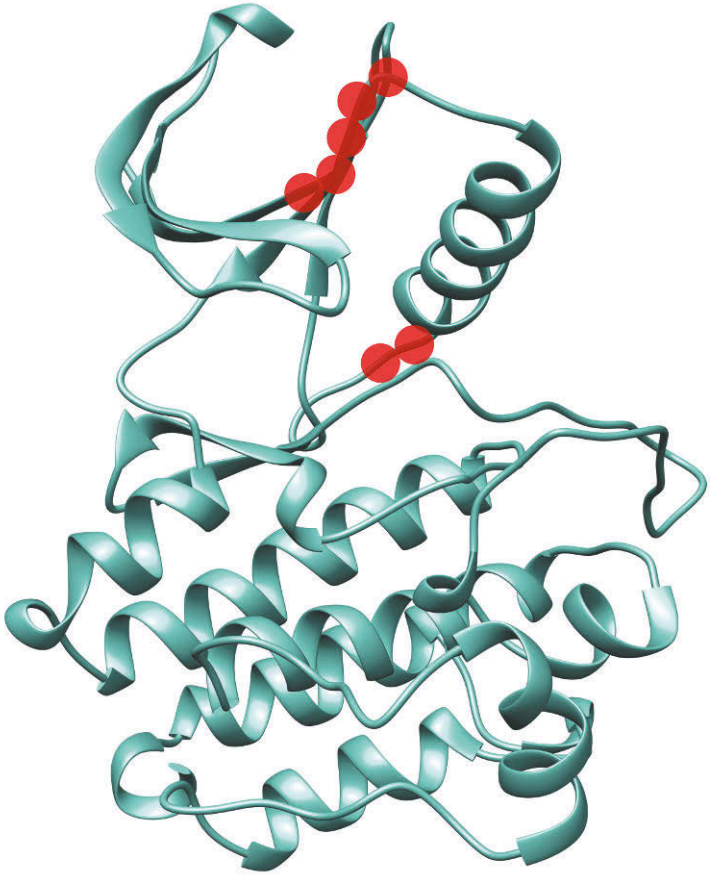
Mahlet Zerihun Tamirat

The Dynamic Structural Effects of Activating ERBB Kinase Somatic Mutations



Åbo Akademi University

Mahlet Zerihun Tamirat

Born 1989, Addis Ababa, Ethiopia

Previous studies and degrees

Master of Science in Bioinformatics, University of Turku, 2016

Bachelor of Pharmacy, Addis Ababa University, 2011

The Dynamic Structural Effects of Activating ERBB Kinase Somatic Mutations

Mahlet Zerihun Tamirat

Biochemistry
Faculty of Science and Engineering
Åbo Akademi University
Åbo, Finland, 2021

From the Faculty of Science and Engineering, Åbo Akademi University; the InFLAMES Research Flagship Center; Åbo Akademi Graduate School and National Doctoral Programme in Informational and Structural Biology

Supervised by

Professor Mark S. Johnson
Structural Bioinformatics Laboratory, Biochemistry
Faculty of Science and Engineering
InFLAMES Research Flagship Center
Åbo Akademi University
Turku, Finland

Reviewed by

Associate Professor André Juffer
Biocenter Oulu
Faculty of Biochemistry and Molecular Medicine
University of Oulu
Oulu, Finland

Adjunct Professor Tuomo Laitinen
School of Pharmacy
Faculty of Health Sciences
University of Eastern Finland
Kuopio, Finland

Opponent

Adjunct Professor Maija Lahtela-Kakkonen
School of Pharmacy
Faculty of Health Sciences
University of Eastern Finland
Kuopio, Finland

ISBN 978-952-12-4104-8 (Printed)
ISBN 978-952-12-4105-5 (Digital)
Painosalama, Turku, Finland, 2021

Acknowledgements

This thesis work was carried out at the Structural Bioinformatics Laboratory (SBL), Faculty of Science and Engineering, Åbo Akademi University, during 2016-2021. I am indebted to Professor *Mark Johnson* and Professor *Tiina Salminen* for allowing me to join SBL and discover the wonderful science and people I have worked with.

