeled protein saturates the binding site on the labeled protein. As revealed in Figure 3.31, left ends of "ICD:CaM ($CaCl_2$)" and "ICD:CaM (EGTA)" are both approaching F_{norm} of 920‰, where labeled ICD is unbound. On the right end of these two groups, "ICD:CaM ($CaCl_2$)" is approximating 910‰, signature of ICD-CaM complex; "ICD:CaM (EGTA)" is still around 920‰, meaning ICD remains unbound in EGTA-containing buffer. Now in "+ CaM" measurement, data points fluctuate around 910‰, which means labeled ICD is still in CaM-bound state. That contradicts the alternative explanation and corroborates the hypothesis that CaM effectively hinders ICD:Sec7 interaction by directly binding to ICD, not to Sec7. Moreover, in "+ CaM", highest concentration of Sec7 is 389μM while CaM is only 15μM, which leaves a far-fetched scenario that CaM blocks the interaction by directly binding to such massive amount of Sec7.
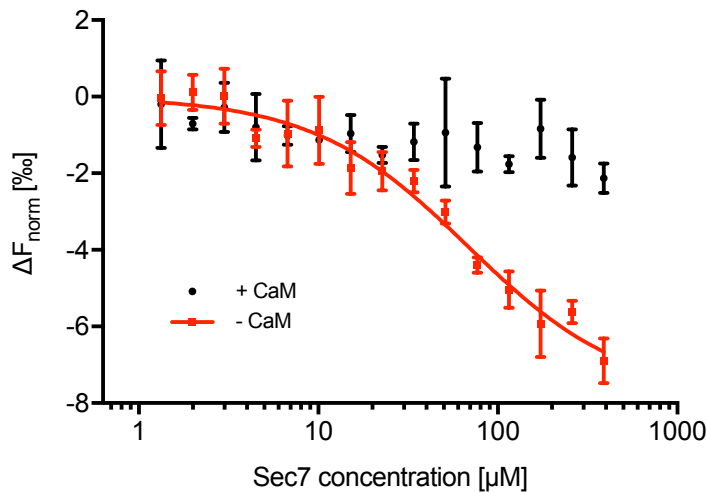
Figure 3.30 MST assay with calmodulin competition. "- CaM" indicates ICD:Sec7 interaction without CaM, and "+ CaM" indicates ICD:Sec7 with CaM as competitor. Assay was performed in buffer containing 2mM $CaCl_2$. Three independent measurements in each dataset are merged and displayed as points and error bars, which represent means and standard deviations of ΔF_{norm} at each concentration, respectively. Individual MST measurements are plotted as scatter graphs in Appendix Figure S2.

Figure 3.31 Superimposition of previous MST results. Three datasets from Figure 3.29 and Figure 3.30 are replotted as F_{norm} *vs.* Constructs concentration, and labeled according to the same nomenclature as in original figures. Three independent measurements in each dataset are merged and displayed as points and error bars, which represent means and standard deviations of F_{norm} at each concentration.

4 Discussion

4.1 PELDOR: a suitable tool for EGFR conformation study

In EGFR conformation study (section 3.1), PELDOR has revealed a single conformation in all tested constructs, be it soluble MBP-tagged ICD or SBP-tagged near-full-length EGFRΔC in micelles. The spin-spin distance agrees with that of crystallized asymmetric dimer (PDB ID: 2J5F) very well. Unlike crystallography, NMR and fashionable cryo-EM, which provide all-atom structure of protein, PELDOR accurately measures the interspin distance in spin-labeled protein, and delivers key information such as inter-/intramolecular conformation changes in response to different conditions (Glaenzer et al., 2017) or location of metal ion cofactor critical for catalysis (Abdullin et al., 2015). As is known, crystallography relies upon the unpredictable outcome of protein crystallization screening, NMR is mainly limited to protein smaller than 50kDa (reviewed by Frueh et al., 2013) with additional stability and concentration requirement, and cryo-EM has an expected lower mass limit of 38kDa (reviewed by Henderson, 1995). In contrast, PELDOR analyzes samples in solution, has no protein size limit, and can be readily applied to various conditions such as detergent, lipid and even cellular environments. The distinctive features make PELDOR a powerful tool complementary to other structural biology techniques. For example, EGFR conformation study in this thesis started from analyzing available cocrystal structure of ICD-PD168393 (PDB ID: 2J5F) and molecular dynamics (MD) simulation files (Arkhipov et al., 2013). By carefully inspecting binding mode of the inhibitor, and calculating the theoretical interspin distances via mtsslWizard, it was proven to be feasible to employ a spin-labeled inhibitor for PELDOR measurement. Later on, PELDOR provided experimental data that faithfully reflect crystal structure of the asymmetric dimer. In this project, crystallography and MD simulation provide protein structure for PELDOR spin label design, while PELDOR extends the conformation observed in densely packed crystal or *in silico* to dilute solution and detergent environments. Potentially PELDOR could also study the structural change of EGFR triggered by various factors, *e.g.* EGF or calmodulin. The dynamic process would be characterized via freeze-quenching technique (reviewed by Cherepanov and De Vries, 2004), by measuring distances from samples obtained at certain intervals over a time period, which is beyond the "snapshot" structure from a crystal and might support predictions from simulation. Therefore, this conformation study of EGFR is an exemplary case of integrative structural biology where individual techniques are incorporated to address challenging issues from multiple perspectives.

Concerning macromolecular distance measurement, fluorescence resonance energy transfer (FRET) has long been known to have the potential as a distance ruler. The optimal range for energy transfer is 10-80 Å (reviewed by Truong and Ikura, 2001), similar to the measurement range of PELDOR, and both methods are not constrained by protein size. However, FRET is more sensitive due to the strong fluorescence signal, which allows detection at single molecule resolution (Hellenkamp et al., 2018). Routinely running at room temperature on liquid solution, FRET in some cases can even monitor target protein from living cells in real time (Mohamed et al., 2018; reviewed by Stumpf and Hoffmann, 2016). The transfer efficiency of FRET is $1/[1+(r/R_0)^6]$ with r being the distance between fluorophores and R_0 being Förster radius which in part depends on the relative orientations between fluorophores (reviewed by Dimura et al., 2016). However, the value of orientation factor is assumed, which adds inaccuracy to the final distance (Schiemann et al., 2004). Also, fluorescence intensity during measurement may be influenced by other factors such as quenching, which requires additional experiment design and data interpretation to eliminate such effects. Furthermore, FRET fluorophores are often designed with longer linker than EPR spin labels, which makes distance between fluorescence centers less defined and more difficult to accurately derive (reviewed by Dimura et al., 2016). In contrast, PELDOR spectroscopy is less sensitive but more quantitative about the distance of interest. Unlike FRET, PELDOR measurement is typically performed on samples in frozen solution, with more protein molecules required for sufficient signal, *e.g.* 7.8×10^{14} molecules for PELDOR in this thesis. During the measurement, the spin orientation selectivity can be inspected by recording time traces at different frequency offsets and distance is calculated without assuming any parameters. Additionally, spin labels are generally smaller than fluorophore and coupled with shorter linker, which means less disruption to the protein structure and more accurate measurement of interspin distance due to structural rigidity.

In current PELDOR project of this thesis, accurate distance distribution is desired to enable a comparison with published crystal structure and simulations for identifying EGFR conformation. Moreover, adequate amount of proteins was enriched via purification procedure to compensate for the relatively low sensitivity of PELDOR. However, for the potential in-cell PELDOR measurement, higher sensitivity is appreciated due to the low EGFR abundance in cells. Also, traditional nitroxide labels face the challenge of cellular reducing environment. In that case, other strategies could serve as remedies, which will be discussed in next section. To sum up, PELDOR demands modest amount of protein without restriction of the size, requires sample preparation in either frozen solution or

other diverse environments, and it accurately measures distance in the range of 18-80 Å. Taken together, PELDOR serves as a suitable tool for EGFR conformation study.

4.2 Attempts of in-cell and in-membrane PELDOR

After the success of PELDOR measurements on purified proteins, it would be more interesting to see whether the same observations could follow in living cells or cellular membrane. Such data under the more native conditions would add more credit to the conclusions drawn so far from *in vitro* samples. Although in-cell PELDOR has been performed in several studies (Azarkh et al., 2011a; Theillet et al., 2016), it still faces several technical challenges such as cell-permeability of the compound, higher demand for labeling specificity in cells compared to purified samples, the reducing cellular environment which destabilizes common nitroxide spin labels and low copy number of target protein in cells. According to a preliminary in-house test, cellular EGFR autophosphorylation is effectively inhibited upon addition of PDsl to the cell culture medium, so PDsl is cell-permeable. Furthermore, clinical EGFR kinase inhibitors, many of which are structurally analogous to PD168393, have shown high selectivity in a chemical proteomics study (Klaeger et al., 2017). Concerning the spin label stability, it has been shown that six-membered-ring nitroxides are less resistant to reducing cellular condition than their five-membered-ring cousins (Azarkh et al., 2011b; Krstic et al., 2011). Additionally, it has been estimated that half-life of a five-membered-ring nitroxide, *i.e.* PROXYL, in oocytes of *Xenopus laevis* is roughly 50min (Igarashi et al., 2010). Thus spin labeling of cells is achievable if PDsl, which contains a five-membered-ring nitroxide, is incubated with cells and frozen in PELDOR tube as soon as possible. However, literature research indicates that the cellular copy number of EGFR may not be enough for PELDOR measurement. Based on classical saturation method and kinetic extrapolation technique, it has been revealed that A431, a human carcinoma cell line, contains approximately 2×10^6 copies of EGFR per cell (Barta et al., 2011; Novy et al., 2012). Considering that 2×10^7 cells reach maximum volume of 100µL that can be measured in PELDOR Q-band tube, the equivalent EGFR concentration is 0.66µM, which is still below the detection limit. Indeed, pilot experiments showed that by this approach no modulation depth for PDsl-labeled EGFR could be observed. So the plan of in-cell PELDOR was aborted.

Consequently, the alternative plan of in-membrane PELDOR was initiated. In this plan, heavy nuclei were removed as pellet from disrupted cells after 1000g centrifugation and cytosol was subsequently discarded after 75000g centrifugation as supernatant. Acquired pellet containing plasma membrane was resuspended and labeled by PDsl. It was hoped

that by discarding nuclei and cytosol, there would be additional space for more EGFR-containing membranes in the PELDOR tube. However, protein quantification via western blot indicates that enriched EGFR is not enough for PELDOR. During the sample resuspension, membrane sample has proven to be intractable and sticky, which explains the heavy loss of EGFR during this procedure.

So far the effort of measuring EGFR PELDOR under cellular or membrane conditions is stalled. One of the bottlenecks is the low EGFR concentration in enriched samples. This obstacle, however, could be in future overcome by using EGFR-overexpressing cell lines or improving membrane enrichment protocol. The other limiting factor is the stability of nitroxide spin label in reducing cellular environment. Until now several new spin labels have been developed and shown to be resistant towards reducing environment, such as Gd^{3+}-complex (Qi et al., 2014), triarylmethyl (trityl) radicals (Jassoy et al., 2017; Shevelev et al., 2014) and tetraethyl-substituted nitroxides (Jagtap et al., 2015). Due to the smaller size and improved reductive stability, tetraethyl-substituted nitroxides promise to be the proper candidate for in-cell EPR in future.

4.3 Influence of inhibitors over EGFR conformation equilibrium

Previous studies have shown asymmetric and symmetric dimers of EGFR kinase domain from either crystal structures (Jura et al., 2009; Zhang et al., 2006) or negative-stain EM images (Mi et al., 2011). Although interspin distances for both conformations are measurable in PELDOR, only one peak corresponding to asymmetric dimer was observed in frozen solution. Two possible reasons would be either PDsl does not fit into the ATP-binding sites of symmetric dimer, or it prefers to bind to asymmetric dimer and thereby shifts the conformation equilibrium. First possibility is ruled out by modeling via mtsslWizard and comparing volumes occupied by each spin label ensemble in symmetric and asymmetric dimers (Yin et al., 2017). Therefore, PD168393-based spin label disrupts EGFR conformation equilibrium in solution by stabilizing asymmetric dimer.

PD168393 belongs to one group of EGFR kinase inhibitors which are based on anilinoquinazoline backbone. Early studies of quinazoline inhibitors such as PD153035, AG-1478 and AG-1517 have indicated that they could increase EGFR dimerization (Arteaga et al., 1997; Lichtner et al., 2001). Several published crystal structures later reveal that two groups of quinazoline inhibitors exist, which favor either active or inactive kinase conformation of EGFR (Blair et al., 2007; Stamos et al., 2002; Wood et al., 2004). Among them are cocrystals of ICD(672-998) with PD168393 (PDB ID: 2J5F) or lapatinib (PDB ID: 1XKK; Wood et al., 2004).
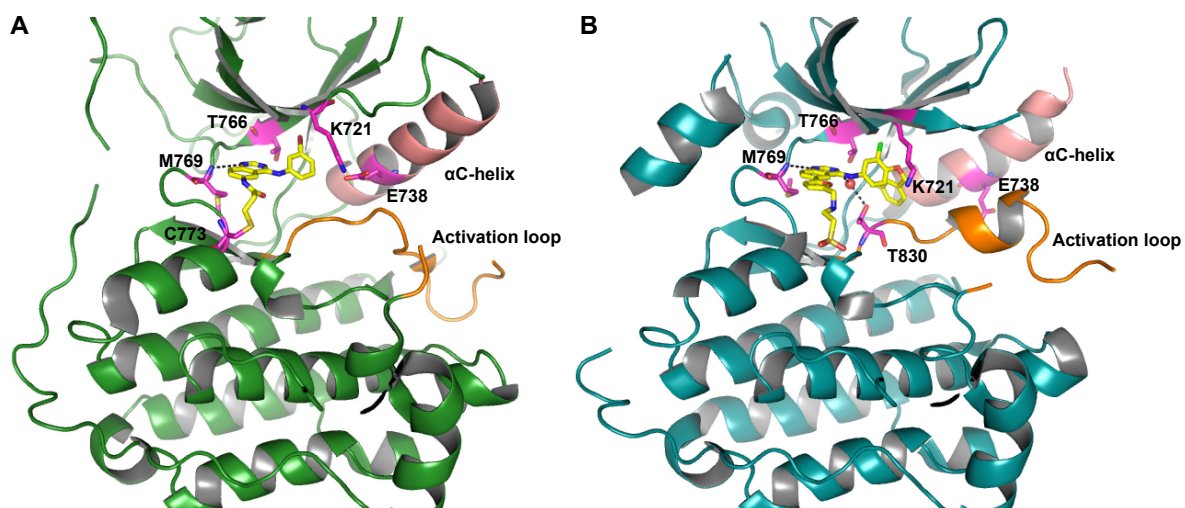
Figure 4.1 Crystal structures of EGFR kinase core in complex with PD168393 (A) or lapatinib (B). Kinase core, *i.e.* ICD(672-998), is depicted as dark green (A) or cyan (B) cartoon. PD168393 and lapatinib are shown as sticks with carbons colored yellow. Relevant residues are labeled and shown as sticks with carbons colored magenta. Activation loops are colored orange and αC-helices are shown as pink. Note the different orientations of αC-helices in (A) and (B). Fragment between L694 and T701 spanning the glycine-rich loop is removed to view ATP-binding site more clearly. Sequence numbering is based on human EGFR without signal peptide. Figure was prepared from PDB ID: 2J5F (A) and 1XKK (B) with PyMOL Molecular Graphics System.
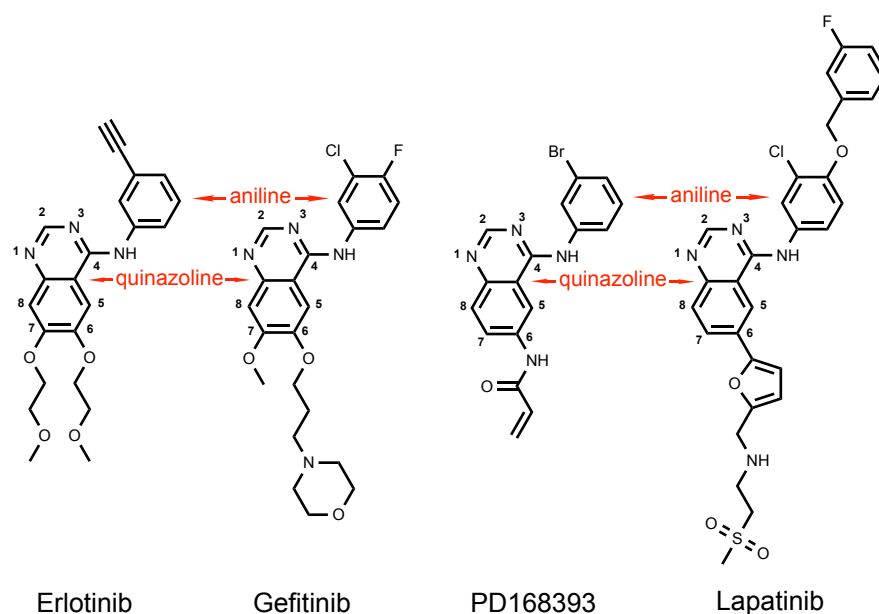
Figure 4.2 Chemical structures of four EGFR kinase inhibitors. Quinazoline backbones are numbered. Figure was prepared with ChemDraw.

For PD168393, the acrylamide warhead reacts with sulfhydryl group of C773 nearby via Michael addition, hence the irreversible binding. Meanwhile, quinazoline N1 forms a hydrogen bond with backbone amide of M769, while 4-aniline reaches into hydrophobic pocket around gatekeeper residue T766, stabilizing kinase domain in its active form,

evidenced by intact K721-E738 salt bridge and "αC-helix in" conformation, as shown in Figure 4.1 (A). Interaction mode of reversible inhibitor lapatinib retains N1 hydrogen bond, whereas quinazoline N3 forms a new water-mediated hydrogen bond with T830. The bulkier substitution on 4-aniline is accommodated in the hydrophobic pocket by "αC-helix out" conformation, in which αC-helix is oriented away from the ATP-binding site, as shown in Figure 4.1 (B). This αC-helix shift and loss of K721-E738 salt bridge indicate inactive kinase conformation and stay in clear contrast to the kinase structure in Figure 4.1 (A). Similarity and difference between these two groups of inhibitors are further illustrated in Figure 4.2, where structures of four compounds (erlotinib, gefitinib, PD168393 and lapatinib) are drawn and compared. On the common anilinoquinazoline backbone, N1 is conservative due to the critical hydrogen bond. Substitutions on C6 and C7 are quite diverse, because they face outwards from the ATP-binding cleft, therefore with less structural restriction. What distinguishes lapatinib from others is the substitution on 4-aniline. One additional aromatic ring is tethered to 4-aniline, and this flexible yet much bulkier substitution leads to distinct αC-helix conformation as explained in Figure 4.1.
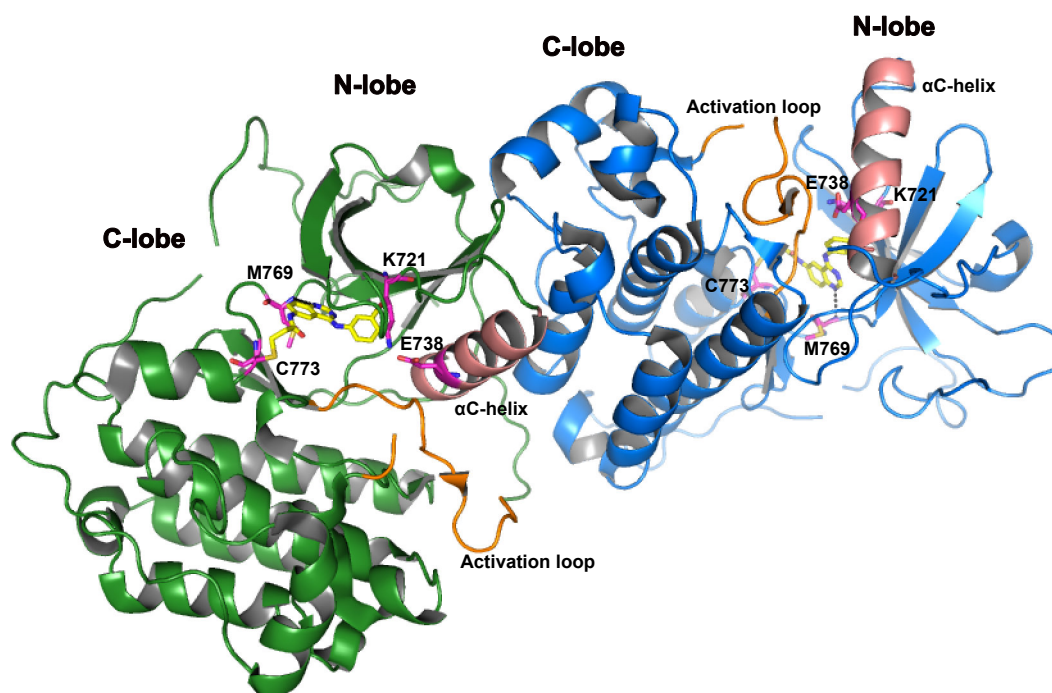
Figure 4.3 Overview of asymmetric dimer formed within kinase core-PD168393 cocrystal. Kinase core, *i.e.* ICD(672-998), is depicted as dark green or dark blue cartoon in either monomer. PD168393 is shown as sticks with carbons colored yellow. Relevant residues are labeled and shown as sticks with carbons colored magenta. Activation loops are colored orange and αC-helices are shown as pink. N-lobe and C-lobe of kinase domain are labeled. Sequence numbering is based on human EGFR without signal peptide. Figure was prepared from PDB ID: 2J5F with PyMOL Molecular Graphics System.
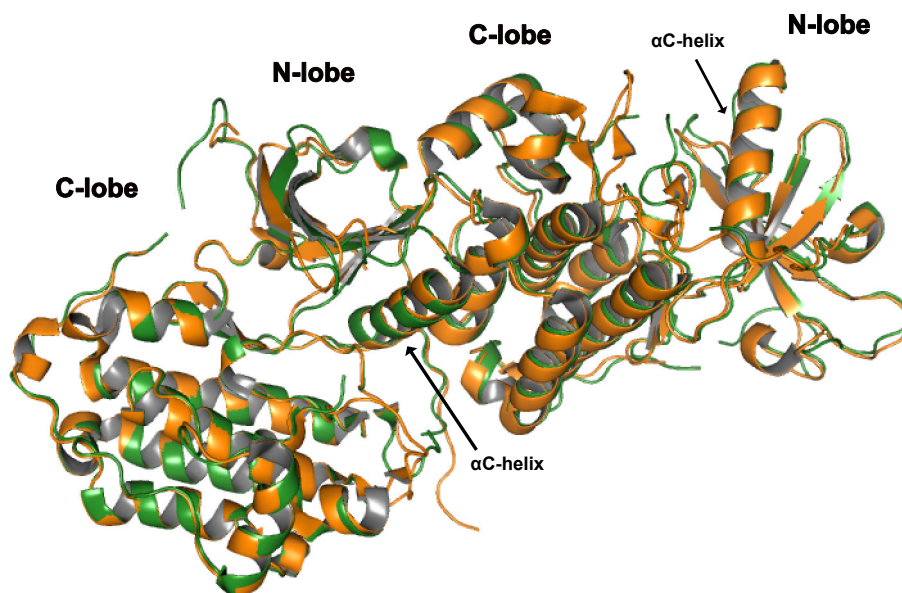
Figure 4.4 Superimposition of two kinase dimer structures. Asymmetric dimer from kinase core-PD168393 cocrystal (dark green) superimposed on that from kinase core-substrate analog cocrystal (orange). PD168393 and substrate analog are omitted. N-lobe and C-lobe of kinase domain as well as αC-helix are labeled. Figure was prepared from PDB ID: 2J5F and 2GS6 with PyMOL Molecular Graphics System.