I would like to thank my supervisor Professor *Mark Johnson* for providing me the opportunity to work under his supervision. I am grateful for his unwavering support and guidance throughout my PhD journey. I am equally grateful for the positive work environment he has created in the lab, which made the PhD journey more enjoyable.

I would like to thank my second supervisor Adjunct Professor *Tomi Airenne* for his encouragement and valuable discussions. I would also like to express my heartfelt gratitude to Adjunct Professor *Outi Salo-Ahen*, whose advice, discussions and kindness have impacted me for the better. I thank Adjunct Professor *Tomi Airenne*, Adjunct Professor *Outi Salo-Ahen* and Dr. *Sanna Soini* for being the members of my thesis committee and for shaping the trajectory of my PhD.

My deep gratitude goes to Associate Professor *André Juffer* and Adjunct Professor *Tuomo Laitinen* for reviewing my thesis and for providing their valuable comments. I would also like to extend my gratitude to Professor *Klaus Elenius* and his research group for our fruitful collaborations. The prompt and continuous scientific IT support I received from Dr. *Jukka Lehtonen* has highly benefited my research at SBL. Thank you *Jukka* for this and for the interesting coffee break chats.

I would like to acknowledge the National Doctoral Programme in Informational and Structural Biology graduate school, which funded me to embark on this scientific adventure. I am thankful for its director Professor *Mark Johnson* and coordinator *Fredrik Karlsson*. Thank you *Fredrik* for all the assistance during the last stages of my PhD.

I am enormously grateful for the financial support I received from multiple parties including Åbo Akademi University, the Magnus Ehrnrooth Foundation, the Maud Kuistila Memorial Foundation, the K. Albin Johansson Foundation and the Orion Research Foundation.

The time I spent at SBL has been enjoyable because of the colleagues I had the privilege to work and socialize with. I want to thank *Vipin*, *Nitin*,

Parthiban and Polytimi for their lovely company at work and outside SBL. I am grateful for *Mia*, *Ida* and *Marion,* who have been the much needed support and source of laughter, especially during the remote working days. Thank you *Mia* for the helpful tips you gave me near the end of my studies and for being a wonderful PhD companion. My friends outside SBL have helped me achieve work-life balance and brought me comfort and happiness. For this, I want to thank *Nebi*, *Bleno*, *Aki*, *Fitse*, *Ida* and *Lilly*.

Finally, I am eternally grateful for my family, who has given me unconditional love and support. አባቴ አቶ ዘርይሁን ታምራት ፣ እናቴ ወ/ሮ አቻምየለሽ ነዋይ ፣ ወንድሞቼ ንሴብሉ ፣ ያሬድ እና አምሀእየሱስ እንዲሁም እህቶቼ ሰብለወንጌል ፣ ዮዲት ፣ ሶስና እና ፍቅርተ ስልድጋፋቹ ከልቤ አመሠግናለሁ። ባለቤቴ ዶ/ር በኃይሉ ሺፈራው ሁል ጊዜ ከጎኔ ሆነህ ስላጠነከርከኝ ፣ ብርታት ስለሆንከኝ፣ በኔ ላይ ስላለህ እምነት ከልቤ አመሠግናለሁ። ይህ መመረቂያ ፅሑፍ ከኔ ልፋት እኩል ያንተ ነው። በመጨረሻም ሁሉ እንዲሳካ የፈቀደ እግዚአብሔር ይመስገን።

Abstract

The ERBB family of receptor tyrosine kinases, epidermal growth factor receptor (EGFR, ERBB1), ERBB2, ERBB3 and ERBB4, are transmembrane signaling proteins that regulate cellular processes such as cell survival, mobility, proliferation and differentiation. Normal activity of the ERBBs is essential for tissue growth and organ development. An array of somatic mutations of the *ERBB* genes have been linked to human malignancies. As a result, the ERBBs are important treatment targets, with multiple ERBB-based drugs currently in effective clinical use against e.g. lung, breast and colorectal cancers.

A comprehensive characterization of the ERBB somatic mutations identified in cancer samples is essential to unveil the molecular level impacts of the genetic alterations that could play a role in tumorigenesis. In this thesis, the structural consequences of four cancer-associated ERBB kinase mutations that aberrantly activate EGFR and ERBB2 proteins were explored. The mutations include three EGFR alterations, 746ΔELREA750 (ΔELREA), V769insASV and D770insNPG, and an ERBB2 missense mutation: E936K. The ERBB receptors and proteins in general are dynamic molecules. Hence, the possible structural changes exerted by the above activating mutations were examined by employing molecular dynamics simulations, which allow the assessment of time-dependent structural motions.

The simulations revealed that the EGFR ΔELREA deletion mutation stabilizes the active state EGFR conformation by conserving the states of key structural units, such as the αC helix and the Lys745-Glu762 salt bridge, which were disrupted in the wild-type EGFR. Furthermore, the deletion resulted in a structural change on the inactive EGFR state, an inward movement of the αC helix, which could drive a conformational change from the inactive towards the active EGFR state. The V769insASV and D770insNPG EGFR insertion mutations also led to the better stability of the active EGFR conformation relative to the wild-type EGFR. Moreover, the insertions obstructed the formation of an autoinhibitory interaction between Ala767 and Arg776 in the inactive EGFR conformation, which would predispose EGFR to transition to the catalytically active EGFR state. The ERBB2 E936K mutation affected the nature of interactions taking place at the ERBB2 homodimer and heterodimer interface, with a new inter-monomer ionic interaction being formed that strengthened the monomer-monomer binding. Consequently, the duration of the activated ERBB2 dimer would be extended, which fuels the phosphorylation of ERBB2.

Taken together, this thesis demonstrated that a series of structural changes are at play that collectively elicit the experimentally reported functional changes by these activating ERBB mutations. A thorough

examination of the mutation-induced structural alterations furthers our knowledge on the role the mutations play in cancer progression and the results are essential knowledge when using structure-based, rational design of ligands that could inhibit ERBB kinase activity and subsequent receptor signaling. Such ligands have potential for further development towards a therapeutic agent in efforts to tackle cancers.

Sammanfattning

ERBB-familjen av tyrosinkinasreceptorer, epidermala tillväxtfaktor-receptorer (EGFR, ERBB1), ERBB2, ERBB3 och ERBB4 är transmembrana signalproteiner som reglerar cellulära processer så som överlevnad, mobilitet, proliferation och differentiering. Normal aktivitet hos ERBB är nödvändigt för vävnadstillväxt och organutveckling. En mängd somatiska mutationer i ERBB generna har kopplats till cancer. På grund av detta är ERBB-proteinerna viktiga mål för behandling och flera ERBB-baserade läkemedel är för tillfället i klinisk användning, t.ex. vid lung-, bröst- och kolorektalcancer.

En omfattande karakterisering av de somatiska mutationerna i ERBB, som identifierats i cancerprover, är nödvändig för att påvisa vilken effekt de genetiska förändringarna på en molekylärnivå kan ha för tumörutveckling. I den här avhandlingen utforskades de strukturella konsekvenserna av fyra cancer-associerade ERBB kinas-mutationer som felaktigt aktiverar EGFR och ERBB2 proteiner. Mutationerna inkluderar tre EGFR förändringar, 746ΔELREA750 (ΔELREA), V769insASV och D770insNPG, och en ERBB2 missense-mutation: E936K. ERBB receptorer och proteiner är dynamiska molekyler och därför användes molekyldynamiska simulationer för att studera de möjliga strukturella förändringar som de ovan nämnda aktiverande mutationerna orsakar, vilket ger möjlighet att utvärdera tidsberoende strukturella rörelser.

Simulationerna visade att EGFR ΔELREA deletion-mutationen stabiliserar det aktiva stadies konformation i EGFR genom att bevara stadierna för viktiga strukturella enheter, så som helix αC och saltbryggan mellan Lys745 och Glu762, som båda avbryts i wild-type EGFR. Deletion-mutationen resulterade dessutom i en strukturell förändring i det inaktiva EGFR stadiet, en inåtgående rörelse i helix αC som kunde orsaka en konformationsförändring från det inaktiva till det aktiva stadiet i EGFR. V769insASV och D770insNPG EGFR insertion-mutationerna ledde också till bättre stabilitet för den aktiva konfirmationen i EGFR i förhållande till wild-type EGFR. Insertion-mutationerna hindrade också bildningen av en autoinhiberande interaktion mellan Ala767 och Arg776 i den inaktiva konformationen för EGFR, vilket skulle göra det mera troligt att EGFR övergår till det katalytiskt aktiva stadiet. E936K-mutationen i ERBB2 påverkade typen av interaktioner som sker vid ERBB2:s homodimera- och heterodimera gränsytor, där en ny jonisk interaktion formas mellan monomererna och stärkte monomer-monomer bindningen. Följaktligen förlängs varaktigheten hos den aktiverade ERBB2-dimeren, vilket ökar fosforyleringen av ERBB2.