According to the allosteric activation mechanism of EGFR, active kinase conformation is associated with asymmetric dimerization (reviewed by Bose and Zhang, 2009). Indeed, it has been reported that the first group of quinazoline inhibitors such as PD168393, which stabilizes active kinase conformation, can cause asymmetric kinase dimerization. In contrast, second group of inhibitors like lapatinib, which favors inactive kinase conformation, promotes no kinase dimerization (Lu et al., 2012). Especially, PD168393-induced dimerization is abolished by V924R mutation, with only monomeric EGFR on gel filtration chromatogram (Lu et al., 2012), which is consistent with our PELDOR and gel filtration data. The asymmetric dimer is also observed in the cocrystal with PD168393 (PDB ID: 2J5F), as illustrated in Figure 4.3, where two neighboring active kinase monomers are oriented in head-to-tail fashion, with activator C-lobe contacting receiver N-lobe. As a comparison, 2J5F is superimposed on another crystal structure (PDB ID: 2GS6), which contains ICD(672-998) in active kinase conformation in complex with ATP analog-peptide conjugate. As can be seen in Figure 4.4, two structures align well, especially so in the αC-helix region, thus confirming asymmetric dimerization. Conversely, no asymmetric dimer is observed in cocrystal with lapatinib (PDB ID: 1XKK). Although not causing aberrant kinase dimerization in the absence of EGF, lapatinib suppresses asymmetric dimerization in the presence of EGF, as revealed by negative-stain EM of a nearly full-length EGFR construct (Mi et al., 2011), probably by blocking formation of active kinase conformation with its bulky substitution on 4-aniline.

As EPR probe, spin label is supposed to be a neutral observer. It should remain unbiased towards conformations in equilibrium, reporting structures without disturbing them. Yet as discussed, it is not quite the case for PDsl, and probably not for lapatinib-based spin label either. On the other hand, PELDOR data from frozen solution accurately report asymmetric dimer observed in crystallization and EM studies, thereby confirming inhibitor-mediated dimerization discussed in other publications (Arteaga et al., 1997; Blair et al., 2007; Lichtner et al., 2001; Lu et al., 2012; Mi et al., 2011). From this perspective, effects of other EGFR inhibitors over its kinase conformation equilibrium can be investigated via PELDOR in future, with the help of IDSL. But still, an unbiased EPR probe is optimal for EGFR conformation study.
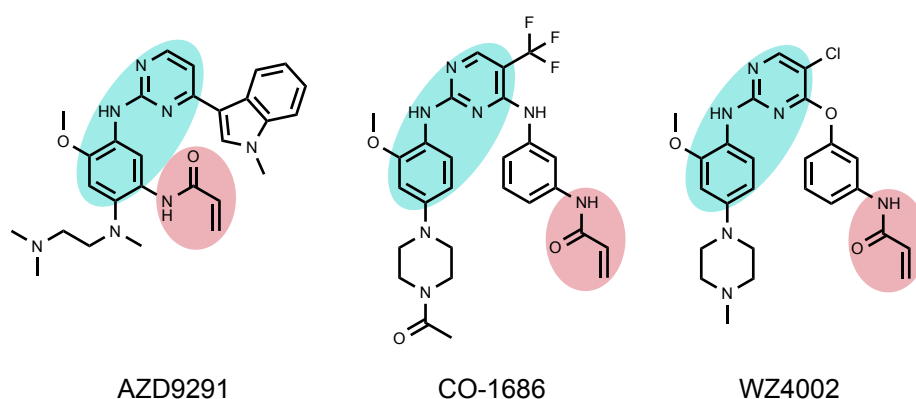
Figure 4.5 Chemical structures of representative third-generation EGFR kinase inhibitors. Anilinopyrimidine backbones are shaded in cyan, and acrylamide moieties are shaded in pink. Figure was prepared with ChemDraw.

At this moment, attention was drawn to the new generation of EGFR inhibitors. As is well known, erlotinib, gefitinib, lapatinib and PD168393 represent first- and second-generation EGFR kinase inhibitors, which share a common quinazoline core but differ in binding reversibility. During treatment with these inhibitors, patients often develop drug resistance via acquired secondary mutations like T766M. The third generation irreversible inhibitors, such as AZD9291, CO-1686 and WZ4002 (Figure 4.5), were developed with new anilinopyrimidine backbone to fight against such mutations (Cross et al., 2014; Engel et al., 2015; Walter et al., 2013; Zhou et al., 2009). They retain acrylamide moiety for the Michael addition with C773, but vacate the space occupied by 4-aniline on quinazoline backbone to tolerate the longer side chain of M766. As noticed in Figure 4.6 (A), the binding mode of WZ4002 (PDB ID: 3IKA; Zhou et al., 2009) is somewhat similar to that of the substrate analog in Figure 4.6 (C), both having no major contact with hydrophobic pocket adjacent to the gatekeeper residue T766. Therefore, question has arisen of whether this analogy means that WZ4002 has less disruption to EGFR kinase conformation equilibrium and qualifies as a candidate for IDSL probe design. Downside is

that the third generation inhibitors are mutant-selective, so it is doubted that they could bind and inhibit wildtype kinase equally effectively. Indeed as reported, quinazoline-based inhibitors are 100-fold more potent than WZ4002 when tested on wildtype kinase (Zhou et al., 2009). Therefore, as part of the follow-up plan for PDsl, inhibition potency of compounds like WZ8004 should be tested first. If it is indeed not so effective against wildtype kinase, incubation procedure could be optimized to achieve better binding of the inhibitor (more inhibitor input, longer incubation time, *etc*.). In the meantime, more effort should be spent on searching for derivatives of these inhibitors with better potency against wildtype kinase.

Kinase (T766M) & WZ4002

WT kinase & PD168393

WT kinase & ATP analog-peptide conjugate
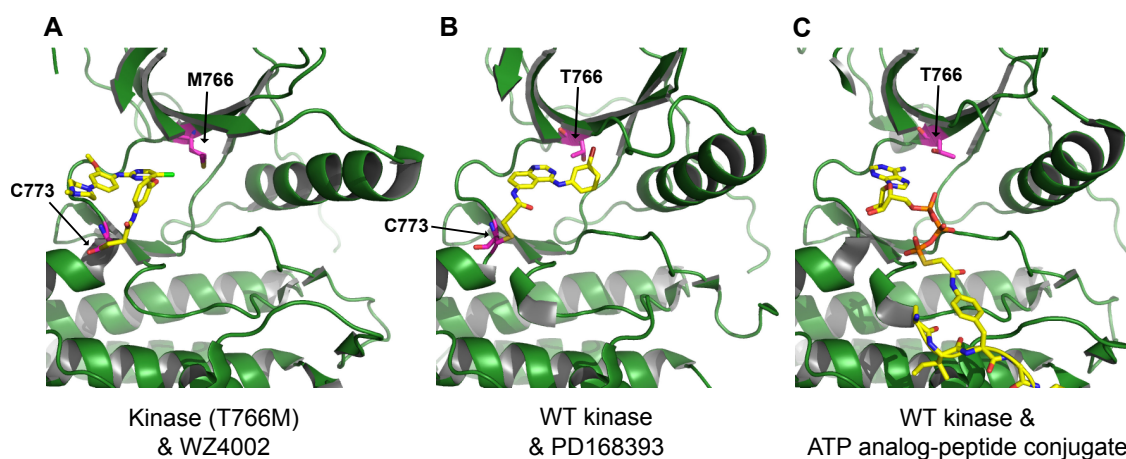
Figure 4.6 Comparison of binding modes of WZ4002 (A), PD168393 (B) and ATP analog-peptide conjugate (C) to kinase core wildtype or T766M mutant. Kinase core, *i.e.* ICD(672-998) is depicted as dark green cartoon. WZ4002, PD168393 and ATP analog-peptide conjugate are shown as sticks with carbons colored yellow. C773 and gatekeeper residue T766 or M766 are labeled and shown as sticks with carbons colored magenta. Fragment between L694 and T701 spanning the glycine-rich loop is removed to view ATP-binding site more clearly. Sequence numbering is based on human EGFR without signal peptide. Figure was prepared from PDB ID: 3IKA (A), 2J5F (B) and 2GS6 (C) with PyMOL Molecular Graphics System.

Another direction would be considering nucleotide-based probe. Since ATP is the natural ligand that fits into the activity site, it seems reasonable to design a probe simply with nitroxide radical tethered to ATP. Yet Km of ATP towards EGFR is in the range of 1-10μM (Honegger et al., 1988), which hinders application of spin-labeled ATP in PELDOR. As reported in our previous study, a spin-labeled ATP with spin marker attached to 2'-OH of the ribose shows no modulation of PELDOR time trace (Yin et al., 2018). A second nucleotide-based labeling strategy is built upon biotinylated acyl ATP or ADP (Figure 4.7), in which an acyl phosphate is situated between nucleotide and biotin tag. Nucleotide functions as binding moiety towards kinase domain, while biotin tag is the labeling group which can be switched into nitroxide radical for EPR measurement. Carbonyl carbon of

the acyl phosphate acts as electrophile and attacks specifically ε-amino group of a conservative lysine near the ATP-binding cleft, after which nucleotide moiety leaves and an amide bond forms between biotin tag and lysine residue (Patricelli et al., 2007). In Figure 4.8 (A), crystal structure of EGFR kinase with ATP analog-peptide conjugate (PDB ID: 2GS6) shows that only one lysine residue K721 is close to ATP β- or γ-phosphate, with around 8Å between terminal nitrogen of lysine side chain and phosphorus of γ-phosphate. This lysine is conserved in CDK2 (PDB ID: 1HCK; Schulze-Gahmen et al., 1996) as K33 in Figure 4.8 (B), with another lysine K129 located in the proximity as well. In EGFR, however, an arginine residue R817 occupies the location corresponding to K129 of CDK2. Further inspection of other kinase structures reveals that this arginine is conserved across tyrosine kinases such as insulin receptor kinase (PDB ID: 3BU5; Wu et al., 2008) or ABL1 kinase (PDB ID: 3CS9; Weisberg et al., 2005), whereas K129 of CDK2 is conserved in serine/threonine kinases like MAP kinase (PDB ID: 4ZSG; Chen et al., 2016). In contrast, K721 of EGFR is conserved in both tyrosine and serine/threonine kinases. However, K721 is known to be critical for maintaining active EGFR kinase conformation as discussed before, so modification of it would be detrimental to kinase activity, which makes this labeling strategy not acceptable for EGFR. On the other hand, acyl ATP probe could in future be re-engineered for EPR conformation studies of serine/threonine kinase, considering that one additional lysine like K129 in CDK2 is available and it can be selectively labeled due to closer proximity to γ-phosphate compared to K33 (Patricelli et al., 2007).
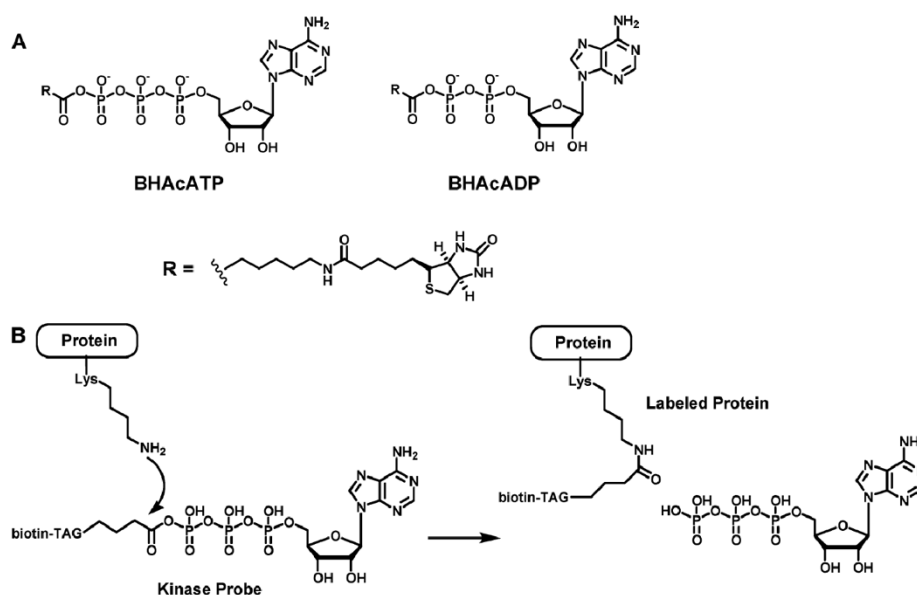
Figure 4.7 Chemical structures and labeling mechanism of ATP and ADP-based probe. BHAcATP: (+)-Biotin-Hex-acyl-ATP, BHAcADP: (+)-Biotin-Hex-acyl-ADP. Figure is adapted from Patricelli et al., 2007.
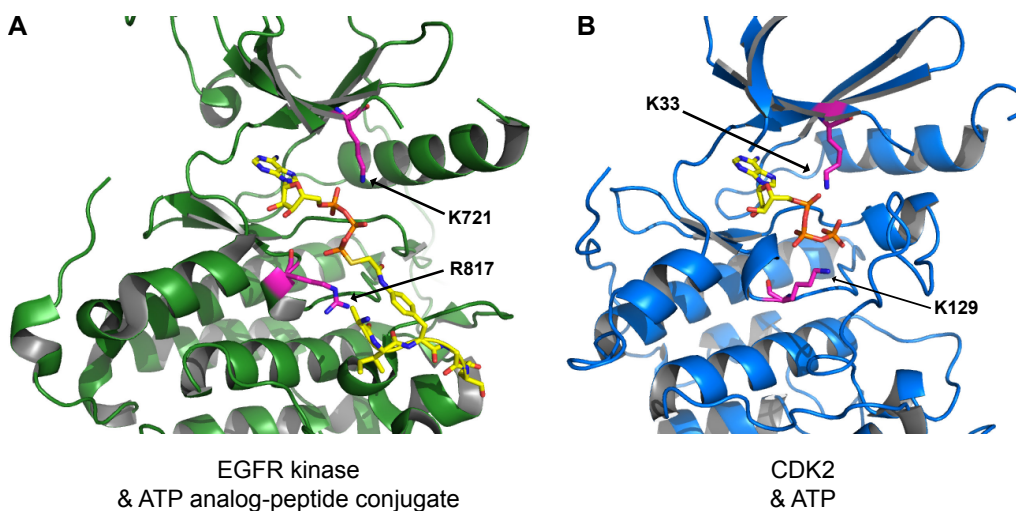
Figure 4.8 Comparison of lysine and arginine residues adjacent to ATP-binding sites in EGFR (A) and CDK2 (B). EGFR and CDK2 kinases are depicted as dark green (A) and dark blue (B) cartoons, respectively. ATP analog-peptide conjugate (A) and ATP (B) are shown as sticks with carbons colored yellow. Neighboring lysine and arginine residues are labeled and shown as sticks with carbons colored magenta. Glycine-rich loop is removed to view ATP-binding cleft more clearly. Figure was prepared from PDB ID: 2GS6 (A) and 1HCK (B) with PyMOL Molecular Graphics System.

4.4 Benefits and limitations of IDSL

In this thesis project, inhibitor-directed spin label rose from the failure of site-directed spin label and witnessed the convenient labeling of EGFR for conformation study. However, there are certain issues looming on the horizon which restrict its further application. In this section, both benefits and limitations of IDSL will be discussed, with possible solutions or alternatives proposed as well.

One of the major benefits from IDSL is that it liberates scientists from laborious trials and errors of mutagenesis work in SDSL. Usage of conventional spin labels like MTSL requires optimal cysteine mutations to be sought out for the downstream labeling. For cysteine-rich proteins, this means mutations of native cysteines to serine or alanine, and mutation to introduce new cysteine as the labeling site. For instance, intracellular domain of EGFR contains six native cysteines in kinase domain and three cysteines in C-terminal tail. The complex cysteine mutation combination is not only time-consuming but also possibly impairing protein activity and stability. Furthermore, natural cysteines may play crucial roles structurally or functionally, such as facilitating protein folding by forming disulfide bridges, or functioning as the catalytic residue in cysteine protease. Cysteine residues as such are simply untouchable during the mutagenesis, thus posing a challenge for EPR spin labeling. Alternatively, one could introduce unnatural amino acid (UAA) as

the labeling site, such as *p*-acetylphenylalanine (Young et al., 2010). However, this approach so far is mainly applied in *E.coli* expression system, while for target protein like EGFR, only eukaryotic cell culture is acceptable for appropriate expression. In addition, mutagenesis is still required, of which the effect on protein expression, folding and activity is unknown. IDSL, on the other hand, rests upon the high affinity of available inhibitors towards wildtype protein, so no mutation is necessary. Another benefit of IDSL is that a sizeable number of inhibitors have already been commercialized and well characterized, which is convenient for repurposing them as EPR spin labels. For example, protein kinase inhibitor has become hotspot for pharmaceutical development since the approval of Fasudil for cerebral vasospasm in Japan in 1995 (reviewed by Fabbro et al., 2015; Shibuya and Suzuki, 1993). Protein kinases play such a pivotal role in cell signaling pathway that their overexpression and overactivation are often linked to various cancer conditions, which makes kinase inhibitor one of the intensively investigated branches of anticancer therapeutics. So far, there are more than 250 candidate kinase inhibitors in clinical development (Klaeger et al., 2017), with many among them already cocrystallized with target kinases. These kinase-inhibitor complex structures provide molecular details about interaction mode which facilitate inhibitor-based spin label design from many perspectives, such as considering where to attach spin label on the inhibitor, and evaluating theoretical interspin distance. EGFR studied in this thesis project is one such research hotspot. With the rich collection of crystal structures and publications about EGFR inhibitors, it is convenient to initiate IDSL for PELDOR. Also, covalent inhibitor is preferred for IDSL application, because irreversible binding eliminates the possibility of spin label being washed off during the subsequent purification steps. However, if the affinity of noncovalent inhibitor is within nanomolar range, labeling procedure with limited washing step could be tolerated. Additionally, IDSL is not confined to the conformation study of kinases. Instead it can be broadly applied to available covalent inhibitors against many protein families, such as omeprazole (Prilosec, AstraZeneca) against H^+/K^+ ATPase and clopidogrel (Plavix, BMS/Sanofi) against $P2Y_{12}$ receptor (reviewed by Singh et al., 2011). Furthermore, it should be mentioned that covalent IDSL is the extension of a more general concept, namely, affinity labeling. It is originally applied to enzyme or antibody labeling, where noncovalent binding moiety of the probe is attracted to catalytic pocket, followed by reaction moiety attacking the nucleophilic side chain of an amino acid, leading to irreversibly modifying protein with fluorophore or radioactive label. By labeling amino acids nearby with different affinity labels, activity sites of the enzyme and antibody were identified and mapped (Wofsy et al., 1962). Applications of affinity labeling as such can be searched for and easily repurposed for PELDOR labeling. Here in this thesis, the idea of a fluorescent affinity label (Blair et al., 2007) is borrowed and converted for EPR purpose, in

which amino acid near the ATP-binding site (C773) was specifically labeled by nitroxide spin.

Despite of the benefits in PELDOR study, IDSL is not without its limitation. In previous section, effects of PDsl over conformational equilibrium of EGFR has been discussed, with alternatives proposed and discussed. In general, influence of inhibitor over protein structure and oligomerization should be carefully investigated before designing inhibitor-based spin label. Additionally, PELDOR is limited to intermolecular distance measurement, if each protein molecule has only one binding site for the inhibitor. To achieve intramolecular measurement with IDSL, second labeling site on the same protein has to be established. If the protein of interest contains few natural cysteine residues, conventional SDSL can be combined with IDSL for the protein labeling. Likewise, UAA-based labeling should be considered when target protein is compatible with its workflow. If the protein contains multiple domains, inhibitors targeting other domains serve as the natural candidates for second label. For some proteins with paramagnetic cofactors such as Cu^{2+}, metal ions can be utilized to measure ion-nitroxide distance in PELDOR (Abdullin et al., 2015). Another concern with monolabeling of protein by IDSL is that monomeric protein, one basic ingredient of oligomerization equilibrium, is missing in the final PELDOR results, which is unfortunate for conformational study involving protein oligomerization. This lack of monomer information could be remedied by combining analytical gel filtration assay to characterize protein oligomerization status. If possible, size exclusion chromatography equipped with multi-angle light scattering (MALS) detectors would even measure molecular weights of eluted species in gel filtration. However, it is not guaranteed that the size exclusion column can thoroughly resolve different oligomeric species, as noticed in the EGFR study. In that case, it is difficult for MALS analysis to calculate molecular weight from a mixed gel filtration peak. Moreover, during drug development, candidate compound could hit more than one targets, a phenomenon termed as polypharmacology (reviewed by Reddy and Zhang, 2013). This off-target effect also poses a challenge for IDSL to be further applied in in-cell PELDOR, where inhibitor specificity is demanded to selectively label the target in the surroundings full of miscellaneous proteins. Although EGFR inhibitors have already shown high selectivity in chemical proteomics study, it is not generally the case for other tyrosine kinase inhibitors (Klaeger et al., 2017). Therefore, more careful literature research on inhibitor specificity is recommended before starting new inhibitor-based spin label design for in-cell PELDOR.