Sammanfattningsvis visar den här avhandlingen att en serie av strukturella förändringar spelar en roll i att kollektivt orsaka de experimentellt rapporterade funktionella förändringarna genom de

aktiverande ERBB-mutationerna. En grundlig genomgång av de mutationsinducerade förändringarna ökar vår kunskap om vilken roll mutationer spelar i cancerprogression, och resultaten är nödvändig kunskap när man använder sig av strukturbaserad, rationell design av ligander som kunde inhibera ERBB kinas-aktiviteten och påföljande receptorsignalering. Dylika ligander kan potentiellt vidareutvecklas till en ny behandlingsmetod för cancer.

Table of Contents

List of original publications

This thesis is based on the following four original publications, which are referred to in Roman numerals in the thesis.

I. **Tamirat MZ**, Koivu M, Elenius K, Johnson MS. Structural characterization of EGFR exon 19 deletion mutation using molecular dynamics simulation. PLoS One. 2019;14(9): e0222814.

II. **Tamirat MZ**, Kurppa KJ, Elenius K, Johnson MS. Structural basis for the functional changes by EGFR exon 20 insertion mutations. Cancers (Basel). 2021;13(5):1–21.

III. Koivu MKA, Chakroborty D, **Tamirat MZ**, Johnson MS, Kurppa KJ, Elenius K. Identification of predictive ERBB mutations by leveraging publicly available cell line databases. Molecular Cancer Therapeutics. 2021;20(3):564–76.

IV. **Tamirat MZ**, Kurppa KJ, Elenius K, Johnson MS. Deciphering the structural effects of activating EGFR somatic mutations with molecular dynamics simulation. Journal of Visualized Experiments. 2020;2020(159):e61125.

Contributions of the author

The author of this thesis conducted the computational structural bioinformatics work covered in publications I-IV. This includes carrying out homology modeling, molecular dynamics simulations and analysis of the structural data. The author was involved in the conceptual design of the studies in publication I and II. The author led the preparation of the manuscripts for publication I, II and IV, and wrote the structural modeling sections in publication III.

Additional publications

Ranga V, Niemelä E, **Tamirat MZ**, Eriksson JE, Airenne TT, Johnson MS. Immunogenic SARS-CoV-2 Epitopes: In Silico Study Towards Better Understanding of COVID-19 Disease—Paving the Way for Vaccine Development. Vaccines. 2020;8(3):408.

Ashok Y, Miettinen M, Oliveira DKH de, **Tamirat MZ**, Näreoja K, Tiwari A, Hottiger MO, Johnson MS, Lehtiö L, Pulliainen AT. Discovery of compounds inhibiting the ADPribosyltransferase activity of pertussis toxin. ACS Infectious Diseases. 2020. acsinfecdis.9b00412.

Chakroborty D, Kurppa KJ, Paatero I, Ojala VK, Koivu M, **Tamirat MZ**, Koivunen JP, Jänne PA, Johnson MS, Elo LL, Elenius K. An unbiased in vitro screen for activating epidermal growth factor receptor mutations. Journal of Biological Chemistry. 2019;294(24):9377–89.

Vuoristo S, Bhagat S, Hydén-Granskog C, Yoshihara M, Gawriyski L, Jouhilahti EM, Ranga V, **Tamirat M**, Huhtala M, Kirjanov I, Nykänen S, Krjutškov K, Damdimopoulos A, Weltner J, Hashimoto K, Recher G, Ezer S, Paluoja P, Paloviita P, Takegami Y, Kanemaru A, Lundin K, Airenne T, Otonkoski T, Tapanainen T, Kawaji H, Murakawa Y, Bürglin TR, Varjosalo M, Johnson MS, Tuuri T, Katayama S, Kere J. DUX4 regulates human embryonic transcripts through repeat-associated enhancers. Cell Reports. 2021. Submitted manuscript.