4.5 ICD, Sec7, calmodulin and lipid: a tale of four

In section 3.2, MST is successfully applied to protein interaction study of EGFR. Emerging evidence from current study has repeatedly shown that ARNO-Sec7 domain binds to JM segment of EGFR-ICD. In the meantime, it is confirmed that calmodulin interacts with the very same JM segment in a Ca^{2+}-dependent manner, and that it effectively competes with Sec7 for JM binding. Yet the story did not end there. The fruitful collaboration with NMR group of Manuel Etzkorn brings some intriguing insights into Sec7:JM interaction, with the role of lipid also clarified. A tale of four has gradually unfolded.

The NMR data of Sec7 and JM peptide was collected and analyzed by Dr. Manuel Etzkorn and Dr. Aldino Viegas. Not only do they confirm Sec7:JM interaction, but also identify key residues involved in this interaction. In the titration assay of Sec7 with JM, concentration of ^{15}N-labeled Sec7 was maintained constantly at 60µM while concentration of unlabeled JM varied from 0 to 420µM. ^{1}H-^{15}N HSQC spectra were acquired, after which combined chemical shift changes were aligned with Sec7 sequence to map amino acids relevant to the binding. Based on the chemical shift perturbations upon JM binding, involved residues are highlighted as dark green in Figure 4.9.
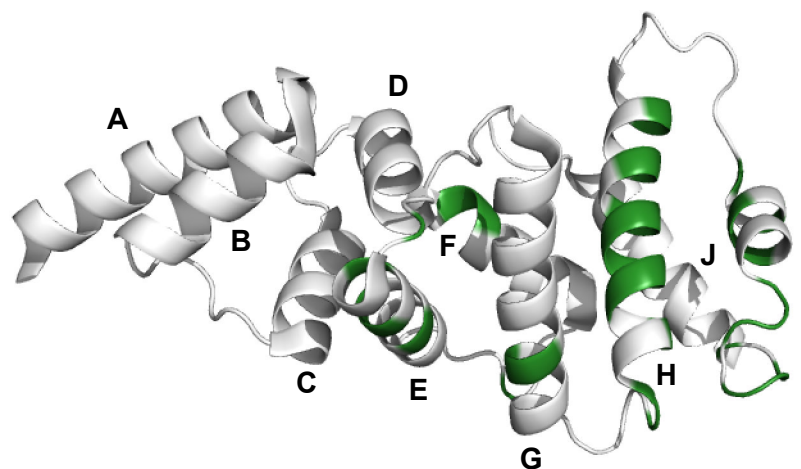
Figure 4.9 Affected residues on ARNO-Sec7 due to Sec7:JM interaction. Crystal structure of ARNO-Sec7 (PDB ID: 4JMI; Rouhana et al., 2013) is depicted as grey cartoon. Amino acids with major changes of combined chemical shift are colored dark green. Helices are denoted alphabetically according to Mossessova et al., 1998. Figure was prepared based on data from Aldino Viegas.

Interestingly, these amino acids partially overlap with the binding site of Arf1 (Cherfils et al., 1998; Mossessova et al., 1998; Renault et al., 2003). Especially, αH-helix is a major part of the binding site for both Arf1 and JM, and is covered by Sec7-PH linker and C-terminal helix in autoinhibited ARNO. Deletion of PBR in ARNOΔPBR relieves the

autoinhibition mechanism and thus opens the binding site for Arf1 docking (DiNitto et al., 2007). The masking of αH-helix in autoinhibited state and exposure of it in open state, fit well with MST data in section 3.2.3.1, where ARNOΔPBR shows binding towards EGFR-ICD while ARNO not. Furthermore, when four residues on the hydrophobic patch of αH-helix are collectively mutated into alanines, the binding capacity of Sec7 towards JM or ICD is lost in MST assay. This is also in agreement with NMR data, and further supports the mapping results from HSQC.

Meanwhile, titration assay of JM with Sec7 was also performed, with ^{15}N-labeled JM at 40μM and unlabeled Sec7 varying from 0 to 280μM. ^{1}H-^{15}N HSQC spectra locate major chemical shift perturbations to the N-terminus of JM segment, *i.e.* JM-A. Since JM-A is rich in basic residues (8 arginine residues and 1 lysine residue) that are positively charged at physiological pH, question has been raised that whether Sec7:JM interaction could be purely caused by electrostatic attraction between JM-A and negatively charged surface of Sec7. Therefore, a scrambled version of JM with random sequence, JMsc, was also tested via ^{1}H-^{15}N HSQC, where JMsc in the presence of three-fold excess of Sec7 shows no chemical shift change. This corroborates the MST data in section 3.2.3.3, and indicates that Sec7:JM interaction is sequence-dependent.

As a known EGFR modulator, calmodulin also recognizes JM segment upon activation (Aifa et al., 2002; Martín-Nieto and Villalobo, 1998). In this thesis project, calmodulin:JM interaction has been confirmed by both MST and NMR. Data acquired from NMR reveal that the affected area upon calmodulin binding is JM-A segment, practically identical to the one in Sec7 interaction. Additionally, it is indicated by NMR that calmodulin:JM interaction is stronger than Sec7:JM, which supports the similar observation from MST assay and explains the efficiency of calmodulin in competition assay (see section 3.2.4.2).

However, Sec7 and calmodulin are not the only factors at play. Acidic (or anionic) phospholipid at the inner leaflet of plasma membrane has also been shown to interact with basic JM segment (Aifa et al., 2002; Sato et al., 2006; Sengupta et al., 2009). To evaluate anionic lipid interaction with JM, and to compare it with Sec7:JM or calmodulin:JM interaction, an *in vitro* lipid bilayer system, *i.e.* nanodisc was introduced in the NMR study. Two types of phospholipids were utilized to study the effect of net charge on JM:lipid interaction, namely neutral phospholipid POPC and anionic phospholipid POPS. Nanodisc assemblies with 100% POPC and 70% POPC + 30% POPS were compared, which shows that only anionic phospholipid-containing nanodisc induces chemical shift perturbations on JM resonance. Moreover, residues with significant perturbations are from JM-A segment again. Interestingly, lipid:JM interaction is strengthened once the negative

charge density on nanodisc is increased by using 50% anionic phospholipid, confirming the electrostatic nature of this interaction. Notably, the exact JM region perturbed by lipid is slightly shorter than those by calmodulin and Sec7, with E661-E666 not affected upon lipid binding. This difference was conveniently exploited to study the competition between lipid and Sec7 for JM binding. When mixture of JM and nanodisc with 50% POPC + 50% POPS was combined with Sec7, distinct chemical shift perturbations of E661-E666 were observed and they are consistent with the perturbations from Sec7:JM interaction. Although Sec7 could not fully shift the binding equilibrium from lipid:JM towards Sec7:JM, this still indicates that Sec7 is capable of interacting with JM in the proximity of nanodisc with a high content of anionic phospholipids. Additionally, although ICD:Sec7 interaction seems quite weak according to MST data, ICD:ARNOΔPBR shows four-fold stronger binding affinity (Kd of 49.2μM *vs*. 12.5μM). The unclear origin of this additional affinity requires further investigation in future. Nevertheless, it means that ARNO, once relieved from its autoinhibition, has the potential of binding to EGFR with a reasonable affinity. Furthermore, it has been shown that EGFR cluster overlaps with PIP2 cluster on cell membrane (Wang et al., 2014). Since PIP2 could enrich ARNO in membrane proximal region through its recognition of ARNO-PH domain (reviewed by Kolanus, 2007), local concentration of ARNO is increased, which further promotes ARNO interaction with EGFR-ICD. Taken together, these data add more credit to the physiological relevance of Sec7:JM interaction at plasma membrane.

4.6 Physiological implication of Sec7:ICD interaction

So far the story has been told around two main characters, namely Sec7 and ICD, yet it is not fully understood in what way the Sec7:ICD interaction is relevant to EGFR regulation under physiological conditions. EGFR has long been known to play a key role in cell growth, proliferation and migration, and its upregulation is frequently involved in non-small cell lung cancer (reviewed by Soria et al., 2012), breast cancer (reviewed by Masuda et al., 2012) and colorectal cancer (reviewed by Markman et al., 2010). Meanwhile, ARNO overexpression has also been reported in human colorectal adenocarcinoma, and it correlates with increased EGFR signal (Pan et al., 2014). Also, SecinH3-treated lung cancer cells show reduced EGFR signaling and inhibited proliferation (Bill et al., 2012). Therefore, it is reasonable to connect Sec7:ICD interaction to EGFR overactivation and carcinogenesis. Here two hypotheses are proposed concerning the potential role(s) played by ARNO in EGFR signaling.

The first hypothesis, dubbed as "endocytosis hypothesis", is based upon several observations from other studies: (1) phosphorylated EGFR can recognize and activate

GEP100, which in turn activates Arf6 as its specific GEF and induces invasion of tumor cells (Morishige et al., 2008; Sabe et al., 2009). (2) Activated Arf6 can recruit ARNO onto plasma membrane via Arf6-PH interaction and facilitates relieving the autoinhibition mechanism of ARNO (Cohen et al., 2007; Stalder et al., 2011). (3) Activated Arf6 regulates plasma membrane/endosome trafficking via clathrin-independent or clathrin-dependent pathway (reviewed by D'Souza-Schorey and Chavrier, 2006). (4) During endocytosis, activated EGFR can maintain long-lived signaling in early endosome (reviewed by Murphy et al., 2009; reviewed by Villaseñor et al., 2016; Tomshine et al., 2009; Wang et al., 2002), a phenomenon known as "endosomal signaling". Thus upon EGF stimulation, EGFR-GEP100-Arf6 cascade is triggered. Arf6, which is activated as Arf6-GTP, recruits and activates ARNO which directly interacts with EGFR on its JM segment. Subsequently, Arf6-GTP mediates endocytosis which through its ARNO connection carries EGFR as well. Internalized EGFR is harbored in endosome vesicles, and continues promoting downstream signaling pathways. Notably, similar model has been applied to Wnt/β-catenin pathway, where Wnt-induced Arf6 activation drives internalization of c-Met, also a receptor tyrosine kinase, which in signaling endosome further mediates downstream signaling to ERK and increases cytoplasmic β-catenin (Pellon-Cardenas et al., 2013). However, for EGFR, it is still in question whether signaling from endosome is more potent than that from plasma membrane. Therefore, endocytosis hypothesis should be tested, *e.g.* via mutagenesis and pharmacological approaches. The former approach involves mutants such as dominant-negative GDP-bound Arf6(T27N) and constitutively active GTP-bound Arf6(Q67L). The latter approach involves compounds such as SecinH3, a GEF inhibitor targeting Sec7 domain. Effects of Arf6 or ARNO on EGFR endocytosis and signaling can be evaluated via cellular assay by comparing EGFR autophosphorylation, internalization and downstream signaling between Arf6 mutants and wildtype Arf6, or between SecinH3-treated cells and DMSO-treated cells.

Considering the critical function of JM in EGFR allosteric activation (Jura et al., 2009; Red Brewer et al., 2009), it is reasonable to postulate that Sec7:JM interaction might play a role more direct than mere "middle man" in Arf6-orchestrated endocytosis. The inspiration for second hypothesis, named "priming hypothesis", comes from electrostatic engine model suggested as EGFR activation mechanism (McLaughlin et al., 2005). During early days of research, calmodulin was claimed to inhibit EGFR activity, based on detergent-solubilized EGFR sample purified via calmodulin affinity chromatography (BENGURÍA et al., 1995; San Jose et al., 1992). However, recent data based on living cells have shown the activating role of calmodulin, which is confirmed by multiple studies (Li et al., 2004, 2012; McLaughlin et al., 2005; Sengupta et al., 2007). To explain this activating function,

Stuart McLaughlin proposed electrostatic engine model. According to this model, during resting phase, positively charged parts of JM and kinase domain are tethered to plasma membrane, where the inner leaflet contains anionic phospholipid such as PS or PIP2, and this electrostatic interaction impedes free movement of JM and kinase domain for EGFR activation. Upon EGF stimulation, calmodulin is activated via PLCγ and phospholipase A2 (reviewed by Sánchez-González et al., 2010), and by directly binding to JM, the negatively charged calmodulin (pI 4.09) reverses the charge on JM. Newly formed calmodulin:JM complex is repelled by anionic phospholipid in the proximity. Both JM and kinase domain fall off the plasma membrane, which allows them to rearrange into active conformation (McLaughlin et al., 2005).

The priming hypothesis assumes a similar role for ARNO-Sec7 in EGFR activation. However, the difference is that JM-binding affinities of Sec7 and ARNOΔPBR are weaker than calmodulin, and it has been shown by NMR that Sec7 can only partially pull JM out of the nanodisc, so Sec7 alone probably cannot fulfil full function as an EGFR activator in this model. Therefore, it is further postulated that Sec7 binds to JM prior to calmodulin binding. Sec7 binding releases JM partly from plasma membrane, and thereby primes it for subsequent calmodulin binding and full release from lipid environment, hence the name "priming" for this hypothesis. However, it is unclear in cells whether Sec7-binding actually happens before calmodulin-binding event. Studies have shown that after EGF stimulation the peaking time of cellular free Ca^{2+} varies between 1 and 3min (Hughes et al., 1991; Nojiri and Hoek, 2000; Uyemura et al., 2005), whereas the peaking time for Arf6-GTP is 30sec (Boulay et al., 2008). However, the locality of Ca^{2+} signaling, due to its steep concentration gradient around Ca^{2+} pool (Clapham, 2007), hinders reliable estimation of the time needed for calmodulin activation in the proximity of plasma membrane. Also, the time for Arf6-GTP to recruit and activate ARNO is not accounted for in the study. Therefore, this comparison based on literature is inconclusive. Additionally, it is possible that on plasma membrane there is pre-bound ARNO independent of EGF stimulation, such as ARNO enriched by PIP2 cluster as mentioned in previous section. In that case, ARNO can function as a priming modulator as well. Practically, priming hypothesis can be verified as follows. First, the role of Sec7 in electrostatic engine model could be verified by using nanodisc containing full length EGFR. Effects of Sec7 over EGFR activity are then compared between controls with varying anionic lipid compositions in nanodisc or salt concentrations in buffer, since these factors influence the electrostatic attraction. Second, to investigate the priming role of ARNO, it should be examined whether calmodulin activation of EGFR depends on ARNO. To serve this purpose, chemical inhibition of ARNO-Sec7 via SecinH3 treatment or gene silencing via RNAi (RNA

interference) could be utilized to suppress ARNO-Sec7 function or expression in cells. Thereafter, comparison of the calmodulin activation effects between SecinH3-/RNAi-treated and untreated cells will reveal whether ARNO renders EGFR activatable to calmodulin, *i.e.* whether there is a priming effect.

5 Materials and Methods

5.1 Materials

5.1.1 Equipment

Name	Manufacturer
ÄKTA FPLC/Pure/avant	GE Healthcare
Analytical balance CPA324S	Sartorius
Autoclave VX-150	Systec
Balance BL 1500S	Sartorius
BioSpectrometer	Eppendorf
Block heater	Stuart Equipment
Blotting paper	Macherey-Nagel
Centrifuge 5417C	Eppendorf
Centrifuge 5427 R	Eppendorf
Centrifuge 5430	Eppendorf
Centrifuge 5804 R	Eppendorf
Centrifuge Multifuge 3 S-R Heraeus	Thermo Fisher Scientific
Chromatography cooling cabinet UNICHROMAT 1500 "PRO"	UniEquip
Desalting column HiPrep^{TM} 26/10	GE Healthcare
Electrophoresis chambers	BIO-RAD
Electrophoresis power supply E802/E865	CONSORT
French press cell disrupter	Thermo Fisher Scientific
High-speed centrifuge Avanti J-26S XP	Beckman Coulter
High-speed centrifuge tube Nalgene^{TM} Oak Ridge for JA-25.50	Thermo Fisher Scientific
Incubation shaker Multitron Cell	INFORS HT
Incubator Heraeus HERAcell 240	Thermo Fisher Scientific
Incubator Heratherm^{TM}	Thermo Fisher Scientific
Incubator shaker Innova® 40	New Brunswick Scientific
Incubator shaker Innova® 42	New Brunswick Scientific
JA-25.50 Rotor Fixed-Angle	Beckman Coulter
JLA-8.1000 Rotor J-LITE® Fixed-Angle	Beckman Coulter
Magnetic stirrer IKAMAG®	IKA
Monolith NT.115	NanoTemper
Multi-Axle Roller-Mixer RM5-30V	neoLab
Multipette® M4	Eppendorf
Multipette® plus	Eppendorf
Odyssey Imaging System	LI-COR
Orbital shaker KS501 digital	IKA
Overhead tumbler PTR-30	Grant-bio
Peristaltic pump 2115 Multiperpex	LKB Bromma

pH meter FiveEasy™	METTLER TOLEDO
Pipette controller accu-jet® pro	BRAND
Pipette Research® plus	Eppendorf
Plate reader Infinite®M1000 PRO	Tecan
Polypropylene Bottle Assembly for JLA-8.1000	Beckman Coulter
SDS-PAGE gel casting stands/frames/combs	BIO-RAD
SEC column HiLoad™ 16/600 Superdex™ 200 pg	GE Healthcare
SEC column HiLoad™ 16/600 Superdex™ 30 pg	GE Healthcare
SEC column HiLoad™ 16/600 Superdex™ 75 pg	GE Healthcare
SEC column Superdex™ 200 10/300 GL	GE Healthcare
Semi-dry blotter"Pegasus"	PHASE
Semi-dry transfer cell TRANS-BLOT® SD	BIO-RAD
Spectrophotometer NanoDrop 2000c	Thermo Fisher Scientific
Thermomixer	HLC
Thermomixer comfort	Eppendorf
Vibra-Cell™ Ultrasonic Liquid Processor VCX400	Sonics & Materials
Vortex 4 basic	IKA
Water bath	GFL®
Water Purification System Barnstead™ MicroPure™	Thermo Fisher Scientific

5.1.2 Chemicals and reagents

Name	**Supplier**
Acetic acid, ROTIPURAN® 100%, p.a.	Carl Roth
Acrylamide/Bis-solution(37.5:1) Rotiphorese® Gel 30	Carl Roth
Adenosine 5'-triphosphoric acid disodium salt, ≥ 98%	PanReac AppliChem
AEBSF	SERVA
Alexa Fluor™ 647 NHS Ester	Thermo Fisher Scientific
APS, ≥ 98%, p.a., ACS	Carl Roth
Bestatin	SERVA
Calcium chloride, ≥ 98%, dehydrated, powder	Carl Roth
Coomassie® Brilliant Blue G 250	SERVA
d-Desthiobiotin, ≥ 98%(TLC)	Sigma-Aldrich
D(+)-Maltose monohydrate, ≥ 95%, for biochemistry	Carl Roth
Deuterium oxide, 99.8 atom% D, D_2O	Sigma-Aldrich
Dimethyl sulfoxide, 99.9+%, ACS reagent, ACROS Organic™	Thermo Fisher Scientific
DPBS, no calcium, no magnesium, gibco®	Thermo Fisher Scientific
DTT, ≥ 99%, p.a.	Carl Roth
E-64	SERVA
EDTA, p.a.	PanReac AppliChem
EGTA, ≥ 99%, p.a.	Carl Roth
Ethylene glycol-d6 99%	Deutero

Gel Filtration Markers Kit MWGF1000	Sigma-Aldrich
Glycerol, ROTIPURAN® ≥ 99.5%, p.a., anhydrous	Carl Roth
Glycine, PUFFERAN® ≥ 99%, p.a.	Carl Roth
Guanosine 5'-diphosphate sodium salt, ≥ 96% (HPLC)	Sigma-Aldrich
Guanosine 5'-triphosphate sodium salt hydrate, ≥ 95% (HPLC)	Sigma-Aldrich
HEPES, Buffer grade	PanReac AppliChem
Hydrochloric acid 37% AnalaR NORMAPUR®	VWR Chemicals
Imidazole	PanReac AppliChem
IPTG, 99%	Carbolution
Kanamycin sulphate, ≥ 750 I.U./mg	Carl Roth
LB Broth (Lennox) for molecular biology	Carl Roth
Magnesium chloride Hexahydrate, ≥ 99.0%, p.a.	Fluka
Monolith Protein Labeling Kit Red-NHS 2nd Generation	NanoTemper
PD168393, ≥ 98% (HPLC)	Sigma-Aldrich
PD168393sl (PDsl)	Synthesized by Jeffrey Hannam
Phosphoramidon	SERVA
Pierce™ Centrifuge Columns and Caps, 10mL	Thermo Fisher Scientific
Protein Ladder, PageRuler™ Plus, prestained	Thermo Fisher Scientific
SDS, grained pure	PanReac AppliChem
Sodium chloride, ≥ 99.5%, p.a., ACS, ISO	Carl Roth
Sodium hydroxide, ≥ 98%, p.a., ISO, in pellets	Carl Roth
Sodium orthovanadate	PanReac AppliChem
TEMED, ≥ 99%, p.a., for electrophoresis	Carl Roth
TES, PUFFERAN® ≥ 99%	Carl Roth
TRIS, PUFFERAN® ≥ 99.9%, ultra quality	Carl Roth
Triton® X-100	PanReac AppliChem
Tween® 20	PanReac AppliChem
β-Mercaptoethanol, Molecular biology grade	PanReac AppliChem

5.1.3 Consumables

Name	Manufacturer
Amylose Resin	NEB
Assay plate 96-well, Half Area, Black Flat Bottom, Non-Binding Surface	Corning
Blotting paper	Macherey-Nagel
Capillaries for MST, Monolith™ NT.115, premium	NanoTemper
Centrifugal concentrator Vivaspin® Turbo 15	Sartorius
Centrifugal Filter, Amicon® Ultra 0.5mL	Merck Millipore
Centrifugal Filter, Amicon® Ultra 4	Merck Millipore
Combitips advanced®	Eppendorf
Disposable cuvette, Rotilabo®, Polystyrene	Carl Roth

Disposable Syringe 5/10mL Discardit™ II	BD
NAP-5 disposable columns, illustra™	GE Healthcare
Ni-NTA Agarose Protino®	Macherey-Nagel
Nitrile Gloves StarGuard® Comfort	STARLAB
Nitrocellulose Blotting Membrane Amersham™ Protran™ premium 0.45μm	GE Healthcare
Protein LoBind Tube 1.5mL	Eppendorf
Reaction tube, SafeSeal 1.5mL/2mL, Polypropylene	Sarstedt
Serological pipette, Costar®Stripette®	Corning
Sf-900™ III Serum Free Medium gibco®	Thermo Fisher Scientific
Single-use hypodermic needles, Sterican®	B.Braun
Strep-Tactin®Superflow® high capacity resin	IBA Lifesciences
TSK SUPRA Hypodermic Needles 2.00×100mm	TSK
Tube 15mL,120×17mm, Polypropylene	Sarstedt
Tube 50mL,114×28mm, Polypropylene	Sarstedt

5.1.4 Enzymes and proteins

Name	Supplier
BSA Fraction V, lyophilized	PAN-Biotech
BSA, ≥ 98%, lyophilized	Sigma-Aldrich
His-tagged NΔ17Arf1	In-house production by Volkmar Fieberg
His-tagged TEV protease	In-house production by Volkmar Fieberg
His-tagged YopH (protein tyrosine phosphatase)	In-house production by Volkmar Fieberg
Human EGF, recombinant from *E.coli*, lyophilized	PeproTech

5.1.5 Antibodies

Name	Dilution	Species	Supplier	Catalog #
Anti-EGFR(D38B1) monoclonal	1:1000	Rabbit	Cell Signaling	4267S
Anti-EGFR(F4) monoclonal	1:1000	Mouse	Santa Cruz	sc-53274
Anti-MBP monoclonal	1:10000	Mouse	NEB	E8032S
Anti-mouse IgG (H+L) DyLight™800	1:15000	Goat	Cell Signaling	5257S
Anti-rabbit IgG (H+L) IRDye®800CW	1:15000	Donkey	LI-COR	926-32213
p-Tyr(PY99) pY monoclonal	1:5000	Mouse	Santa Cruz	sc-7020
Phospho-EGFR(Tyr992) pY992 polyclonal	1:1000	Rabbit	Cell Signaling	2235S
Streptavidin IRDye®800CW(no antibody)	1:10000	—	LI-COR	926-32230

5.1.6 Bacteria strains and cell lines

Strain	Genotype	Provider
E.coli BL21(DE3)	*fhuA2 [lon] ompT gal (λ DE3) [dcm] ΔhsdS λ DE3 = λ sBamHIo ΔEcoRI-B int::(lacI::PlacUV5::T7 gene1) i21 Δnin5*	Merck Millipore
E.coli NEB®10-beta	*Δ(ara-leu) 7697 araD139 fhuA ΔlacX74 galK16 galE15 e14- Φ80dlacZΔM15 recA1 relA1 endA1 nupG rpsL (StrR) rph spoT1 Δ(mrr-hsdRMS-mcrBC)*	NEB
E.coli DH10MultiBac^{Turbo}	*F– mcrA Δ(mrr-hsdRMS-mcrBC) Φ80lacZΔM15 ΔlacX74 recA1 endA1 araD139 Δ(ara, leu)7697 galU galK λ– rpsL nupG /pMON14272 v-cath::Amp^r chiA::LoxP */ pMON7124*	ATG:biosynthetics
Spodoptera frugiperda Sf9	Derived from IPLB-Sf21-AE	Thermo Fisher Scientific

5.1.7 Constructs

Name	Vector	Antibiotic	Expression host	Tag	Cleavage site
EGFRΔC	pACE-Bac1	Gm	Sf9	SBP	TEV
HTint-ICDΔJM27	pACE-Bac1	Gm	Sf9	His	Factor Xa/TEV
MBPT-ICD	pACE-Bac1	Gm	Sf9	MBP	TEV
MBPT-ICD998	pACE-Bac1	Gm	Sf9	MBP	TEV
MBPT-ICD998(4CSA S744C)	pACE-Bac1	Gm	Sf9	MBP	TEV
MBPT-ICD998(I682Q)	pACE-Bac1	Gm	Sf9	MBP	TEV
MBPT-ICD998(V924R)	pACE-Bac1	Gm	Sf9	MBP	TEV
HT-ARNO	pET28	Kan	*E.coli*	His	TEV
HT-ARNO-PH	pET28	Kan	*E.coli*	His	TEV
HT-Sec7	pET28	Kan	*E.coli*	His	TEV
HT-Sec7-4A	pET28	Kan	*E.coli*	His	TEV
HT-ARNOΔPBR	pET28	Kan	*E.coli*	His	TEV
HT-Calmodulin	pET28	Kan	*E.coli*	His	TEV
MBPT-JM	pET28	Kan	*E.coli*	MBP	TEV
MBPT-JMsc	pET28	Kan	*E.coli*	MBP	TEV
HT-ICD	pFastBac1	Amp/Gm	Sf9	His	TEV

5.1.8 Software

Name	Provider
Affinity Designer 1.6.1	Serif (Europe) Ltd.
ChemDraw®17.1.1.0(1)	PerkinElmer Informatics, Inc.
DataGraph 4.3	Visual Data Tools, Inc.
GraphPad Prism®7.0a	GraphPad Software, Inc.
i-control™	Tecan Group, Ltd.
Image Studio™ Lite	LI-COR, Inc.
MO.AffinityAnalysis 2.3.0.7385	NanoTemper Technologies GmbH
PyMOL Molecular Graphics System	Schrödinger, LLC.
SnapGene®4.2.4	GSL Biotech LLC.
UNICORN™6.4	GE Healthcare Europe GmbH

5.2 Methods

5.2.1 Protein expression

All the ARNO-related constructs, calmodulin and JM/JMsc constructs were expressed in prokaryotic *E.coli* system. In contrast, all the EGFR-related constructs except JM/JMsc were expressed in eukaryotic Sf9 system. Notably, before large-scale protein production, small test expression and "quick and dirty" purification (see section 5.2.2.1) were routinely conducted to examine protein expression level.

5.2.1.1 Protein expression in *E.coli* system

Protein expression in *E.coli* is summarized as follows:

Table 5.1 Overlook of protein expression

Protein	Resistance	OD$_{600}$ before IPTG	Expression conditions
HT-ARNO	Kan	0.6-0.8	250µM IPTG, 4h @37℃
HT-ARNO-PH	Kan	0.6-0.8	250µM IPTG, 4h @37℃
HT-Sec7	Kan	0.6-0.8	250µM IPTG, 4h @37℃
HT-Sec7(4A)	Kan	0.6-0.8	250µM IPTG, 4h @37℃
HT-ARNOΔPBR	Kan	0.2-0.3	100µM IPTG, overnight @20℃
HT-Calmodulin	Kan	0.6-0.8	250µM IPTG, 4h @37℃
MBPT-JM	Kan	0.2-0.3	100µM IPTG, overnight @20℃
MBPT-JMsc	Kan	0.2-0.3	100µM IPTG, overnight @20℃

Part of protein expression work in *E.coli* system was assisted by Volkmar Fieberg. To start with, *E.coli* preculture was initiated in 20mL LB medium (50µg/mL Kanamycin) via inoculation of BL21(DE3) glycerol stock of different constructs. After overnight incubation at 37℃ with 110RPM shaking, the preculture was transferred to 2L LB medium with 50µg/mL Kanamycin to start large culture. The cells grew further at 37℃ until reaching target "OD_{600} before IPTG" (see table above), followed by IPTG addition. The large cultures were kept in incubation shaker with 110RPM shaking, under corresponding "expression conditions" (see table above). In the end, cells were harvested in Beckman high-speed centrifuge with JLA-8.1000 Rotor at 4℃, 4000RPM for 20min. Supernatant was discarded and pellets were deposited in -80℃ freezer until purification.

5.2.1.2 Protein expression in Sf9 system

Sf9 cell maintenance and protein expression were done by Yvonne Aschenbach. The Sf9 system used in the lab is *Multi*Bac^{Turbo} modified from Bac-to-Bac® system. Briefly, the recombinant donor plasmid (pFastBac1 or pACE-Bac1) containing cloned gene of interest was delivered via transformation into DH10MultiBac^{Turbo}. After transposition and antibiotic selection, recombinant bacmids were extracted, followed by Sf9 cell transfection to produce V0 recombinant baculovirus. Viral stock was typically amplified for at least one more round to obtain V1 before protein expression.

To start large-scale protein expression, Sf9 cell culture was maintained in Sf-900™ III Serum Free Medium until reaching cell density of $2×10^6$cells/mL. Protein expression was initiated by infecting Sf9 cells with baculovirus stock at volume ratio 1000:1 (cell *vs.* virus). The culture volume was kept at maximum 30% of flask volume during the expression. After 72h incubation at 27℃ with 80RPM shaking, cells were harvested in Beckman high-speed centrifuge with JLA-8.1000 Rotor at 4℃, 1000g for 10min. Supernatant was discarded while cell pellets were deposited in -80℃ freezer until purification.

5.2.2 Protein purification

Before going through large-scale protein purification, a small-scale test purification was routinely performed to check whether target protein was properly expressed. Simplified and fast steps of this procedure, especially washing steps, lead to unclean elution samples visualized on SDS-PAGE gel, hence the name "quick and dirty" purification.

5.2.2.1 "Quick and dirty" test purification

Typically pellets from 30mL culture were taken as test purification material. Cell pellets were resuspended with 1mL lysis buffer (50mM Tris pH8 @4℃, 300mM NaCl, 10% Glycerol) in a 2mL reaction tube, followed by cell lysis on Vibra-Cell™ Ultrasonic Liquid Processor. The sonification program is as follows: 1s working, 4s pause with 10 cycles and 30% power. During sonification, reaction tubes were sitting in the ice/water bath to keep samples cooled.

After the lysis, cell debris was spinned down in a bench-top centrifuge: 20,000g for 5min @4℃. The supernatant was added to 100μL Ni-NTA or amylose resin slurry for incubation of 30min @4℃. Thereafter, the incubation mixture was transferred to a Pierce™ Centrifuge Column connected to the peristaltic pump, followed by one round washing step with lysis buffer. Centrifuge column was disconnected, with bottom end sealed by the cap. Next, 50μL elution buffer (lysis buffer with 500Mm imidazole or 10mM maltose) was added for elution. Mixtures were incubated for 5min @4℃, followed by centrifugation to collect the eluted samples, which were subsequently analyzed via SDS-PAGE to check the protein expression levels. If they are too low, protein expression protocol should be improved. Otherwise, proceed to large-scale protein purification.

5.2.2.2 Large-scale purification of soluble proteins

Purification buffers:

Lysis buffer	**Elution buffer**
50mM Tris pH 7.8 @4℃	50mM Tris pH 7.8 @4℃
300mM NaCl	300mM NaCl
10% glycerol	10% glycerol
25mM imidazole	500mM imidazole or 10mM maltose

Desalting buffer	**SEC buffer**
50mM Tris pH 7.8 @4℃	20mM HEPES pH 7.8 @4℃
300mM NaCl	150mM NaCl
25mM Imidazole	

In general, cells were disrupted via French press in lysis buffer supplemented with protease inhibitor cocktail. Lysate was cleared via high-speed centrifugation at 70,000g for 20min, after which corresponding affinity chromatography purification (Ni-NTA- or amylose-based) was performed. Elution sample went through buffer exchange to switch

the buffer into desalting buffer, before TEV cleavage overnight at 4℃. Notably, during cleavage 14.3mM β-mercaptoethanol and 1mM EDTA were added. For all the His-tagged constructs, protein samples were subjected to reverse Ni-NTA chromatography, and flow-through samples were concentrated to less than 5mL, followed by size exclusion chromatography in the appropriate columns (see table below). Thereafter, another concentration step was taken when necessary.

Overlook of key facts in purification workflow is summarized here (see table below). General protocol has been modified to serve specific purposes. In addition to listed modifications, there have been the following changes. TEV cleavage of HT-ICD was done in the presence of 0.5μM His-tagged YopH and 0.5mM $MgCl_2$ to dephosphorylate the kinase, and TEV cleavage of MBPT-JM and MBPT-JMsc was done at RT for 72h. For calmodulin purification, lysis buffer is 50mM HEPES pH 7.8 @4℃, 300mM NaCl, 10% Glycerol, 25mM imidazole. After French press, cleared lysate was heated at 80℃ for 5min, then cooled down on ice for 10min, before centrifugation to remove denatured protein. Moreover, the protein sample was supplemented with 1mM $CaCl_2$ right before SEC.

Table 5.2 Overlook of protein purification workflow

Protein	Tag	TEV cleavage	Reverse purification	SEC columns
EGFRΔC*	SBP	No	No	None
HTint-ICDΔJM27	His	Yes	Yes	HiLoad™ 16/600 Superdex™ 200 pg
MBPT-ICD	MBP	No	No	HiLoad™ 16/600 Superdex™ 200 pg
MBPT-ICD998	MBP	No	No	HiLoad™ 16/600 Superdex™ 200 pg
MBPT-ICD998(4CSA S744C)	MBP	No	No	HiLoad™ 16/600 Superdex™ 200 pg
MBPT-ICD998(I682Q)	MBP	No	No	HiLoad™ 16/600 Superdex™ 200 pg
MBPT-ICD998(V924R)	MBP	No	No	HiLoad™ 16/600 Superdex™ 200 pg
HT-ICD	His	Yes	Yes	HiLoad™ 16/600 Superdex™ 200 pg
HT-ARNO	His	Yes	Yes	HiLoad™ 16/600 Superdex™ 200 pg
HT-ARNO-PH	His	Yes	Yes	HiLoad™ 16/600 Superdex™ 75 pg
HT-Sec7	His	Yes	Yes	HiLoad™ 16/600 Superdex™ 75 pg
HT-Sec7(4A)	His	Yes	Yes	HiLoad™ 16/600 Superdex™ 75 pg
HT-ARNOΔPBR	His	Yes	Yes	HiLoad™ 16/600 Superdex™ 200 pg
HT-Calmodulin	His	Yes	Yes	HiLoad™ 16/600 Superdex™ 75 pg
MBPT-JM	MBP	Yes	No	HiLoad™ 16/600 Superdex™ 30 pg
MBPT-JMsc	MBP	Yes	No	HiLoad™ 16/600 Superdex™ 30 pg

* EGFRΔC was purified based on a different protocol which will be explained in next section.

5.2.2.3 Large-scale purification of membrane protein

Unlike other constructs, EGFRΔC contains transmembrane segment, which means it cannot be purified with detergent-free buffer. Therefore, a customized membrane protein purification procedure was established.

To begin with, cells were homogenized and lysed in Buffer A (50mM Tris pH8 @4℃, 300mM NaCl, 10% Glycerol, 1mM EDTA and 1% Triton® X-100), supplemented with protease inhibitor cocktail. The volume ratio between pellet and lysis buffer was 1:2, and lysis process took 1h with magnetic stirring in the ice/water bath. Thereafter, lysate went through centrifugation to remove the cell debris. The supernatant was diluted by five-fold with detergent-free buffer (Buffer A without Triton® X-100). After incubation of supernatant with Strep-Tactin® for 1h at 4℃, two rounds of washing were performed with wash buffer (same as Buffer A except that it has 0.1% Triton® X-100). Then resin was transferred to a Pierce™ Centrifuge Column connected to the peristaltic pump, and it was washed once more with D_2O-containing Buffer B (100mm TES pH 7.8 @4℃, 300mM NaCl, 1mM EDTA, 0.1% Triton® X-100 in D_2O) before elution with Buffer B plus 5mM Desthiobiotin. Eluted sample was concentrated with centrifugal filter (100kDa cutoff), until the volume reach 100μL. Requirement of D_2O-containing Buffer for PELDOR is explained in Results section.

Because Triton® X-100 has pronounced absorption at 280nm, concentration of EGFRΔC cannot be determined via NanoDrop. Instead, Triton® X-100-dissolved EGFRΔC was quantified via western blot with anti-EGFR F4 antibody for detection, using another soluble EGFR protein as concentration standard. More details about western blot are in section 5.2.4.2, and protein concentration determination in section 5.2.4.3.

5.2.3 Protein labeling

Purified protein samples were labeled to serve the purpose of EPR measurement or fluorescence-based MST assay.

5.2.3.1 Protein labeling by MTSL, PD168393 or PDsl

Labeling of protein by MTSL or PDsl, which carries a nitroxide spin label, is suited to EPR spectroscopy. PD168393 labeling serves as a control in western blot to check whether spin label moiety impairs binding of the inhibitor towards EGFR kinase domain.

Labeling process was integrated into protein purification procedure. In general, protein concentrations in elution samples from affinity chromatography was measured via NanoDrop, and then labeling was immediately conducted by mixing certain amount of label under corresponding conditions (see table below). Thereafter, labeled samples went through SEC to isolate the target protein, which was rebuffered into final D_2O-containing buffer (100mM TES pH 7.8 @4℃, 150mM NaCl), followed by concentration and quantification via NanoDrop. Details of labeling procedure are summarized as follows:

Table 5.3 Summary of protein labeling (A)

Protein	Label	Protein: label molar ratio	Labeling conditions
MBPT-ICD998(4CSA S744C)	MTSL	1:10	Dark, at 4℃, 2h
EGFRΔC	PDsl	1:1.1	Dark, on ice, 10min
MBPT-ICD	PDsl	1:2	Dark, on ice, 10min
MBPT-ICD998	PD168393/PDsl	1:2	Dark, on ice, 10min
MBPT-ICD998(I682Q)	PDsl	1:2	Dark, on ice, 10min
MBPT-ICD998(V924R)	PDsl	1:2	Dark, on ice, 10min

For EGFRΔC, a difference labeling procedure was followed. After the purification and concentration determination via western blot, protein was labeled by mixing with PDsl in 1:1.1 molar ratio on ice for 10min.

For PELDOR sample preparation, all proteins were mixed in 1:1 volume ratio with deuterated ethylene glycol, before being flash-frozen in Q-band PELDOR tubes.

5.2.3.2 Protein labeling by fluorophores

Fluorescent labeling was performed for MST assay. Theory and practice of MST will be introduced in section 5.2.6. The labeled constructs and labeling parameters are listed below:

Table 5.4 Summary of protein labeling (B)

Protein	Dye	Protein: dye molar ratio	Labeling conditions
Sec7	Alexa Fluor™ 647 NHS Ester	1:10	Dark, on ice, 1h
Sec7(4A)	Alexa Fluor™ 647 NHS Ester	1:10	Dark, on ice, 1h
ICD	RED-NHS 2nd generation	1:3	Dark, on ice, 0.5h
ICDΔJM27	RED-NHS 2nd generation	1:3	Dark, on ice, 0.5h

Labeling reactions here are based on fluorophore NHS ester, which reacts with primary amine in the side chain of lysine. Reaction scheme is as follows:
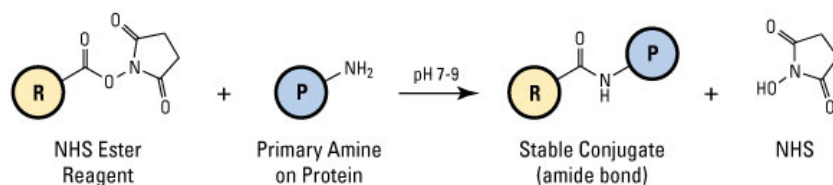
Figure 5.1 Reaction of NHS ester with primary amine on protein. Here R represents fluorophore, whereas P represents protein. Figure source: https://www.thermofisher.com/de/en/home/life-science/protein-biology/protein-biology-learning-center/protein-biology-resource-library/pierce-protein-methods/amine-reactive-crosslinker-chemistry.html

Labeling of Sec7 and Sec7(4A) requires mixing of 10μM protein with 100μM Alexa Fluor™ 647 NHS Ester in labeling buffer A (20mM HEPES pH7.8 @4℃, 150mM NaCl, 100mM NaHCO$_3$). As for the labeling of ICD and ICDΔJM27, 10μM protein was mixed with 30μM RED-NHS 2nd generation in labeling buffer B (20mM HEPES pH7.8 @4℃, 150mM NaCl). Labeling mixtures were kept on ice in darkness for 0.5 or 1h (see table above), before being terminated by 100mM Tris (pH8 @4℃) as quencher. Next, samples were applied to pre-equilibrated (in labeling buffer B) illustra™ NAP-5 column to remove residual free dye, followed by elution with labeling buffer B. Protein concentration and DOL were determined on NanoDrop 2000c Spectrophotometer, after which labeled samples were aliquoted and flash-frozen in liquid nitrogen.

5.2.4 Protein visualization and quantification

Protein samples need to be characterized before downstream experiments. In this regard, SDS-PAGE and western blot are routinely conducted to examine the sample purity, phosphorylation level, *etc.*, sometimes even in a quantitative way. Additionally, NanoDrop provides a fast yet relatively accurate approach to quantify protein concentration and in the case of labeled proteins, degree of labeling as well.

5.2.4.1 SDS-PAGE analysis

Sodium dodecyl sulfate polyacrylamide gel electrophoresis (SDS-PAGE) is one of the fundamental methods to separate proteins. SDS is utilized as anionic detergent to unfold and coat protein uniformly. Unfolded protein carries the negative charges approximately in proportion to its size so different proteins can be separated via electrophoresis based on

sizes. The SDS-PAGE was performed according to Laemmli protocol (Laemmli, 1970). Buffers, solutions and gel casting procedures are shown below. Gel samples were prepared by mixing protein and loading buffer at volume ratio 5:1, followed by boiling on block heater for 10min. Gels were run at 200V for 45-50min, followed by gel staining and destaining. Thoroughly destained gels were scanned by LI-COR Odyssey imaging system.

4 × Separating gel buffer

181.7 g/L Tris pH 8.8 @RT

4 g/L SDS

5 × Running buffer

15.1 g/L Tris

72 g/L glycine

5 g/L SDS

4 × Stacking gel buffer

60.6 g/L Tris pH 6.8 @RT

4 g/L SDS

6 × Loading buffer

7 mL 4x Stacking gel buffer

3 mL glycerol

1.5 g SDS

0.93 g DTT

1.2 mg bromophenol blue

Staining solution

150mL methanol

50mL acetic acid

300mL water

0.5g Coomassie$^{®}$ Brilliant Blue G

Destaining solution

100mL acetic acid

900mL water

Table 5.5 Gel casting setup

	Separating gel [%]							Stacking gel [%]	
	5.0	**6.0**	**7.5**	**8.0**	**10.0**	**12.5**	**15.0**	**3.0**	**4.0**
Acrylamide	834	1000	1250	1333	1667	2083	2500	160	213
Water	2878	2712	2462	2379	2045	1629	1212	1028	975
Separating gel buffer	1250							—	
Stacking gel buffer	—							400	
TEMED	8							2	
APS	30							10.4	

5.2.4.2 Western blot analysis

Western blot analysis was performed based on the semi-dry blotting protocol from Kyhse-Andersen (Kyhse-Andersen, 1984). All the buffers needed are shown below.