Chakroborty D, Ojala VK, Knittle AK, Drexler J, **Tamirat MZ**, Ruzicka R, Bosch K, Woertl J, Schmittner S, Elo LL, Johnson MS, Kurppa KJ, Solca F, Elenius K. An unbiased screen identifies rare activating ERBB4 mutations. 2021. Manuscript.

Abbreviations

3D	Three-dimensional
AKT	Ak strain transforming
AR	Amphiregulin
ATP	Adenosine triphosphate
BTC	Betacellulin
CDK2	Cyclin dependent kinase 2
cDNA	Complementary DNA
DNA	Deoxyribonucleic acid
EGF	Epidermal growth factor
EGFR	Epidermal growth factor receptor
EGFRvIII	EGFR variant III
EPG	Epigen
EPR	Epiregulin
ERBB	Erythroblastic oncogene B
ERK	Extracellular signal-regulated kinase
Gab1	Grb2-associated binder
HB-EGFR	Hairpin-binding EGF-like growth factor
HER	Human epidermal growth factor receptor
INSR	Insulin receptor
IP3	Inositol 1,4,5-trisphosphate
IRS1	Insulin receptor substrate 1
JAK	Janus kinase
JM	Juxtamembrane
JNK	c-Jun N-terminal kinase
mAB	Monoclonal antibody
MAPK	Mitogen-activated protein kinase
MDS	Molecular dynamics simulation
MM	Molecular mechanics
MM-GBSA	Molecular mechanics generalized Born surface area
NRG	Neuregulin
NSCLC	Non-small cell lung cancer
PC	Principal component
PCA	Principal component analysis
PDB	Protein Data Bank
PDGFR	Platelet-derived growth factor receptor
PDK-1	Phosphoinositide-dependent kinase-1
PH	Pleckstrin homology
PI3K	Phosphatidylinositol 3-kinase
PIP	Phosphatidylinositol (3,4,5)-trisphosphate

PLC-γ1	Phospholipase C-gamma1
PTB	Protein tyrosine binding
RAF	Rapidly accelerated fibrosarcoma
RAS	Rat sarcoma
RMSD	Root-mean-square deviation
RMSF	Root-mean-square fluctuation
RTK	Receptor tyrosine kinase
SH2	Src homology 2
STAT	Signal transducer and activator of transcription
TFGα	Transforming growth factor α
TKI	Tyrosine kinase inhibitor
TM	Transmembrane

1 Introduction

Cell surface receptor proteins, embeded in the cell membrane, transmit signals from the exterior to the interior of the cell (and also *vice versa* e.g. integrins) facilitating communication between the cell's interior and the extracellular environment (Uings and Farrow 2000). Signal transduction through membrane receptors is evoked by binding of extracellular ligands, such as growth factors and peptide hormones, to the receptor, which induces a variety of cellular functions. Membrane receptors are generally catagorized into three classes: G-protein-coupled receptors, ligand-gated ion channel-linked receptors and enzyme-linked receptors (O'Connor and Adams 2010). A prominent family member of the enzyme-linked cell surface receptors is the receptor tyrosine kinases (RTKs), which have a cytoplasmic tyrosine kinase activity (Robinson et al. 2000; Uings and Farrow 2000). As a protein kinase, RTKs catalyze the transfer of the ɣ-phosphate of ATP to tyrosine residues of substrate molecules, which may be the RTK itself (Figure 1). Indeed, amino acid (mainly serine, threonine or tyrosine) phosphorylation is one of the most common post-translational modifications implemented in signal transduction.