Anode I buffer	**Anode II buffer**
36.34 g/L Tris pH 10.4 @RT	3.03 g/L Tris pH 10.4 @RT
Cathode buffer	**10 × TBS**
3.03 g/L Tris pH 9.4 @RT	24.2 g Tris pH 7.6 @RT
3 g/L Glycine	80 g NaCl

After SDS-PAGE, gel was soaked in Cathode buffer, along with 3 pieces of blotting paper (C). Meanwhile, one piece of Nitrocellulose Blotting Membrane and one piece of blotting paper (B) were soaked in Anode II buffer, and another piece of blotting paper (A) was soaked in Anode I buffer. After 10min of soaking, stack the transfer unit in following orders: cathode/paper(C)/gel/membrane/paper(B)/paper(A)/anode. Set the voltage at maximum allowed by blotting device, and current at 2mA/cm^2, for 45min. Thereafter, the transfer unit was disassembled, and the blot went through blocking step by incubating with 5% BSA-containing TBST (TBST: 1×TBS with 0.1% Tween$^®$ 20) on the shaker at RT for 1h. Then the blot was incubated with primary and secondary antibodies, before being scanned by LI-COR Odyssey imaging system. Primary antibody incubation took 2h at RT or overnight at 4℃, whereas secondary antibody incubation took 1h at RT. After each antibody incubation, the blot was washed with TBST 3 times.

5.2.4.3 Protein quantification

Spectrophotometer NanoDrop provides UV/Vis measurement mode that conveniently determines protein concentrations according to the Lambert-Beer law:

$$A = \varepsilon \times b \times c, \tag{6.1}$$

where A represents measured absorbance, ε represents extinction coefficient, b is path length and c is protein concentration. ε is wavelength- and protein sequence-dependent. Details for protein concentration determination are listed below (Table 5.6). Routinely, protein absorption at 280nm was utilized to determine concentration. In several cases (calmodulin, JM and JMsc), there is no tryptophan in protein, which means computed ε at 280nm is not accurate, so another approach based on 214nm absorption of amide bonds was employed. In the case of EGFRΔC, absorption of Triton$^®$ X-100 at 280nm interferes with NanoDrop measurement, so western blot analysis was performed to quantify protein concentration.

In the case of fluorescent-labeled proteins, fluorophores typically have strong absorption at 280nm, so absorbance of protein should be corrected to remove the fluorophore absorption at 280nm. Thus in protein concentration determination using Lambert-Beer

law, A is actually calculated as $A_{280} - A_{650} \times$ correction factor, where A_{650} is the maximum absorbance of fluorophores at 650nm and correction factors are listed below (Table 5.7). Meanwhile, fluorophore concentrations can be computed from Lambert-Beer law as well, using A_{650} and ε_{650}. Degree of labeling for each labeled protein was then calculated as ratio of fluorophore concentration to protein concentration. It indicates number of fluorophore per protein molecule on average.

Table 5.6 Summary of protein concentration determination

Protein	Method	Detection wavelength	Extinction coefficient (M^{-1} cm^{-1}) *
EGFRΔC	Western blot	—	—
HTint-ICDΔJM27**	NanoDrop	280nm	68300
MBPT-ICD	NanoDrop	280nm	136140
MBPT-ICD998	NanoDrop	280nm	120210
MBPT-ICD998(4CSA S744C)	NanoDrop	280nm	120210
MBPT-ICD998(I682Q)	NanoDrop	280nm	120210
MBPT-ICD998(V924R)	NanoDrop	280nm	120210
HT-ICD**	NanoDrop	280nm	68300
HT-ARNO**	NanoDrop	280nm	45380
HT-ARNO-PH**	NanoDrop	280nm	29450
HT-Sec7**	NanoDrop	280nm	12950
HT-Sec7(4A)**	NanoDrop	280nm	11460
HT-ARNOΔPBR**	NanoDrop	280nm	45380
HT-Calmodulin**	NanoDrop	214nm	214835
MBPT-JM**	NanoDrop	214nm	49684
MBPT-JMsc**	NanoDrop	214nm	49684

* Extinction coefficients at 280nm were computed using ExPaSy ProtParam tool (https://web.expasy.org/protparam/). Extinction coefficients at 214nm were calculated as described (Kuipers and Gruppen, 2007).

** For these constructs, extinction coefficients are based on protein sequences after TEV cleavage.

Table 5.7 Spectroscopic parameters of fluorophores

Fluorophore	Correction factor	ε_{650} (M^{-1} cm^{-1})
Alexa Fluor™ 647	0.03	239,000
RED 2nd generation	0.04	195,000

5.2.5 Protein activity assays

5.2.5.1 Autophosphorylation assay

Autophosphorylation assay was performed in non-binding 96-well plate. EGFR constructs were pretreated by 40nM YopH to dephosphorylate phosphotyrosines at RT for 5min,

followed by 1mM Na_3VO_4 treatment to terminate YopH. After RT incubation with Na_3VO_4 for 3min, autophosphorylation was started by incubating proteins in reaction buffer (20mM HEPES pH7.6 @4℃, 50mM NaCl, 5mM $MgCl_2$, 1mM ATP), except that reaction buffer for EGFRΔC contains 0.1% Triton® X-100. At indicated time-points, reaction was terminated by mixing immediately with SDS-PAGE loading buffer (supplemented with 25mM EDTA) and incubating on block heater for 5min. Protein phosphorylation was visualized via western blot immunodetection in which phosphorylated protein was detected by pY or pY992 antibody (anti-phosphotyrosine), while total protein was detected by anti-MBP, anti-EGFR antibody or Streptavidin.

5.2.5.2 Guanine nucleotide exchange assay

ARNO-Sec7 acts as GEF upon GTPase such as Arf1, switching it from GDP-bound to GTP-bound state. Such structural switch of Arf1 results in change of intrinsic tryptophan fluorescence intensity, which is detected by Tecan plate reader.

First, GTP, Arf1 and ARNO mixtures were prepared as follows:

GTP mix			
	Volume	Stock concentration	Working concentration
DPBS + 3mM $MgCl_2$	4975 μL	—	—
GTP	25 μL	100 mM	500 μM

Arf1 mix			
	Volume	Stock concentration	Working concentration
DPBS	676.5 μL	—	—
His-NΔ17Arf1	8.25 μL	136.22 μM	1.5 μM
EDTA pH8 @4℃	3 μL	500 mM	2 mM
GDP	60 μL	1000 μM	80 μM

ARNO mix			
	Volume	Stock concentration	Working concentration
DPBS + 3mM $MgCl_2$	496.25 μL	—	—
ARNO constructs	3.75 μL	8 μM	60 nM

Arf1 mixture with total volume of 747.75μL was incubated at 37℃ with shaking for 15min, to unload intrinsically bound nucleotide. Then 2.25μL of 1M $MgCl_2$ was added to the mixture, followed by incubation for 5min at 37℃, to facilitate binding of external GDP to Arf1.

Thereafter, 20µL ARNO mix was combined with 25µL Arf1 mix and 45µL reaction buffer (DPBS + 3mM MgCl$_2$) in non-binding 96-well plate. Exchange reaction was started by dispensing 10µL GTP mix into each well and tryptophan fluorescence change was monitored by plate reader every 10 second for 15min. Fluorescence curves were plotted using GraphPad Prism®, and exchange activities of various ARNO constructs were compared.

5.2.6 Microscale thermophoresis assay

MST offers a quantitative way of analyzing protein-protein interaction, a method that is less time-consuming, immobilization-free and requires minimal protein consumption. Thermophoresis refers to the phenomenon in which different kinds of particles (here protein molecules) move along temperature gradient at different speeds, depending on charge, size, conformation or hydration shell of the moving particles. A generic MST scheme is illustrated in Figure 5.2.
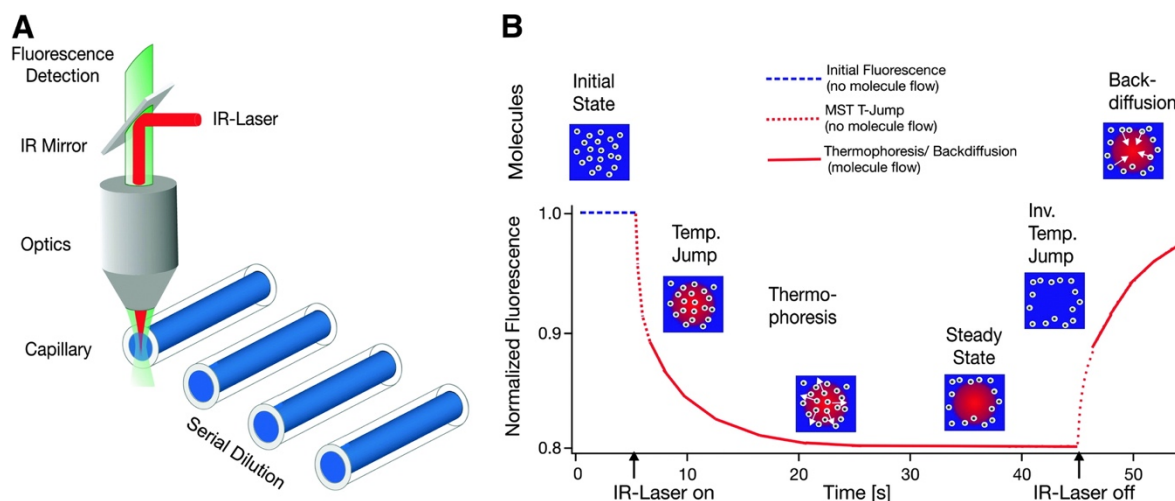
Figure 5.2 MST scheme. (A) Setup of an MST measuring device. In brief, infrared (IR) laser is reflected by IR Mirror and shines upon the middle points of capillaries, creating a temperature gradient. Meanwhile, fluorescence change due to particle thermophoresis is monitored at specific excitation and emission wavelengths depending on the fluorophore. (B) One example of thermophoretic time trace. At 5 second, IR laser is turned on to start thermophoresis, and particles diffuse away from middle point of capillaries. At the end of MST phase, when IR laser is turned off, particles start to diffuse back. Figure is taken from Jerabek-Willemsen et al., 2011.

Principle of MST is well explained in Seidel et al., 2013. When thermophoresis starts, fluorophore-labeled particles begin to move along temperature gradient created by IR laser, leading to decreased concentration in the focus of the IR laser where fluorescence is monitored. In Figure 5.2, relative fluorescence intensity at a chosen time point after "IR-

Laser on" is denoted as F_1, whereas the relative fluorescence intensity right before "IR-Laser on" is denoted as F_0. The time points of F_1 and F_0 are also referred to as hot cursor and cold cursor. Then F_{norm} is defined as:

$$F_{norm} = \frac{F_1}{F_0} \qquad . \qquad (5.1)$$

In a typical MST assay setup for protein-protein interaction, one of the protein is fluorescent-labeled and held at constant concentration, while the other is unlabeled and serially diluted, which is referred to as "ligand" in the MST analysis software. If one of the factors (size, charge, protein conformation or hydration shell) is changed upon binding, F_{norm} is typically changed as well, which means $F_{norm,bound}$ differs from $F_{norm,unbound}$, and they are the signatures of two distinct species, *i.e.* labeled protein-ligand complex and labeled protein alone. In reality, MST sample is a mixture of these two species and F_{norm} is the linear combination of $F_{norm,unbound}$ and $F_{norm,bound}$ as shown below (5.2), where FB means fraction bound of labeled protein.

$$F_{norm} = (1 - FB)F_{norm,unbound} + (FB)F_{norm,bound} \qquad (5.2)$$

In the following mathematical deduction, A represents unlabeled protein while B represents labeled one. In a 1:1 protein interaction equilibrium (5.3), dissociation constant K_d is calculated as in formula (5.4), where [AB] represents concentration of complex AB, while [A] and [B] represent total input concentrations of A and B.

$$AB \rightleftharpoons A + B \qquad (5.3)$$

$$K_d = \frac{[A]_{free}[B]_{free}}{[AB]} = \frac{([A]-[AB])([B]-[AB])}{[AB]} \qquad (5.4)$$

The equation is solvable and solution for fraction bound of B is expressed as follows:

$$FB = \frac{[AB]}{[B]} = \frac{[A]+[B]+K_d-\sqrt{([A]+[B]+K_d)^2-4[A][B]}}{2[B]} \qquad (5.5)$$

On the right-hand side of formula (5.5), only K_d is unknown, while on the left-hand side, FB is linearly related to F_{norm} (5.2) which can be calculated directly from experimental data. Hence K_d can be fitted based on the obtained data.

For each MST assay, unlabeled protein was used to prepare 15-step serial dilution with final volume of 5µL in assay buffer (20mM HEPES pH7.8 @4℃, 150mM NaCl, 0.005% Triton® X-100, 10µM BSA). Next, 5µL of 200nM fluorescent-labeled protein was added to each dilution. For the ICD:CaM interaction, ICDΔJM27:CaM interaction and calmodulin competition assay, 2mM $CaCl_2$ or 2mM EGTA was added to the assay buffer. The calmodulin titration was performed in 1:2 serial dilutions, while Sec7 titration was performed in 1:3 dilutions. For the calmodulin competition assay, 30µM calmodulin was

premixed with 200nM labeled protein, before being added to 15 serial dilutions of ARNO-Sec7. Mixed samples were loaded into Monolith NT.115 premium capillaries and MST measurements were performed on Monolith NT.115 system at 25℃. For assays using labeled Sec7 and Sec7(4A), samples were pre-incubated at room temperature for 10min and measured with 60% excitation power, 50% MST power. For assays using labeled ICD and ICDΔJM27, samples were pre-incubated at room temperature for 5min and measured with 20% excitation power, 40% MST power. Sample preparation and measurement were repeated in triplicate, and data from each protein pair were grouped together for the analysis, which was performed using MO.Affinity Analysis v2.3. For the calculation of F_{norm}, hot cursor was set at 5 seconds for assays including labeled Sec7 and Sec7(4A), while for assays including labeled ICD and ICDΔJM27, hot cursor was set at 2.5 seconds. Data fitting was conducted using "Dose Response Fit" module with Kd Model. Generated fitting result is displayed as F_{norm} *vs*. ligand concentration. For more convenient comparison of different samples, final results are converted to ΔF_{norm} mode by subtracting $F_{norm,unbound}$ from each curve using the "Compare Results" module. Exported data from "Compare Results" module were used for graph preparation in Prism® 7.0a.

5.2.7 EPR

5.2.7.1 CW-EPR

The PDsl-bound EGFR samples were loaded into 10µl glass capillaries (0.6mm inner diameter). The capillaries were sealed with glue to avoid leaking of the sample. Continuous wave EPR spectra were then recorded on an EMXnano X-band EPR Spectrometer from Bruker (Billerica, MA). The samples were measured at room temperature with a microwave power of 2.5mW, a video amplifier gain of 30dB, a modulation amplitude of 1G, a time constant of 81.9ms, a conversion time of 82.05ms, and a resolution of 10 points per G.

5.2.7.2 PELDOR

PELDOR data in this thesis were collected as described (Yin et al., 2017).

5.2.8 Analytical gel filtration

Analytical gel filtration chromatography in this thesis was performed as described (Yin et al., 2017).

6 Appendix

6.1 Protein sequences and parameters

Affinity tags are underlined for all constructs, and cleavage sites are indicated by asterisks for cleaved constructs.

EGFRΔC (SBP-tagged):

MW: 117021.35, pI: 6.39, ε_{280} = 116200 M^{-1}cm^{-1}

LEEKKVCQGTSNKLTQLGTFEDHFLSLQRMFNNCEVVLGNLEITYVQRNYDLSFLKTIQEVAGYVLI
ALNTVERIPLENLQIIRGNMYYENSYALAVLSNYDANKTGLKELPMRNLQEILHGAVRFSNNPALCN
VESIQWRDIVSSDFLSNMSMDFQNHLGSCQKCDPSCPNGSCWGAGEENCQKLTKIICAQQCSGR
CRGKSPSDCCHNQCAAGCTGPRESDCLVCRKFRDEATCKDTCPPLMLYNPTTYQMDVNPEGKY
SFGATCVKKCPRNYVVTDHGSCVRACGADSYEMEEDGVRKCKKCEGPCRKVCNGIGIGEFKDSL
SINATNIKHFKNCTSISGDLHILPVAFRGDSFTHTPPLDPQELDILKTVKEITGFLLIQAWPENRTDLH
AFENLEIIRGRTKQHGQFSLAVVSLNITSLGLRSLKEISDGDVIISGNKNLCYANTINWKKLFGTSGQ
KTKIISNRGENSCKATGQVCHALCSPEGCWGPEPRDCVSCRNVSRGRECVDKCNLLEGEPREFV
ENSECIQCHPECLPQAMNITCTGRGPDNCIQCAHYIDGPHCVKTCPAGVMGENNTLVWKYADAG
HVCHLCHPNCTYGCTGPGLEGCPTNGPKIPSIATGMVGALLLLLVVALGIGLFMRRRHIVRKRTLR
RLLQERELVEPLTPSGEAPNQALLRILKETEFKKIKVLGSGAFGTVYKGLWIPEGEKVKIPVAIKELR
EATSPKANKEILDEAYVMASVDNPHVCRLLGICLTSTVQLITQLMPFGCLLDYVREHKDNIGSQYLL
NWCVQIAKGMNYLEDRRLVHRDLAARNVLVKTPQHVKITDFGLAKLLGAEEKEYHAEGGKVPIKW
MALESILHRIYTHQSDVWSYGVTVWELMTFGSKPYDGIPASEISSILEKGERLPQPPICTIDVYMIMV
KCWMIDADSRPKFRELIIEFSKMARDPQRYLVIQGDERMHLPSPTDSNFYRALMDEEDMDDVVDA
DEYLIPQQGGVPENLYFQG<u>MDEKTTGWRGGHVVEGLAGELEQLRARLEHHPQGQREP</u>SG

MBPT-ICD:

MW: 102870.94, pI: 5.46, ε_{280} = 136140 M^{-1}cm^{-1}

<u>MAMKIEEGKLVIWINGDKGYNGLAEVGKKFEKDTGIKVTVEHPDKLEEKFPQVAATGDGPDIIFWA
HDRFGGYAQSGLLAEITPDKAFQDKLYPFTWDAVRYNGKLIAYPIAVEALSLIYNKDLLPNPPKTWE
EIPALDKELKAKGKSALMFNLQEPYFTWPLIAADGGYAFKYENGKYDIKDVGVDNAGAKAGLTFLV
DLIKNKHMNADTDYSIAEAAFNKGETAMTINGPWAWSNIDTSKVNYGVTVLPTFKGQPSKPFVGVL
SAGINAASPNKELAKEFLENYLLTDEGLEAVNKDKPLGAVALKSYEEELAKDPRIAATMENAQKGEI
MPNIPQMSAFWYAVRTAVINAASGRQTVDEALKDAQT</u>SSGENLYFQGSRRRHIVRKRTLRRLLQE
RELVEPLTPSGEAPNQALLRILKETEFKKIKVLGSGAFGTVYKGLWIPEGEKVKIPVAIKELREATSP
KANKEILDEAYVMASVDNPHVCRLLGICLTSTVQLITQLMPFGCLLDYVREHKDNIGSQYLLNWCV
QIAKGMNYLEDRRLVHRDLAARNVLVKTPQHVKITDFGLAKLLGAEEKEYHAEGGKVPIKWMALES
ILHRIYTHQSDVWSYGVTVWELMTFGSKPYDGIPASEISSILEKGERLPQPPICTIDVYMIMVKCWMI
DADSRPKFRELIIEFSKMARDPQRYLVIQGDERMHLPSPTDSNFYRALMDEEDMDDVVDADEYLIP
QQGFFSSPSTSRTPLLSSLSATSNNSTVACIDRNGLQSCPIKEDSFLQRYSSDPTGALTEDSIDDTF

LPVPEYINQSVPKRPAGSVQNPVYHNQPLNPAPSRDPHYQDPHSTAVGNPEYLNTVQPTCVNST
FDSPAHWAQKGSHQISLDNPDYQQDFFPKEAKPNGIFKGSTAENAEYLRVAPQSSEFIGA

MBPT-ICD998:

MW: 82342.54, pl: 5.63, ε₂₈₀ = 120210 M⁻¹cm⁻¹

MA<u>MKIEEGKLVIWINGDKGYNGLAEVGKKFEKDTGIKVTVEHPDKLEEKFPQVAATGDGPDIIFWA
HDRFGGYAQSGLLAEITPDKAFQDKLYPFTWDAVRYNGKLIAYPIAVEALSLIYNKDLLPNPPKTWE
EIPALDKELKAKGKSALMFNLQEPYFTWPLIAADGGYAFKYENGKYDIKDVGVDNAGAKAGLTFLV
DLIKNKHMNADTDYSIAEAAFNKGETAMTINGPWAWSNIDTSKVNYGVTVLPTFKGQPSKPFVGVL
SAGINAASPNKELAKEFLENYLLTDEGLEAVNKDKPLGAVALKSYEEELAKDPRIAATMENAQKGEI
MPNIPQMSAFWYAVRTAVINAASGRQTVDEALKDAQT</u>SSGENLYFQGSRRRHIVRKRTLRRLLQE
RELVEPLTPSGEAPNQALLRILKETEFKKIKVLGSGAFGTVYKGLWIPEGEKVKIPVAIKELREATSP
KANKEILDEAYVMASVDNPHVCRLLGICLTSTVQLITQLMPFGCLLDYVREHKDNIGSQYLLNWCV
QIAKGMNYLEDRRLVHRDLAARNVLVKTPQHVKITDFGLAKLLGAEEKEYHAEGGKVPIKWMALES
ILHRIYTHQSDVWSYGVTVWELMTFGSKPYDGIPASEISSILEKGERLPQPPICTIDVYMIMVKCWMI
DADSRPKFRELIIEFSKMARDPQRYLVIQGDERMHLPSPTDSNFYRALMDEEDMDDVVDADEYLIP
QQG

MBPT-ICD998(I682Q):

MW: 82357.51, pl: 5.63, ε₂₈₀ = 120210 M⁻¹cm⁻¹

MA<u>MKIEEGKLVIWINGDKGYNGLAEVGKKFEKDTGIKVTVEHPDKLEEKFPQVAATGDGPDIIFWA
HDRFGGYAQSGLLAEITPDKAFQDKLYPFTWDAVRYNGKLIAYPIAVEALSLIYNKDLLPNPPKTWE
EIPALDKELKAKGKSALMFNLQEPYFTWPLIAADGGYAFKYENGKYDIKDVGVDNAGAKAGLTFLV
DLIKNKHMNADTDYSIAEAAFNKGETAMTINGPWAWSNIDTSKVNYGVTVLPTFKGQPSKPFVGVL
SAGINAASPNKELAKEFLENYLLTDEGLEAVNKDKPLGAVALKSYEEELAKDPRIAATMENAQKGEI
MPNIPQMSAFWYAVRTAVINAASGRQTVDEALKDAQT</u>SSGENLYFQGSRRRHIVRKRTLRRLLQE
RELVEPLTPSGEAPNQALLRQLKETEFKKIKVLGSGAFGTVYKGLWIPEGEKVKIPVAIKELREATS
PKANKEILDEAYVMASVDNPHVCRLLGICLTSTVQLITQLMPFGCLLDYVREHKDNIGSQYLLNWC
VQIAKGMNYLEDRRLVHRDLAARNVLVKTPQHVKITDFGLAKLLGAEEKEYHAEGGKVPIKWMALE
SILHRIYTHQSDVWSYGVTVWELMTFGSKPYDGIPASEISSILEKGERLPQPPICTIDVYMIMVKCW
MIDADSRPKFRELIIEFSKMARDPQRYLVIQGDERMHLPSPTDSNFYRALMDEEDMDDVVDADEYL
IPQQG

MBPT-ICD998(V924R):

MW: 82399.60, pl: 5.69, ε₂₈₀ = 120210 M⁻¹cm⁻¹

MA<u>MKIEEGKLVIWINGDKGYNGLAEVGKKFEKDTGIKVTVEHPDKLEEKFPQVAATGDGPDIIFWA
HDRFGGYAQSGLLAEITPDKAFQDKLYPFTWDAVRYNGKLIAYPIAVEALSLIYNKDLLPNPPKTWE
EIPALDKELKAKGKSALMFNLQEPYFTWPLIAADGGYAFKYENGKYDIKDVGVDNAGAKAGLTFLV
DLIKNKHMNADTDYSIAEAAFNKGETAMTINGPWAWSNIDTSKVNYGVTVLPTFKGQPSKPFVGVL
SAGINAASPNKELAKEFLENYLLTDEGLEAVNKDKPLGAVALKSYEEELAKDPRIAATMENAQKGEI</u>

MPNIPQMSAFWYAVRTAVINAASGRQTVDEALKDAQTSSGENLYFQGSRRRHIVRKRTLRRLLQE
RELVEPLTPSGEAPNQALLRILKETEFKKIKVLGSGAFGTVYKGLWIPEGEKVKIPVAIKELREATSP
KANKEILDEAYVMASVDNPHVCRLLGICLTSTVQLITQLMPFGCLLDYVREHKDNIGSQYLLNWCV
QIAKGMNYLEDRRLVHRDLAARNVLVKTPQHVKITDFGLAKLLGAEEKEYHAEGGKVPIKWMALES
ILHRIYTHQSDVWSYGVTVWELMTFGSKPYDGIPASEISSILEKGERLPQPPICTIDVYMIMRKCWMI
DADSRPKFRELIIEFSKMARDPQRYLVIQGDERMHLPSPTDSNFYRALMDEEDMDDVVDADEYLIP
QQG

MBPT-ICD998(4CSA/S744C):

MW: 82262.36, pI: 5.63, ε_{280} = 120210 M^{-1}cm^{-1}

MAMKIEEGKLVIWINGDKGYNGLAEVGKKFEKDTGIKVTVEHPDKLEEKFPQVAATGDGPDIIFWA
HDRFGGYAQSGLLAEITPDKAFQDKLYPFTWDAVRYNGKLIAYPIAVEALSLIYNKDLLPNPPKTWE
EIPALDKELKAKGKSALMFNLQEPYFTWPLIAADGGYAFKYENGKYDIKDVGVDNAGAKAGLTFLV
DLIKNKHMNADTDYSIAEAAFNKGETAMTINGPWAWSNIDTSKVNYGVTVLPTFKGQPSKPFVGVL
SAGINAASPNKELAKEFLENYLLTDEGLEAVNKDKPLGAVALKSYEEELAKDPRIAATMENAQKGEI
MPNIPQMSAFWYAVRTAVINAASGRQTVDEALKDAQTSSGENLYFQGSRRRHIVRKRTLRRLLQE
RELVEPLTPSGEAPNQALLRILKETEFKKIKVLGSGAFGTVYKGLWIPEGEKVKIPVAIKELREATSP
KANKEILDEAYVMACVDNPHVSRLLGIALTSTVQLITQLMPFGSLLDYVREHKDNIGSQYLLNWCV
QIAKGMNYLEDRRLVHRDLAARNVLVKTPQHVKITDFGLAKLLGAEEKEYHAEGGKVPIKWMALES
ILHRIYTHQSDVWSYGVTVWELMTFGSKPYDGIPASEISSILEKGERLPQPPIATIDVYMIMVKCWMI
DADSRPKFRELIIEFSKMARDPQRYLVIQGDERMHLPSPTDSNFYRALMDEEDMDDVVDADEYLIP
QQG

ARNO (His-tagged):

Before TEV cleavage, MW: 48583.26, pI: 5.71

After TEV cleavage, MW: 46690.23, pI: 5.43, ε_{280} = 45380 M^{-1}cm^{-1}

MGSHHHHHHENLYFQ*GSMEDGVYEPPDLTPEERMELENIRRRKQELLVEIQRLREELSEAMSEV
EGLEANEGSKTLQRNRKMAMGRKKFNMDPKKGIQFLVENELLQNTPEEIARFLYKGEGLNKTAIG
DYLGEREELNLAVLHAFVDLHEFTDLNLVQALRQFLWSFRLPGEAQKIDRMMEAFAQRYCLCNPG
VFQSTDTCYVLSFAVIMLNTSLHNPNVRDKPGLERFVAMNRGINEGGDLPEELLRNLYDSIRNEPF
KIPEDDGNDLTHTFFNPDREGWLLKLGGGRVKTWKRRWFILTDNCLYYFEYTTDKEPRGIIPLENL
SIREVDDPRKPNCFELYIPNNKGQLIKACKTEADGRVVEGNHMVYRISAPTQEEKDEWIKSIQAAVS
VDPFYEMLAARKKRISVKKKQEQP

ARNO-PH (His-tagged):

Before TEV cleavage, MW: 15812.86, pI: 6.65

After TEV cleavage, MW: 13919.83, pI: 6.80, ε_{280} = 29450 M^{-1}cm^{-1}

MGSHHHHHHENLYFQ*GSDREGWLLKLGGGRVKTWKRRWFILTDNCLYYFEYTTDKEPRGIIPLE
NLSIREVDDPRKPNCFELYIPNNKGQLIKACKTEADGRVVEGNHMVYRISAPTQEEKDEWIKSIQAA
VSVD

Sec7 (His-tagged):

Before TEV cleavage, MW: 23536.92, pI: 6.50

After TEV cleavage, MW: 21643.88, pI: 6.38, ε_{280} = 12950 M^{-1}cm^{-1}

MGS<u>HHHHHH</u>ENLYFQ*GSRNRKMAMGRKKFNMDPKKGIQFLVENELLQNTPEEIARFLYKGEGL
NKTAIGDYLGEREELNLAVLHAFVDLHEFTDLNLVQALRQFLWSFRLPGEAQKIDRMMEAFAQRY
CLCNPGVFQSTDTCYVLSFAVIMLNTSLHNPNVRDKPGLERFVAMNRGINEGGDLPEELLRNLYD
SIRNEPFKIP

Sec7(4A) (His-tagged):

Before TEV cleavage, MW: 23266.53, pI: 6.50

After TEV cleavage, MW: 21373.50, pI: 6.38, ε_{280} = 11460 M^{-1}cm^{-1}

MGS<u>HHHHHH</u>ENLYFQ*GSRNRKMAMGRKKFNMDPKKGIQFLVENELLQNTPEEIARFLYKGEGL
NKTAIGDYLGEREELNLAVLHAFVDLHEFTDLNLVQALRQFLWSFRLPGEAQKIDRMMEAFAQRY
CLCNPGVFQSTDTCAVLSAAVAALNTSLHNPNVRDKPGLERFVAMNRGINEGGDLPEELLRNLYD
SIRNEPFKIP

ARNOΔPBR (His-tagged):

Before TEV cleavage, MW: 46848.15, pI: 5.29

After TEV cleavage, MW: 44955.12, pI: 5.07, ε_{280} = 45380 M^{-1}cm^{-1}

MGS<u>HHHHHH</u>ENLYFQ*GSMEDGVYEPPDLTPEERMELENIRRRKQELLVEIQRLREELSEAMSEV
EGLEANEGSKTLQRNRKMAMGRKKFNMDPKKGIQFLVENELLQNTPEEIARFLYKGEGLNKTAIG
DYLGEREELNLAVLHAFVDLHEFTDLNLVQALRQFLWSFRLPGEAQKIDRMMEAFAQRYCLCNPG
VFQSTDTCYVLSFAVIMLNTSLHNPNVRDKPGLERFVAMNRGINEGGDLPEELLRNLYDSIRNEPF
KIPEDDGNDLTHTFFNPDREGWLLKLGGGRVKTWKRRWFILTDNCLYYFEYTTDKEPRGIIPLENL
SIREVDDPRKPNCFELYIPNNKGQLIKACKTEADGRVVEGNHMVYRISAPTQEEKDEWIKSIQAAVS
VDPFYEMLAA

ICD (His-tagged):

Before TEV cleavage, MW: 63140.71, pI: 6.01

After TEV cleavage, MW: 61304.73, pI: 5.85, ε_{280} = 68300 M^{-1}cm^{-1}

MS<u>HHHHHH</u>ENLYFQ*GARRRHIVRKRTLRRLLQERELVEPLTPSGEAPNQALLRILKETEFKKIKVL
GSGAFGTVYKGLWIPEGEKVKIPVAIKELREATSPKANKEILDEAYVMASVDNPHVCRLLGICLTST
VQLITQLMPFGCLLDYVREHKDNIGSQYLLNWCVQIAKGMNYLEDRRLVHRDLAARNVLVKTPQH
VKITDFGLAKLLGAEEKEYHAEGGKVPIKWMALESILHRIYTHQSDVWSYGVTVWELMTFGSKPYD
GIPASEISSILEKGERLPQPPICTIDVYMIMVKCWMIDADSRPKFRELIIEFSKMARDPQRYLVIQGDE
RMHLPSPTDSNFYRALMDEEDMDDVVDADEYLIPQQGFFSSPSTSRTPLLSSLSATSNNSTVACID
RNGLQSCPIKEDSFLQRYSSDPTGALTEDSIDDTFLPVPEYINQSVPKRPAGSVQNPVYHNQPLNP

APSRDPHYQDPHSTAVGNPEYLNTVQPTCVNSTFDSPAHWAQKGSHQISLDNPDYQQDFFPKEA
KPNGIFKGSTAENAEYLRVAPQSSEFIGA

HTint-ICDΔJM27:

Before TEV cleavage, MW: 63729.39, pI: 6.01

After TEV cleavage, MW: 57785.56, pI: 5.34, ε_{280} = 68300 M^{-1}cm^{-1}

MASS<u>HHHHHH</u>IEGRCRRRHIVRKRTLRRLLQERELVEPLTPSENLYFQ*GEAPNQALLRILKETEFK
KIKVLGSGAFGTVYKGLWIPEGEKVKIPVAIKELREATSPKANKEILDEAYVMASVDNPHVCRLLGIC
LTSTVQLITQLMPFGCLLDYVREHKDNIGSQYLLNWCVQIAKGMNYLEDRRLVHRDLAARNVLVKT
PQHVKITDFGLAKLLGAEEKEYHAEGGKVPIKWMALESILHRIYTHQSDVWSYGVTVWELMTFGSK
PYDGIPASEISSILEKGERLPQPPICTIDVYMIMVKCWMIDADSRPKFRELIIEFSKMARDPQRYLVIQ
GDERMHLPSPTDSNFYRALMDEEDMDDVVDADEYLIPQQGFFSSPSTSRTPLLSSLSATSNNSTV
ACIDRNGLQSCPIKEDSFLQRYSSDPTGALTEDSIDDTFLPVPEYINQSVPKRPAGSVQNPVYHNQ
PLNPAPSRDPHYQDPHSTAVGNPEYLNTVQPTCVNSTFDSPAHWAQKGSHQISLDNPDYQQDFF
PKEAKPNGIFKGSTAENAEYLRVAPQSSEFIGA

MBPT-JM:

Before TEV cleavage, MW: 46122.61, pI: 5.67

After TEV cleavage, MW: 4572.39, pI: 11.91, ε_{214} = 49684 M^{-1}cm^{-1}

<u>MAMKIEEGKLVIWINGDKGYNGLAEVGKKFEKDTGIKVTVEHPDKLEEKFPQVAATGDGPDIIFWA
HDRFGGYAQSGLLAEITPDKAFQDKLYPFTWDAVRYNGKLIAYPIAVEALSLIYNKDLLPNPPKTWE
EIPALDKELKAKGKSALMFNLQEPYFTWPLIAADGGYAFKYENGKYDIKDVGVDNAGAKAGLTFLV
DLIKNKHMNADTDYSIAEAAFNKGETAMTINGPWAWSNIDTSKVNYGVTVLPTFKGQPSKPFVGVL
SAGINAASPNKELAKEFLENYLLTDEGLEAVNKDKPLGAVALKSYEEELAKDPRIAATMENAQKGEI
MPNIPQMSAFWYAVRTAVINAASGRQTVDEALKDAQT</u>SSGENLYFQ*RRRHIVRKRTLRRLLQER
ELVEPLTPSGEAPNQALLRI

MBPT-JMsc:

Before TEV cleavage, MW: 46122.61, pI: 5.67

After TEV cleavage, MW: 4572.39, pI: 11.91, ε_{214} = 49684 M^{-1}cm^{-1}

<u>MAMKIEEGKLVIWINGDKGYNGLAEVGKKFEKDTGIKVTVEHPDKLEEKFPQVAATGDGPDIIFWA
HDRFGGYAQSGLLAEITPDKAFQDKLYPFTWDAVRYNGKLIAYPIAVEALSLIYNKDLLPNPPKTWE
EIPALDKELKAKGKSALMFNLQEPYFTWPLIAADGGYAFKYENGKYDIKDVGVDNAGAKAGLTFLV
DLIKNKHMNADTDYSIAEAAFNKGETAMTINGPWAWSNIDTSKVNYGVTVLPTFKGQPSKPFVGVL
SAGINAASPNKELAKEFLENYLLTDEGLEAVNKDKPLGAVALKSYEEELAKDPRIAATMENAQKGEI
MPNIPQMSAFWYAVRTAVINAASGRQTVDEALKDAQT</u>SSGENLYFQ*RELKHIQVRLRTERQLEPL
EIRAVNRSRLTPRLAGLPR

Calmodulin (His-tagged):

Before TEV cleavage, MW: 18599.43, pI: 4.35

After TEV cleavage, MW: 16706.39, pI: 4.09, ε_{214} = 214835 M^{-1}cm^{-1}

MGS<u>HHHHHH</u>ENLYFQ*ADQLTEEQIAEFKEAFSLFDKDGDGTITTKELGTVMRSLGQNPTEAELQ
DMINEVDADGNGTIDFPEFLTMMARKMKDTDSEEEIREAFRVFDKDGNGYISAAELRHVMTNLGE
KLTDEEVDEMIREADIDGDGQVNYEEFVQMMTAK

6.2 Uncorrected PELDOR time traces from EPR study
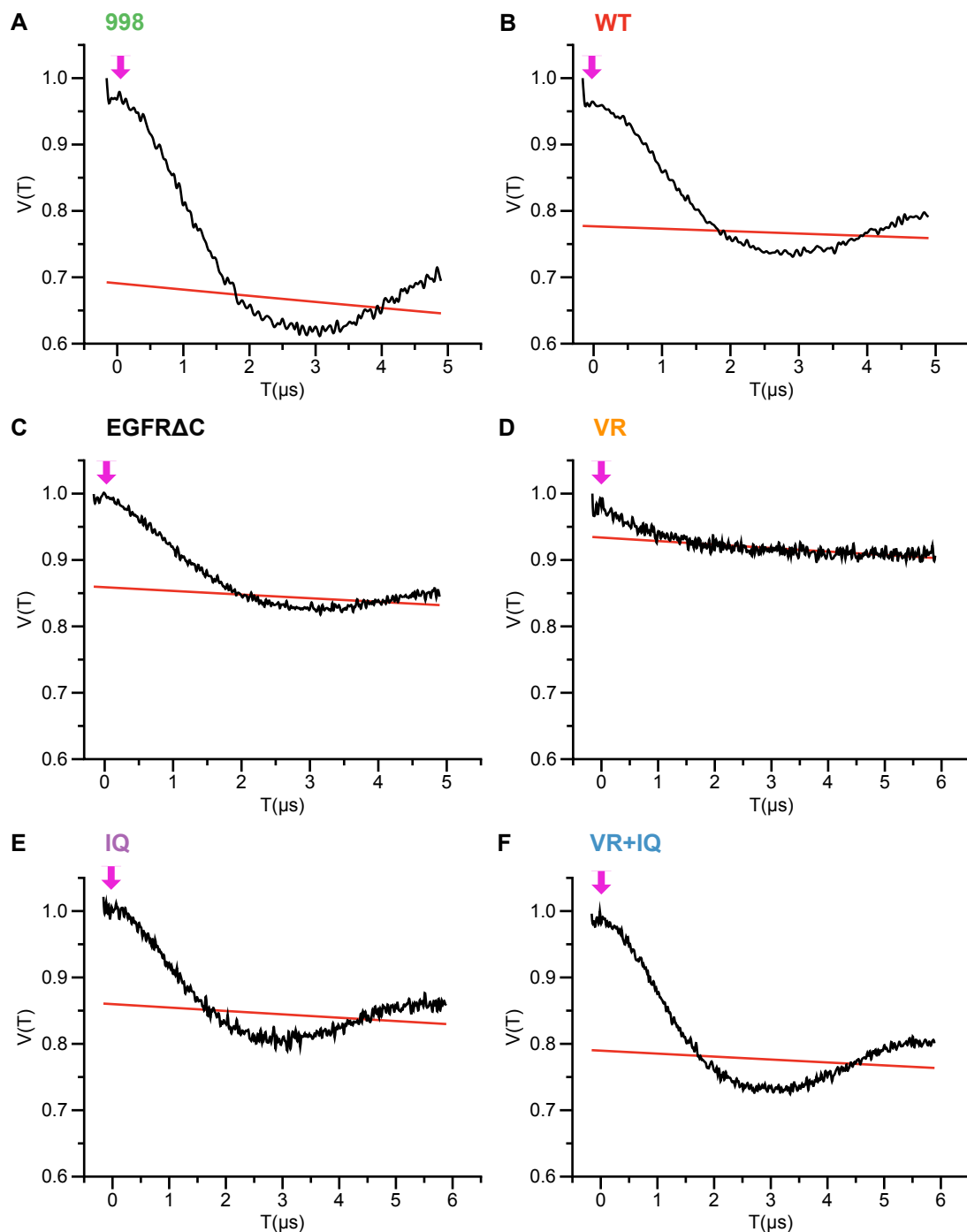
Figure S1 Uncorrected PELDOR time traces of all EGFR constructs used in conformation study. In each graph, red curve denotes the background function, while magenta arrow marks the zero measurement time point.

6.3 Individual data points from MST assay
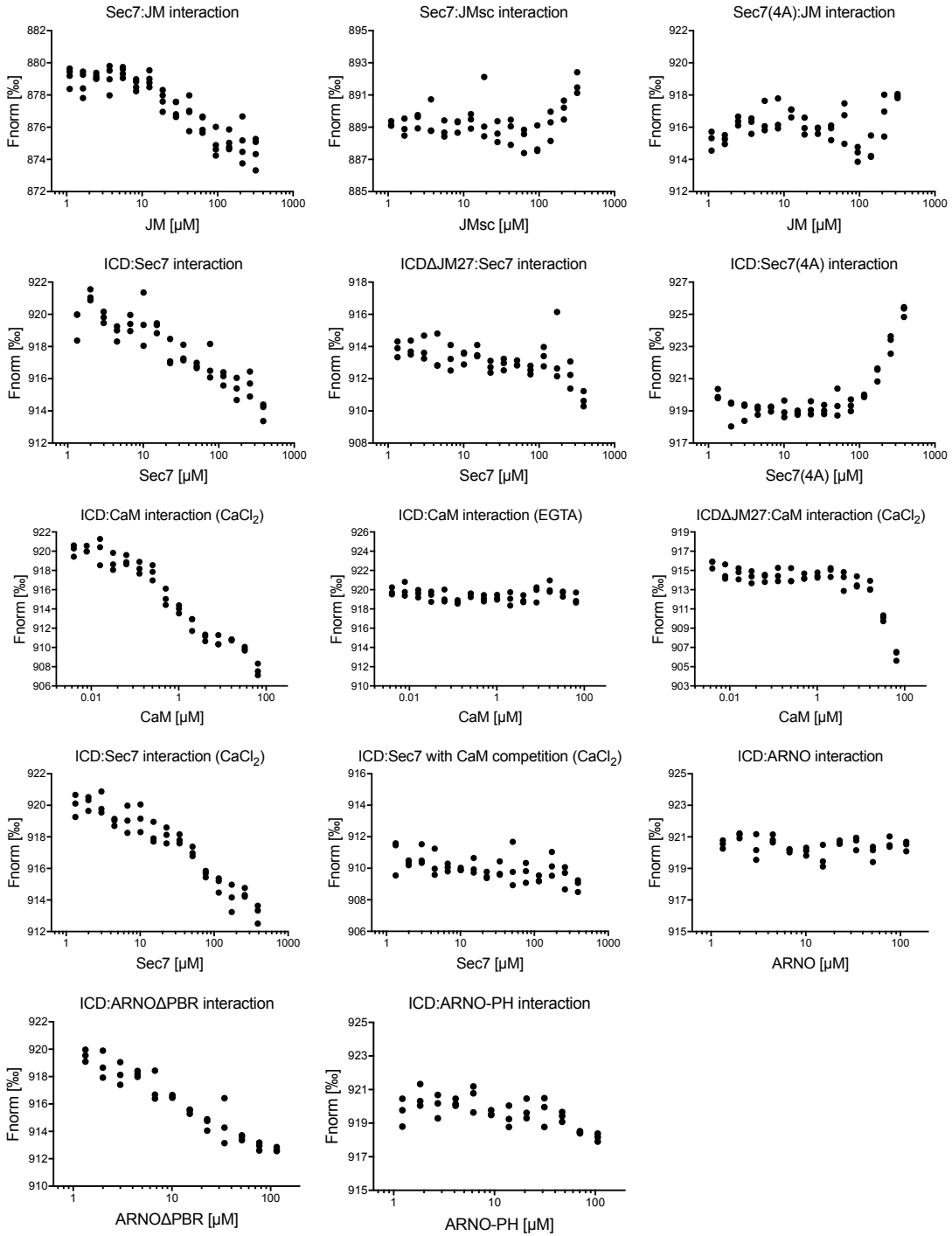
Figure S2 Individual MST measurements are grouped as protein interaction pairs. MST of each pair was repeated in triplicate and plotted as scatter graph. CaM: calmodulin.