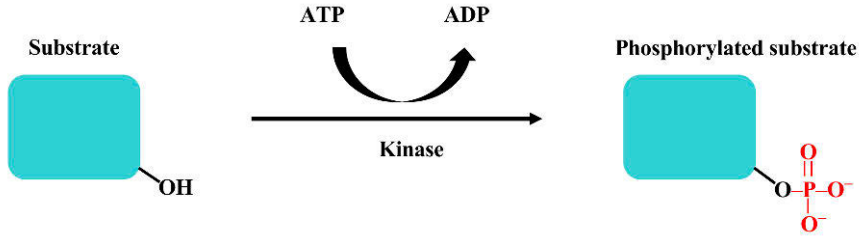


Figure 1. Phosphorylation reaction catalyzed by protein kinases.

A sub-family of the RTKs that are the focus of this thesis are the human ERBB receptor tyrosine kinases. In humans, ERBBs include epidermal growth factor receptor (EGFR)/ERBB1/HER1, ERBB2/HER2, ERBB3/HER3 and ERBB4/HER4 (Roskoski 2014b). The term ERBB is derived from the avian erythroblastosis virus oncogene, v-ERBB, which is homologous to EGFR (Downward et al. 1984). The binding of growth factors to ERBB receptor kinases induces a conformational change that allows the receptors to form homodimers and heterodimers (Leahy 2004); although ERBB2 does not bind a growth factor and ERBB3 is kinase dead, both are still capable of activity via heterodimer formation with other ERBB members. The dimerization of the ERBBs results in the autophosphorylation of the receptors and activation of downstream signaling pathways that modulate biological processes such as cell proliferation, differentiation and apoptosis. The ERBBs are essential for

normal tissue development and homeostasis, however, dysregulation of their activity links the receptors with a variety of diseases.

The ERBB family of receptors are commonly identified oncogenic proteins in cancer (Sweeney et al. 2017; Sanchez-Vega et al. 2018) and hence are pursued for therapeutic interventions, and the tyrosine kinase domain of ERBBs are a potent cancer inducing enzyme transmitted by infectious ocogenic retroviruses (Yamamoto et al. 1983; Downward et al. 1984). Moreover, gene amplification and/or somatic mutations that enhance the activity of the ERBBs play a crucial role in the pathogenesis of cancer (Mishra et al. 2017). The frequency of somatic alterations in EGFR and ERBB2 is more than three-fold higher than ERBB3 and ERBB4, which is consistent with the abundant data describing the oncogenic significance of EGFR and ERBB2. Mutational hot spots have been identified for EGFR, ERBB2 and ERBB3, unlike ERBB4, which has only been linked to ”mini-hotspots” in melanoma (Lau et al. 2014). Activating somatic mutations of the ERBBs are widely associated with non-small cell lung cancer (NSCLC) (Del Re et al. 2020), but there are also inactivating somatic mutations that can dramatically affect signaling in heterodimers (Soung et al. 2006, Tvorogov et al. 2009). Gain-of-function ERBB mutations have also been reported in other cancers such as breast, brain, skin and colorectal cancers (Arteaga and Engelman 2014). One way to gain a better understanding of the consequences of these genetic alterations that have potential to promote the development of cancer, i.e. kinase activating ERBB mutations, is to assess their structural and functional roles in conjunction with experimental studies. This thesis aimed to investigate the dynamic structural effects of four cancer-related activating ERBB mutations and to provide a mechanistic insight on how the mutations result in increased kinase activity. Classical molecular dynamics simulation (MDS) was utilized to examine the mutation-induced structural changes on the ERBB proteins.

2 Review of the literature

2.1 Discovery of the ERBB tyrosine kinases

The discovery of the ERBB family of receptor tyrosine kinases was ignited in 1962 when Stanley Cohen identified epidermal growth factor (EGF) in an effort to describe the factor inducing the new-born mice precocious eyelid opening and tooth eruption (Cohen 1962; Cohen 1965). In 1975, Carpenter and colleagues showed the presence of a membrane-bound receptor that binds EGF, now known as the epidermal growth factor receptor (EGFR) (Carpenter et al. 1975). Later in 1978, it was shown that binding of EGF resulted in an enhanced phosphorylation of EGFR (Carpenter et al. 1978). In 1980, Ushiro and Cohen revealed that EGFR is a tyrosine kinase receptor after identifying phospho-tyrosine residues in the isolated membrane protein that resulted from the binding of EGF (Ushiro and Cohen 1980).