References

Abdullin, D., Florin, N., Hagelueken, G., and Schiemann, O. (2015). EPR-based approach for the localization of paramagnetic metal ions in biomolecules. Angew Chem Int Ed Engl *54*, 1827–1831.

Aifa, S., Johansen, K., Nilsson, U.K., Liedberg, B., Lundström, I., and Svensson, S.P.S. (2002). Interactions between the juxtamembrane domain of the EGFR and calmodulin measured by surface plasmon resonance. Cell. Signal.

Alwan, H.A.J., Van Zoelen, E.J.J., and Van Leeuwen, J.E.M. (2003). Ligand-induced lysosomal epidermal growth factor receptor (EGFR) degradation is preceded by proteasome-dependent EGFR de-ubiquitination. J. Biol. Chem.

Arkhipov, A., Shan, Y., Das, R., Endres, N.F., Eastwood, M.P., Wemmer, D.E., Kuriyan, J., and Shaw, D.E. (2013). Architecture and membrane interactions of the EGF receptor. Cell *152*, 557–569.

Arteaga, C.L., Ramsey, T.T., Shawver, L.K., and Guyer, C.A. (1997). Unliganded epidermal growth factor receptor dimerization induced by direct interaction of quinazolines with the ATP binding site. J. Biol. Chem.

Azarkh, M., Okle, O., Singh, V., Seemann, I.T., Hartig, J.S., Dietrich, D.R., and Drescher, M. (2011a). Long-range distance determination in a DNA model system inside Xenopus laevis oocytes by in-cell spin-label EPR. Chembiochem *12*, 1992–1995.

Azarkh, M., Okle, O., Eyring, P., Dietrich, D.R., and Drescher, M. (2011b). Evaluation of spin labels for in-cell EPR by analysis of nitroxide reduction in cell extract of Xenopus laevis oocytes. J Magn Reson *212*, 450–454.

Bakker, J., Spits, M., Neefjes, J., and Berlin, I. (2017). The EGFR odyssey – from activation to destruction in space and time. J. Cell Sci.

Barta, P., Bjorkelund, H., and Andersson, K. (2011). Circumventing the requirement of binding saturation for receptor quantification using interaction kinetic extrapolation. Nucl. Med. Commun. *32*, 863–867.

BENGURÍA, A., MARTÍN-NIETO, J., BENAIM, G., and VILLALOBO, A. (1995). Regulatory Interaction between Calmodulin and the Epidermal Growth Factor Receptor. Ann. N. Y. Acad. Sci. *766*, 472–476.

Béraud-Dufour, S., Robineau, S., Chardin, P., Paris, S., Chabre, M., Cherfils, J., and Antonny, B. (1998). A glutamic finger in the guanine nucleotide exchange factor ARNO displaces Mg2+and the β-phosphate to destabilize GDP on ARF1. EMBO J.

Béraud-Dufour, S., Paris, S., Chabre, M., and Antonny, B. (1999). Dual interaction of ADP ribosylation factor 1 with Sec7 domain and with lipid membranes during catalysis of guanine nucleotide exchange. J. Biol. Chem.

Bill, A., Schmitz, A., König, K., Heukamp, L.C., Hannam, J.S., and Famulok, M. (2012). Anti-proliferative effect of cytohesin inhibition in gefitinib-resistant lung cancer cells. PLoS One.

Blair, J.A., Rauh, D., Kung, C., Yun, C.H., Fan, Q.W., Rode, H., Zhang, C., Eck, M.J., Weiss, W.A., and Shokat, K.M. (2007). Structure-guided development of affinity probes for tyrosine kinases

References

using chemical genetics. Nat. Chem. Biol.

Blechman, J.M., Lev, S., Barg, J., Eisenstein, M., Vaks, B., Vogel, Z., Givol, D., and Yarden, Y. (1995). The fourth immunoglobulin domain of the stem cell factor receptor couples ligand binding to signal transduction. Cell.

Boehm, T., Hofer, S., Winklehner, P., Kellersch, B., Geiger, C., Trockenbacher, A., Neyer, S., Fiegl, H., Ebner, S., Ivarsson, L., et al. (2003). Attenuation of cell adhesion in lymphocytes is regulated by CYTIP, a protein which mediates signal complex sequestration. EMBO J.

De Bondt, H.L., Rosenblatt, J., Jancarik, J., Jones, H.D., Morgant, D.O., and Kim, S.H. (1993). Crystal structure of cyclin-dependent kinase 2. Nature.

Bose, R., and Zhang, X. (2009). The ErbB kinase domain: Structural perspectives into kinase activation and inhibition. Exp. Cell Res.

Boulay, P.L., Cotton, M., Melançon, P., and Claing, A. (2008). ADP-ribosylation factor 1 controls the activation of the phosphatidylinositol 3-kinase pathway to regulate epidermal growth factor-dependent growth and migration of breast cancer cells. J. Biol. Chem.

Bublil, E.M., Pines, G., Patel, G., Fruhwirth, G., Ng, T., and Yarden, Y. (2010). Kinase-mediated quasi-dimers of EGFR. FASEB J.

Cai, G., Zhu, L., Chen, X., Sun, K., Liu, C., Sen, G.C., Stark, G.R., and Qin, J. (2018). TRAF4 binds to the juxtamembrane region of EGFR directly and promotes kinase activation [Cell Biology]. Proc. Natl. Acad. Sci. U. S. A.

Casanova, J.E. (2007). Regulation of Arf activation: The Sec7 family of guanine nucleotide exchange factors. Traffic.

Chen, H., Tucker, J., Wang, X., Gavine, P.R., Phillips, C., Augustin, M.A., Schreiner, P., Steinbacher, S., Preston, M., and Ogg, D. (2016). Discovery of a novel allosteric inhibitor-binding site in ERK5: Comparison with the canonical kinase hinge ATP-binding site. Acta Crystallogr. Sect. D Struct. Biol.

Cherepanov, A. V., and De Vries, S. (2004). Microsecond freeze-hyperquenching: Development of a new ultrafast micro-mixing and sampling technology and application to enzyme catalysis. Biochim. Biophys. Acta - Bioenerg.

Cherfils, J., Ménétrey, J., Mathieu, M., Le Bras, G., Robineau, S., Béraud-Dufour, S., Antonny, B., and Chardin, P. (1998). Structure of the Sec 7 domain of the Arf exchange factor ARNO. Nature.

Cherfils, J., and Melancon, P. (2005). On the action of Brefeldin A on Sec7-stimulated membrane-recruitment and GDP/GTP exchange of Arf proteins. Biochem. Soc. Trans.

Clapham, D.E. (2007). Calcium Signaling. Cell.

Cohen, L.A., Honda, A., Varnai, P., Brown, F.D., Balla, T., and Donaldson, J.G. (2007). Active Arf6 Recruits ARNO/Cytohesin GEFs to the PM by Binding Their PH Domains. Mol. Biol. Cell.

Cox, R., Mason-Gamer, R.J., Jackson, C.L., and Segev, N. (2004). Phylogenetic analysis of Sec7-Domain-containing arf nucleotide exchangers. Mol. Biol. Cell.

Cross, D.A.E., Ashton, S.E., Ghiorghiu, S., Eberlein, C., Nebhan, C.A., Spitzler, P.J., Orme, J.P., Finlay, M.R. V., Ward, R.A., Mellor, M.J., et al. (2014). AZD9291, an irreversible EGFR TKI, overcomes T790M-mediated resistance to EGFR inhibitors in lung cancer. Cancer Discov.

References

D'Souza-Schorey, C., and Chavrier, P. (2006). ARF proteins: Roles in membrane traffic and beyond. Nat. Rev. Mol. Cell Biol.

Dimura, M., Peulen, T.O., Hanke, C.A., Prakash, A., Gohlke, H., and Seidel, C.A. (2016). Quantitative FRET studies and integrative modeling unravel the structure and dynamics of biomolecular systems. Curr Opin Struct Biol *40*, 163–185.

DiNitto, J.P., Delprato, A., Gabe Lee, M.T., Cronin, T.C., Huang, S., Guilherme, A., Czech, M.P., and Lambright, D.G. (2007). Structural Basis and Mechanism of Autoregulation in 3-Phosphoinositide-Dependent Grp1 Family Arf GTPase Exchange Factors. Mol. Cell.

Eberth, A., and Ahmadian, M.R. (2009). In vitro GEF and GAP assays. Curr. Protoc. Cell Biol.

Eden, E.R., White, I.J., Tsapara, A., and Futter, C.E. (2010). Membrane contacts between endosomes and ER provide sites for PTP1B-epidermal growth factor receptor interaction. Nat. Cell Biol.

Endres, N.F., Das, R., Smith, A.W., Arkhipov, A., Kovacs, E., Huang, Y., Pelton, J.G., Shan, Y., Shaw, D.E., Wemmer, D.E., et al. (2013). Conformational coupling across the plasma membrane in activation of the EGF receptor. Cell *152*, 543–556.

Endres, N.F., Barros, T., Cantor, A.J., and Kuriyan, J. (2014). Emerging concepts in the regulation of the EGF receptor and other receptor tyrosine kinases. Trends Biochem. Sci.

Engel, J., Richters, A., Getlik, M., Tomassi, S., Keul, M., Termathe, M., Lategahn, J., Becker, C., Mayer-Wrangowski, S., Grütter, C., et al. (2015). Targeting Drug Resistance in EGFR with Covalent Inhibitors: A Structure-Based Design Approach. J. Med. Chem.

Fabbro, D., Cowan-Jacob, S.W., and Moebitz, H. (2015). Ten things you should know about protein kinases: IUPHAR Review 14. Br J Pharmacol *172*, 2675–2700.

Ferguson, K.M., Berger, M.B., Mendrola, J.M., Cho, H.S., Leahy, D.J., and Lemmon, M.A. (2003). EGF activates its receptor by removing interactions that autoinhibit ectodomain dimerization. Mol. Cell.

Frederick, L., Wang, X.-Y.Y., Eley, G., and James, C.D. (2000). Diversity and frequency of epidermal growth factor receptor mutations in human glioblastomas. Cancer Res.

Frueh, D.P., Goodrich, A.C., Mishra, S.H., and Nichols, S.R. (2013). NMR methods for structural studies of large monomeric and multimeric proteins. Curr Opin Struct Biol *23*, 734–739.

Glaenzer, J., Peter, M.F., Thomas, G.H., and Hagelueken, G. (2017). PELDOR Spectroscopy Reveals Two Defined States of a Sialic Acid TRAP Transporter SBP in Solution. Biophys J *112*, 109–120.

Goldberg, J. (1998). Structural basis for activation of ARF GTPase: Mechanisms of guanine nucleotide exchange and GTP-myristoyl switching. Cell.

Gotoh, N., Tojo, A., Hino, M., Yazaki, Y., and Shibuya, M. (1992). 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.

Gsandtner, I., Charalambous, C., Stefan, E., Ogris, E., Freissmuth, M., and Zezula, J. (2005). Heterotrimeric G protein-independent signaling of a G protein-coupled receptor: Direct binding of ARNO/cytohesin-2 to the carboxyl terminus of the A2A adenosine receptor is necessary for

References

sustained activation of the ERK/MAP kinase pathway. J. Biol. Chem.

Hackel, P.O., Gishizky, M., and Ullrich, A. (2001). Mig-6 is a negative regulator of the epidermal growth factor receptor signal. Biol. Chem.

Hafner, M., Schmitz, A., Grüne, I., Srivatsan, S.G., Paul, B., Kolanus, W., Quast, T., Kremmer, E., Bauer, I., and Famulok, M. (2006). Inhibition of cytohesins by SecinH3 leads to hepatic insulin resistance. Nature.

Hagelueken, G., Ward, R., Naismith, J.H., and Schiemann, O. (2012). MtsslWizard: In Silico Spin-Labeling and Generation of Distance Distributions in PyMOL. Appl. Magn. Reson.

Hellenkamp, B., Schmid, S., Doroshenko, O., Opanasyuk, O., Kuhnemuth, R., Rezaei Adariani, S., Ambrose, B., Aznauryan, M., Barth, A., Birkedal, V., et al. (2018). Precision and accuracy of single-molecule FRET measurements-a multi-laboratory benchmark study. Nat Methods *15*, 669–676.

Henderson, R. (1995). The potential and limitations of neutrons, electrons and X-rays for atomic resolution microscopy of unstained biological molecules. Q Rev Biophys *28*, 171–193.

Honegger, a, Dull, T.J., Szapary, D., Komoriya, a, Kris, R., Ullrich, a, and Schlessinger, J. (1988). Kinetic parameters of the protein tyrosine kinase activity of EGF-receptor mutants with individually altered autophosphorylation sites. EMBO J.

Hubbard, S.R., and Miller, W.T. (2007). Receptor tyrosine kinases: mechanisms of activation and signaling. Curr. Opin. Cell Biol.

Hughes, A.R., Bird, G.S., Obie, J.F., Thastrup, O., and Putney Jr., J.W. (1991). Role of inositol (1,4,5)trisphosphate in epidermal growth factor- induced Ca2+ signaling in A431 cells. Mol.Pharmacol.

Hurtado-Lorenzo, A., Skinner, M., El Annan, J., Futai, M., Sun-Wada, G.H., Bourgoin, S., Casanova, J., Wildeman, A., Bechoua, S., Ausiello, D.A., et al. (2006). V-ATPase interacts with ARNO and Arf6 in early endosomes and regulates the protein degradative pathway. Nat. Cell Biol.

Ibach, J., Radon, Y., Gelléri, M., Sonntag, M.H., Brunsveld, L., Bastiaens, P.I.H., and Verveer, P.J. (2015). Single particle tracking reveals that EGFR signaling activity is amplified in clathrin-coated pits. PLoS One.

Igarashi, R., Sakai, T., Hara, H., Tenno, T., Tanaka, T., Tochio, H., and Shirakawa, M. (2010). Distance Determination in Proteins inside Xenopus laevis Oocytes by Double Electron-Electron Resonance Experiments. J. Am. Chem. Soc. *132*, 8228–+.

Jackson, C.L., and Casanova, J.E. (2000). Turning on ARF: The Sec7 family of guanine-nucleotide-exchange factors. Trends Cell Biol.

Jagtap, A.P., Krstic, I., Kunjir, N.C., Hansel, R., Prisner, T.F., and Sigurdsson, S.T. (2015). Sterically shielded spin labels for in-cell EPR spectroscopy: analysis of stability in reducing environment. Free Radic Res *49*, 78–85.

Jassoy, J.J., Berndhauser, A., Duthie, F., Kuhn, S.P., Hagelueken, G., and Schiemann, O. (2017). Versatile Trityl Spin Labels for Nanometer Distance Measurements on Biomolecules In Vitro and within Cells. Angew. Chemie-International Ed. *56*, 177–181.

Jeffrey, P.D., Russo, A.A., Polyak, K., Gibbs, E., Hurwitz, J., Massagué, J., and Pavletich, N.P. (1995). Mechanism of CDK activation revealed by the structure of a cyclinA-CDK2 complex. Nature.

References

Jerabek-Willemsen, M., Wienken, C.J., Braun, D., Baaske, P., and Duhr, S. (2011). Molecular Interaction Studies Using Microscale Thermophoresis. Assay Drug Dev. Technol.

Jeschke, G., and Polyhach, Y. (2007). Distance measurements on spin-labelled biomacromolecules by pulsed electron paramagnetic resonance. Phys. Chem. Chem. Phys.

Jeschke, G., Chechik, V., Ionita, P., Godt, A., Zimmermann, H., Banham, J., Timmel, C.R., Hilger, D., and Jung, H. (2006). DeerAnalysis2006 - A comprehensive software package for analyzing pulsed ELDOR data. Appl. Magn. Reson.

Jiang, X. (2003). Grb2 Regulates Internalization of EGF Receptors through Clathrin-coated Pits. Mol. Biol. Cell.

Jura, N., Endres, N.F., Engel, K., Deindl, S., Das, R., Lamers, M.H., Wemmer, D.E., Zhang, X., and Kuriyan, J. (2009). Mechanism for activation of the EGF receptor catalytic domain by the juxtamembrane segment. Cell *137*, 1293–1307.

Jura, N., Zhang, X., Endres, N.F., Seeliger, M.A., Schindler, T., and Kuriyan, J. (2011). Catalytic Control in the EGF Receptor and Its Connection to General Kinase Regulatory Mechanisms. Mol. Cell.

Klaeger, S., Heinzlmeir, S., Wilhelm, M., Polzer, H., Vick, B., Koenig, P.A., Reinecke, M., Ruprecht, B., Petzoldt, S., Meng, C., et al. (2017). The target landscape of clinical kinase drugs. Science (80-.). *358*.

Klarlund, J.K., Tsiaras, W., Holik, J.J., Chawla, A., and Czech, M.P. (2000). Distinct polyphosphoinositide binding selectivities for pleckstrin homology domains of GRP1-like proteins based on diglycine versus triglycine motifs. J. Biol. Chem.

Kolanus, W. (2007). Guanine nucleotide exchange factors of the cytohesin family and their roles in signal transduction. Immunol. Rev.

Kolanus, W., Nagel, W., Schiller, B., Zeitlmann, L., Godar, S., Stockinger, H., and Seed, B. (1996). αLβ2 integrin/LFA-1 binding to ICAM-1 induced by cytohesin-1, a cytoplasmic regulatory molecule. Cell.

Krstic, I., Hansel, R., Romainczyk, O., Engels, J.W., Dotsch, V., and Prisner, T.F. (2011). Long-range distance measurements on nucleic acids in cells by pulsed EPR spectroscopy. Angew Chem Int Ed Engl *50*, 5070–5074.

Kuipers, B.J.H., and Gruppen, H. (2007). Prediction of molar extinction coefficients of proteins and peptides using UV absorption of the constituent amino acids at 214 nm to enable quantitative reverse phase high-performance liquid chromatography-mass spectrometry analysis. J. Agric. Food Chem.

Kyhse-Andersen, J. (1984). Electroblotting of multiple gels: a simple apparatus without buffer tank for rapid transfer of proteins from polycrylamide to nitrocellulose. J. Biochem. Biophys. Methods.

Laemmli, U.K. (1970). Cleavage of structural proteins during the assembly of the head of bacteriophage T4. Nature.

Lemmon, M.A., and Schlessinger, J. (2010). Cell signaling by receptor tyrosine kinases. Cell.

Lemmon, M.A., Schlessinger, J., and Ferguson, K.M. (2014). The EGFR family: Not so prototypical receptor tyrosine kinases. Cold Spring Harb. Perspect. Biol.

Lev, S., Blechman, J., Nishikawa, S., Givol, D., and Yarden, Y. (1993). Interspecies molecular chimeras of kit help define the binding site of the stem cell factor. Mol. Cell. Biol.

Levkowitz, G. (1996). Coupling of the c-Cbl protooncogene product to ErbB-1/EGF-receptor but not to other ErbB proteins. Oncogene.

Li, H., and Villalobo, A. (2002). Evidence for the direct interaction between calmodulin and the human epidermal growth factor receptor. Biochem. J.

Li, H., Ruano, M.J., and Villalobo, A. (2004). Endogenous calmodulin interacts with the epidermal growth factor receptor in living cells. FEBS Lett.

Li, H., Panina, S., Kaur, A., Ruano, M.J., Sánchez-González, P., La Cour, J.M., Stephan, A., Olesen, U.H., Berchtold, M.W., and Villalobo, A. (2012). Regulation of the ligand-dependent activation of the epidermal growth factor receptor by calmodulin. J. Biol. Chem.

Lichtner, R.B., Menrad, A., Sommer, A., Klar, U., and Schneider, M.R. (2001). Signaling-inactive epidermal growth factor receptor/ligand complexes in intact carcinoma cells by quinazoline tyrosine kinase inhibitors. Cancer Res.

Lietzke, S.E., Bose, S., Cronin, T., Klarlund, J., Chawla, A., Czech, M.P., and Lambright, D.G. (2000). Structural basis of 3-phosphoinositide recognition by Pleckstrin homology domains. Mol. Cell.

Lu, C., Mi, L.-Z., Grey, M.J., Zhu, J., Graef, E., Yokoyama, S., and Springer, T.A. (2010). Structural Evidence for Loose Linkage between Ligand Binding and Kinase Activation in the Epidermal Growth Factor Receptor. Mol. Cell. Biol.

Lu, C., Mi, L.Z., Schurpf, T., Walz, T., and Springer, T.A. (2012). Mechanisms for kinase-mediated dimerization of the epidermal growth factor receptor. J Biol Chem *287*, 38244–38253.

Malaby, A.W., van den Berg, B., and Lambright, D.G. (2013). Structural basis for membrane recruitment and allosteric activation of cytohesin family Arf GTPase exchange factors. Proc. Natl. Acad. Sci.

Malaby, A.W., Das, S., Chakravarthy, S., Irving, T.C., Bilsel, O., and Lambright, D.G. (2018). Structural Dynamics Control Allosteric Activation of Cytohesin Family Arf GTPase Exchange Factors. Structure.

Markman, B., Javier Ramos, F., Capdevila, J., and Tabernero, J. (2010). EGFR and KRAS in Colorectal Cancer.

Martín-Nieto, J., and Villalobo, A. (1998). The human epidermal growth factor receptor contains a juxtamembrane calmodulin-binding site. Biochemistry.

Masuda, H., Zhang, D., Bartholomeusz, C., Doihara, H., Hortobagyi, G.N., and Ueno, N.T. (2012). Role of epidermal growth factor receptor in breast cancer. Breast Cancer Res. Treat.

McLaughlin, S., Smith, S.O., Hayman, M.J., and Murray, D. (2005). An Electrostatic Engine Model for Autoinhibition and Activation of the Epidermal Growth Factor Receptor (EGFR/ErbB) Family. J. Gen. Physiol.

Meacci, E., Tsai, S.-C., Adamik, R., Moss, J., and Vaughan, M. (1997). Cytohesin-1, a cytosolic guanine nucleotide-exchange protein for ADP-ribosylation factor. Proc. Natl. Acad. Sci.

De Meyts, P. (2008). The insulin receptor: a prototype for dimeric, allosteric membrane receptors?

References

Trends Biochem. Sci.

Mi, L.Z., Grey, M.J., Nishida, N., Walz, T., Lu, C., and Springer, T.A. (2008). Functional and structural stability of the epidermal growth factor receptor in detergent micelles and phospholipid nanodiscs. Biochemistry.

Mi, L.Z., Lu, C., Li, Z., Nishida, N., Walz, T., and Springer, T.A. (2011). Simultaneous visualization of the extracellular and cytoplasmic domains of the epidermal growth factor receptor. Nat Struct Mol Biol *18*, 984–989.

Michalczyk, A., Klüter, S., Rode, H.B., Simard, J.R., Grütter, C., Rabiller, M., and Rauh, D. (2008). Structural insights into how irreversible inhibitors can overcome drug resistance in EGFR. Bioorganic Med. Chem.

Mohamed, Z.H., Rhein, C., Saied, E.M., Kornhuber, J., and Arenz, C. (2018). FRET probes for measuring sphingolipid metabolizing enzyme activity. Chem Phys Lipids *216*, 152–161.

Morishige, M., Hashimoto, S., Ogawa, E., Toda, Y., Kotani, H., Hirose, M., Wei, S., Hashimoto, A., Yamada, A., Yano, H., et al. (2008). GEP100 links epidermal growth factor receptor signalling to Arf6 activation to induce breast cancer invasion. Nat. Cell Biol.

Mossessova, E., Gulbis, J.M., and Goldberg, J. (1998). Structure of the guanine nucleotide exchange factor Sec7 domain of human Arno and analysis of the interaction with ARF GTPpase. Cell.

Mossessova, E., Corpina, R.A., and Goldberg, J. (2003). Crystal Structure of ARF1•Sec7 Complexed with Brefeldin A and Its Implications for the Guanine Nucleotide Exchange Mechanism. Mol. Cell.

Murphy, J.E., Padilla, B.E., Hasdemir, B., Cottrell, G.S., and Bunnett, N.W. (2009). Endosomes: A legitimate platform for the signaling train. Proc. Natl. Acad. Sci.

Nojiri, S., and Hoek, J. (2000). Suppression of epidermal growth factor-induced phospholipase C activation associated with actin rearrangement in rat hepatocytes in primary culture. Hepatology.

Novy, Z., Barta, P., Mandikova, J., Laznicek, M., and Trejtnar, F. (2012). A comparison of in vitro methods for determining the membrane receptor expression in cell lines. Nucl. Med. Biol. *39*, 893–896.

Pan, T., Sun, J., Hu, J., Hu, Y., Zhou, J., Chen, Z., Xu, D., Xu, W., Zheng, S., and Zhang, S. (2014). Cytohesins/ARNO: The function in colorectal cancer cells. PLoS One.

Park, E., Kim, N., Ficarro, S.B., Zhang, Y., Lee, B.I., Cho, A., Kim, K., Park, A.K.J., Park, W.Y., Murray, B., et al. (2015). Structure and mechanism of activity-based inhibition of the EGF receptor by Mig6. Nat. Struct. Mol. Biol.

Patricelli, M.P., Szardenings, A.K., Liyanage, M., Nomanbhoy, T.K., Wu, M., Weissig, H., Aban, A., Chun, D., Tanner, S., and Kozarich, J.W. (2007). Functional interrogation of the kinome using nucleotide acyl phosphates. Biochemistry.

Pawson, T. (2004). Specificity in Signal Transduction: From Phosphotyrosine-SH2 Domain Interactions to Complex Cellular Systems. Cell.

Pellon-Cardenas, O., Clancy, J., Uwimpuhwe, H., and D'Souza-Schorey, C. (2013). ARF6-Regulated Endocytosis of Growth Factor Receptors Links Cadherin-Based Adhesion to Canonical

References

Wnt Signaling in Epithelia. Mol. Cell. Biol.

Pines, G., Huang, P.H., Zwang, Y., White, F.M., and Yarden, Y. (2010). EGFRvIV: A previously uncharacterized oncogenic mutant reveals a kinase autoinhibitory mechanism. Oncogene.

Qi, M., Gross, A., Jeschke, G., Godt, A., and Drescher, M. (2014). Gd(III)-PyMTA Label Is Suitable for In-Cell EPR. J. Am. Chem. Soc. *136*, 15366–15378.

Randazzo, P.A., Jian, X., Chen, P.-W., Zhai, P., Soubias, O., and Northup, J.K. (2013). Quantitative Analysis of Guanine Nucleotide Exchange Factors (GEFs) as Enzymes. Cell. Logist.

Red Brewer, M., Choi, S.H., Alvarado, D., Moravcevic, K., Pozzi, A., Lemmon, M.A., and Carpenter, G. (2009). The juxtamembrane region of the EGF receptor functions as an activation domain. Mol Cell *34*, 641–651.

Reddy, A.S., and Zhang, S. (2013). Polypharmacology: drug discovery for the future. Expert Rev Clin Pharmacol *6*, 41–47.

Renault, L., Guibert, B., and Cherfils, J. (2003). Structural snapshots of the mechanism and inhibition of a guanine nucleotide exchange factor. Nature.

Robinson, M.S. (2015). Forty Years of Clathrin-coated Vesicles. Traffic.

Roskoski, R. (2014). The ErbB/HER family of protein-tyrosine kinases and cancer. Pharmacol. Res.

Rouhana, J., Hoh, F., Estaran, S., Henriquet, C., Boublik, Y., Kerkour, A., Trouillard, R., Martinez, J., Pugnière, M., Padilla, A., et al. (2013). Fragment-based identification of a locus in the Sec7 domain of arno for the design of protein-protein interaction inhibitors. J. Med. Chem.

Sabe, H., Hashimoto, S., Morishige, M., Ogawa, E., Hashimoto, A., Nam, J.M., Miura, K., Yano, H., and Onodera, Y. (2009). The EGFR-GEP100-Arf6-AMAP1 signaling pathway specific to breast cancer invasion and metastasis. Traffic.

Sako, Y., Minoghchi, S., and Yanagida, T. (2000). Single-molecule imaging of EGFR signalling on the surface of living cells. Nat. Cell Biol.

San Jose, E., Benguria, A., Geller, P., and Villalobo, A. (1992). Calmodulin inhibits the epidermal growth factor receptor tyrosine kinase. J. Biol. Chem.

Sánchez-González, P., Jellali, K., and Villalobo, A. (2010). Calmodulin-mediated regulation of the epidermal growth factor receptor. FEBS J. *277*, 327–342.

Sato, T., Pallavi, P., Golebiewska, U., McLaughlin, S., and Smith, S.O. (2006). Structure of the membrane reconstituted transmembrane-juxtamembrane peptide EGFR(622-660) and its interaction with Ca2+/Calmodulin. Biochemistry.

Schiemann, O., and Prisner, T.F. (2007). Long-range distance determinations in biomacromolecules by EPR spectroscopy. Q Rev Biophys *40*, 1–53.

Schiemann, O., Piton, N., Mu, Y.G., Stock, G., Engels, J.W., and Prisner, T.F. (2004). A PELDOR-based nanometer distance ruler for oligonucleotides. J. Am. Chem. Soc. *126*, 5722–5729.

Schulze-Gahmen, U., De Bondt, H.L., and Kim, S.H. (1996). High-resolution crystal structures of human cyclin-dependent kinase 2 with and without ATP: Bound waters and natural ligand as guides for inhibitor design. J. Med. Chem.

Seidel, S.A.I., Dijkman, P.M., Lea, W.A., van den Bogaart, G., Jerabek-Willemsen, M., Lazic, A., Joseph, J.S., Srinivasan, P., Baaske, P., Simeonov, A., et al. (2013). Microscale thermophoresis

quantifies biomolecular interactions under previously challenging conditions. Methods.

Sengupta, P., Ruano, M.J., Tebar, F., Golebiewska, U., Zaitseva, I., Enrich, C., McLaughlin, S., and Villalobo, A. (2007). Membrane-permeable calmodulin inhibitors (e.g. W-7/W-13) bind to membranes, changing the electrostatic surface potential: Dual effect of W-13 on epidermal growth factor receptor activation. J. Biol. Chem.

Sengupta, P., Bosis, E., Nachliel, E., Gutman, M., Smith, S.O., Mihályné, G., Zaitseva, I., and McLaughlin, S. (2009). EGFR juxtamembrane domain, membranes, and calmodulin: Kinetics of their interaction. Biophys. J.

Shevelev, G.Y., Krumkacheva, O.A., Lomzov, A.A., Kuzhelev, A.A., Rogozhnikova, O.Y., Trukhin, D. V, Troitskaya, T.I., Tormyshev, V.M., Fedin, M. V, Pyshnyi, D. V, et al. (2014). Physiological-Temperature Distance Measurement in Nucleic Acid using Triarylmethyl-Based Spin Labels and Pulsed Dipolar EPR Spectroscopy. J. Am. Chem. Soc. *136*, 9874–9877.

Shibuya, M., and Suzuki, Y. (1993). [Treatment of cerebral vasospasm by a protein kinase inhibitor AT 877]. No To Shinkei *45*, 819–824.

Sigismund, S., Woelk, T., Puri, C., Maspero, E., Tacchetti, C., Transidico, P., Di Fiore, P.P., and Polo, S. (2005). Clathrin-independent endocytosis of ubiquitinated cargos. Proc. Natl. Acad. Sci.

Singh, J., Petter, R.C., Baillie, T.A., and Whitty, A. (2011). The resurgence of covalent drugs. Nat Rev Drug Discov *10*, 307–317.

Soria, J.C., Mok, T.S., Cappuzzo, F., and Jänne, P.A. (2012). EGFR-mutated oncogene-addicted non-small cell lung cancer: Current trends and future prospects. Cancer Treat. Rev.

Sorkin, A., and Carpenter, G. (1993). Interaction of activated EGF receptors with coated pit adaptins. Science (80-.).

Sorkin, A., Mazzotti, M., Sorkina, T., Scotto, L., and Beguinot, L. (1996). Epidermal growth factor receptor interaction with clathrin adaptors is mediated by the Tyr974-containing internalization motif. J. Biol. Chem.

Stalder, D., Barelli, H., Gautier, R., Macia, E., Jackson, C.L., and Antonny, B. (2011). Kinetic studies of the arf activator arno on model membranes in the presence of arf effectors suggest control by a positive feedback loop. J. Biol. Chem.

Stamos, J., Sliwkowski, M.X., and Eigenbrot, C. (2002). Structure of the epidermal growth factor receptor kinase domain alone and in complex with a 4-anilinoquinazoline inhibitor. J. Biol. Chem.

Stumpf, A.D., and Hoffmann, C. (2016). Optical probes based on G protein-coupled receptors - added work or added value? Br J Pharmacol *173*, 255–266.

Taylor, S.S., and Radzio-Andzelm, E. (1994). Three protein kinase structures define a common motif. Structure.

Theillet, F.X., Binolfi, A., Bekei, B., Martorana, A., Rose, H.M., Stuiver, M., Verzini, S., Lorenz, D., van Rossum, M., Goldfarb, D., et al. (2016). Structural disorder of monomeric alpha-synuclein persists in mammalian cells. Nature *530*, 45–50.

Theis, M.G., Knorre, A., Kellersch, B., Moelleken, J., Wieland, F., Kolanus, W., and Famulok, M. (2004). Discriminatory aptamer reveals serum response element transcription regulated by cytohesin-2. Proc. Natl. Acad. Sci. U. S. A.

References

Tomshine, J.C., Severson, S.R., Wigle, D.A., Sun, Z., Beleford, D.A.T., Shridhar, V., and Horazdovsky, B.F. (2009). Cell proliferation and epidermal growth factor signaling in non-small cell lung adenocarcinoma cell lines are dependent on Rin1. J. Biol. Chem.

Traut, T.W. (1994). Physiological concentrations of purines and pyrimidines. Mol. Cell. Biochem.

Truong, K., and Ikura, M. (2001). The use of FRET imaging microscopy to detect protein-protein interactions and protein conformational changes in vivo. Curr. Opin. Struct. Biol. *11*, 573–578.

Tsvetkov, Y.D., Milov, A.D., and Maryasov, A.G. (2008). Pulsed electron–electron double resonance (PELDOR) as EPR spectroscopy in nanometre range. Russ. Chem. Rev.

Ullrich, A., Coussens, L., Hayflick, J.S., Dull, T.J., Gray, A., Tam, A.W., Lee, J., Yarden, Y., Libermann, T.A., Schlessinger, J., et al. (1984). Human epidermal growth factor receptor cDNA sequence and aberrant expression of the amplified gene in A431 epidermoid carcinoma cells. Nature.

Ullrich, O., Reinsch, S., Urbé, S., Zerial, M., and Parton, R.G. (1996). Rab11 regulates recycling through the pericentriolar recycling endosome. J. Cell Biol.

Urferl, R., Tsoulfas2 ', P., O 'connell, L., Shelton4, D.L., Parada2, L.F., and Presta, L.G. (1995). An immunoglobulin-like domain determines the specificity of neurotrophin receptors. EMBO J.

Uyemura, T., Takagi, H., Yanagida, T., and Sako, Y. (2005). Single-molecule analysis of epidermal growth factor signaling that leads to ultrasensitive calcium response. Biophys. J.

Villaseñor, R., Kalaidzidis, Y., and Zerial, M. (2016). Signal processing by the endosomal system. Curr. Opin. Cell Biol.

Walter, A.O., Sjin, R.T.T., Haringsma, H.J., Sun, J., Ohashi, K., Lee, K., Dubrovskiy, A., Labenski, M., Wang, Z., Zhu, Z., et al. (2013). Discovery of a mutant-selective covalent inhibitor of EGFR that overcomes T790M mediated resistance in NSCLC. Cancer Discov.

Wang, Y., Pennock, S., Chen, X., and Wang, Z. (2002). Endosomal Signaling of Epidermal Growth Factor Receptor Stimulates Signal Transduction Pathways Leading to Cell Survival. Mol. Cell. Biol.

Wang, Y., Gao, J., Guo, X., Tong, T., Shi, X., Li, L., Qi, M., Wang, Y., Cai, M., Jiang, J., et al. (2014). Regulation of EGFR nanocluster formation by ionic protein-lipid interaction. Cell Res.

Ward, C.W., Lawrence, M.C., Streltsov, V.A., Adams, T.E., and McKern, N.M. (2007). The insulin and EGF receptor structures: new insights into ligand-induced receptor activation. Trends Biochem. Sci.

Wee, P., and Wang, Z. (2017). Epidermal growth factor receptor cell proliferation signaling pathways. Cancers (Basel).

Weisberg, E., Manley, P.W., Breitenstein, W., Brüggen, J., Cowan-Jacob, S.W., Ray, A., Huntly, B., Fabbro, D., Fendrich, G., Hall-Meyers, E., et al. (2005). Characterization of AMN107, a selective inhibitor of native and mutant Bcr-Abl. Cancer Cell.

Wlesmann, C., Ultsch, M.H., Bass, S.H., and De Vos, A.M. (1999). Crystal structure of nerve growth factor in complex with the ligand- binding domain of the TrkA receptor. Nature.

Wofsy, L., Metzger, H., and Singer, S.J. (1962). Affinity labeling-a general method for labeling the active sites of antibody and enzyme molecules. Biochemistry *1*, 1031–1039.

Wood, E.R., Truesdale, A.T., McDonald, O.B., Yuan, D., Hassell, A., Dickerson, S.H., Ellis, B.,

References

Pennisi, C., Horne, E., Lackey, K., et al. (2004). 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 Res.

Wu, J., Tseng, Y.D., Xu, C.F., Neubert, T.A., White, M.F., and Hubbard, S.R. (2008). Structural and biochemical characterization of the KRLB region in insulin receptor substrate-2. Nat. Struct. Mol. Biol.

Xu, W., Doshi, A., Lei, M., Eck, M.J., and Harrison, S.C. (1999). Crystal structures of c-Src reveal features of its autoinhibitory mechanism. Mol. Cell.

Yarden, Y., and Schlessinger, J. (1987a). Self-Phosphorylation of Epidermal Growth Factor Receptor: Evidence for a Model of Intermolecular Allosteric Activation. Biochemistry.

Yarden, Y., and Schlessinger, J. (1987b). Epidermal Growth Factor Induces Rapid, Reversible Aggregation of the Purified Epidermal Growth Factor Receptor. Biochemistry.

Yarden, Y., and Sliwkowski, M.X. (2001). Untangling the ErbB signalling network. Nat. Rev. Mol. Cell Biol.

Yin, D.M., Hannam, J.S., Schmitz, A., Schiemann, O., Hagelueken, G., and Famulok, M. (2017). Studying the Conformation of a Receptor Tyrosine Kinase in Solution by Inhibitor-Based Spin Labeling. Angew. Chemie - Int. Ed.

Yin, D.M., Hammler, D., Peter, M.F., Marx, A., Schmitz, A., and Hagelueken, G. (2018). Inhibitor-Directed Spin Labelling—A High Precision and Minimally Invasive Technique to Study the Conformation of Proteins in Solution. Chem. - A Eur. J.

Young, T.S., Ahmad, I., Yin, J.A., and Schultz, P.G. (2010). An enhanced system for unnatural amino acid mutagenesis in E. coli. J Mol Biol *395*, 361–374.

Yu, X., Sharma, K.D., Takahashi, T., Iwamoto, R., and Mekada, E. (2002). Ligand-independent Dimer Formation of Epidermal Growth Factor Receptor (EGFR) Is a Step Separable from Ligand-induced EGFR Signaling. Mol. Biol. Cell.

Yun, C.H., Boggon, T.J., Li, Y., Woo, M.S., Greulich, H., Meyerson, M., and Eck, M.J. (2007). Structures of Lung Cancer-Derived EGFR Mutants and Inhibitor Complexes: Mechanism of Activation and Insights into Differential Inhibitor Sensitivity. Cancer Cell.

Zhang, X., Gureasko, J., Shen, K., Cole, P.A., and Kuriyan, J. (2006). An allosteric mechanism for activation of the kinase domain of epidermal growth factor receptor. Cell *125*, 1137–1149.

Zhang, X., Pickin, K.A., Bose, R., Jura, N., Cole, P.A., and Kuriyan, J. (2007). Inhibition of the EGF receptor by binding of MIG6 to an activating kinase domain interface. Nature.

Zhou, W., Ercan, D., Chen, L., Yun, C.H., Li, D., Capelletti, M., Cortot, A.B., Chirieac, L., Iacob, R.E., Padera, R., et al. (2009). Novel mutant-selective EGFR kinase inhibitors against EGFR T790M. Nature.

Acknowledgments

Studying overseas is difficult, studying as Ph.D. in a foreign land over 9000 kilometers away from home is only harder. Without the help and support from colleagues, friends and families, I could not walk through this long journey. I dedicate this chapter to those who helped me or lent a hand along the road, and I am deeply grateful for their warm hearts and kindness.

To begin with, I would like to thank Prof. Michael Famulok for hosting such an interdisciplinary lab, which provides a highly dynamic and interactive environment for creative scientific work. As the leader of multiple academic organizations, Prof. Famulok managed to discuss project details with students on a regular basis, with which he demonstrates to me his motivation and passion as a senior scientist.

I want to thank Prof. Matthias Geyer, who leads the neighboring lab of structural immunology, for the discussions concerning our collaboration projects, during which he often impressed me with his knowledge and deep understanding of protein structures, as well as his sense of humor.

I also thank Prof. Olav Schiemann and Dr. Gregor Hagelüken for the successful collaboration, from which I learnt so much about EPR techniques for studying protein dynamics. As my collaboration partner, Dr. Hagelüken impressed me with his broad knowledge of protein science, creative thinking and efficient execution. I truly enjoyed our interesting and fruitful collaboration.

I am grateful to Dr. Anton Schmitz, who supervised all the operational sides of my projects. I learnt from Toni his logical thinking and rigorous control of the design and execution of a scientific project, as well as his open-mindedness for ideas from other research fields. Toni especially surprised me with his keen observation during the data analyses, which sometimes gave a completely different perspective and opportunities for our next move. We also exchanged opinions about politics, history and cultures, and through him I get familiarized with this country. I also want to thank him and Dr. Sven Freudenthal for their thoughtfulness regarding my visa, contract and other administrative issues.

I thank Dr. Manuel Etzkorn and Dr. Aldino Viegas for our successful collaboration.

I also thank my beloved colleagues/friends from Famulok group and Caesar institute, especially Christian Sieg and Volkmar Fieberg. Christian helped me really a lot both in

and out of the lab on so many small things, and I am constantly impressed by his thoughtfulness and attention to details. Volki assisted me on protein purification quite often and surprised me once in a while with his knowledge of ÄKTA systems. Also, out of the lab we are the best hiking mates, and I am still looking forward to our next hiking trip after my defense. I also want to thank Sebastian ("Basti") Schmitt from Matthias Geyer lab for cheering me up with his endless jokes whenever I was frustrated by experiments, and giving me project/career advice as a postdoc. Hussein Hamzeh from Benjamin Kaupp lab also lent me hands on so many occasions. We started the Ph.D. in the same month, and helped each other throughout the journey. I wish him a successful postdoc in the lab and good luck with job hunting in future.

I will dedicate this paragraph to my dear flat mate Stefan Metzler. Stefano has introduced me to the German culture, which was completely strange to me five years ago. I enjoyed his cooking and our theater/concert nights a lot. Also, in everyday life, he helped me with language learning and all kinds of administrative paperwork. Thank you Stefano, no one can find a better flat mate than you!

Next, I thank my friends from the good old China, Zhuoyan ("Anthony") Chen, An Gong, Zhiyang Tan, Jinpeng Su, Huipeng Jiao and Eva Xu. As expatriates, we supported and took care of each other along the road. I will cherish the good time we had together.

Last but not least, I would thank my parents and families, for their open-mindedness to support my study overseas. You are the strength for my scientific pursuit and for that, I shall be forever grateful.