In the years to follow the role of tyrosine phosphorylation in signal transduction was actively investigated. The identification of other tyrosine kinases such as the insulin receptor (INSR) and the platelet-derived growth factor receptor (PDGFR) further galvanized the study of the RTK family proteins, which includes 20 subfamilies and 58 known members (Robinson et al. 2000; Ségaliny et al. 2015). The *EGFR* cDNA was cloned and sequenced in 1984 (Ullrich et al. 1984). This study also revealed that the human A431 epidermoid carcinoma cells exhibit an aberrant amplification of *EGFR*, which linked *EGFR* with cancer. Indeed, the *EGFR* sequence had previously shown high similarity with the avian erythroblastosis retroviral oncogene, v-ERBB, that caused cancer in chickens infected by the virus (Yamamoto et al. 1983; Downward et al. 1984); indeed, numerous retorviruses were characterized that harbor the tyrosine kinase domain from a species (avian, feline, murine, human) that when infecting the species leads to uncontrolled tyrosine-kinase activity and the association with cancers (reviewed in Lipsick 2019).

The remaining ERBB family receptors, ERBB2, ERBB3 and ERBB4, were discovered in subsequent studies (Schechter et al. 1984; Kraus et al. 1989; Plowman et al. 1993).

2.2 Physiological role of the ERBB tyrosine kinases

The ERBBs are ubiquitously expressed in various cell types, such as epithelial, cardiac, mesenchymal and neuronal cells (Roskoski 2014b). The ERBBs are essential in the embryogenesis of vertebrates as they regulate the development of different organs (Britsch 2007). Null mutations in *ERBB* genes of mice have resulted in embryonic or perinatal lethality (Citri and Yarden 2006). Knockout of *ERBB* genes in mice have

shown varying phenotypes based on the targeted ERBB family member. For instance, EGFR null mice exhibited brain, liver, kidney and gastrointestinal tract abnormalities (Miettinen et al. 1995). Mice with compromised ERBB2 or ERBB4 function showed fatal cardiac problems related to the lack of myocardial trabeculae development, which is needed for proper contractions and blood flow (Gassmann et al. 1995; Lee et al. 1995). The importance of ERBBs in cardiac development and function is further substantiated by the cardiotoxicity side effects of cancer therapeutics that target the ERBBs, such as the monoclonal antibody Trastuzumab and tyrosine kinase inhibitors (TKIs) (Sanchez-Soria and Camenisch 2010).

EGFR and ERBB2 are reported to affect embryonic hair follicle development and skin maturation. Genetically engineered mouse models of EGFR and ERBB2 have e.g. exhibited delayed hair development, altered hair follicle morphology, wavey hair, thickened skin and alopecia (Schneider et al. 2008). ERBBs are also critical in the development of mammary glands (Hynes and Watson 2010), with EGFR, ERBB2 and ERBB3 having roles in ductal outgrowth. ERBB4 is important in lactation and alveolar differentiation. Furthermore, ERBBs are key for the development and regulation of the central nervous system, with mouse studies revealing their role in peripheral myelination, circuit assembly, synaptic plasticity and neuronal communication (Gassmann et al. 1995; Lee et al. 1995; Mei and Nave 2014). ERBB3 mutations in mice have led to neuropathies linked to abnormalities in Schwann cell development, cells that allow electrical insulation by wrapping around nerve axons (Riethmacher et al. 1997). Complicating matters, heterodimers are formed within the family, which will be discussed in section 2.3.5.

2.3 ERBB receptor structure and mechanism of activation

When Ullrich and colleagues first identified the EGFR amino acid sequence, they suggested that EGFR is composed of three domains, namely an extracellular domain, a transmembrane domain and a cytoplasmic domain with a protein kinase activity (Ullrich et al. 1984). Decades of experimental evidence has indeed confirmed their hypothesis. The four ERBB members share a common architecture; they contain a signal peptide (23 to 26 residues) that is cleaved to give rise to the mature protein, which consists of the growth factor binding ectodomain, the transmembrane (TM) domain, the juxtamembrane (JM) segment, the intracellular kinase domain and the C-terminal tail (Figure 2). Orchestrated conformational changes involving the different ERBB

domains and motifs – extracellular, TM and intracellular – regulate the activation of the proteins.

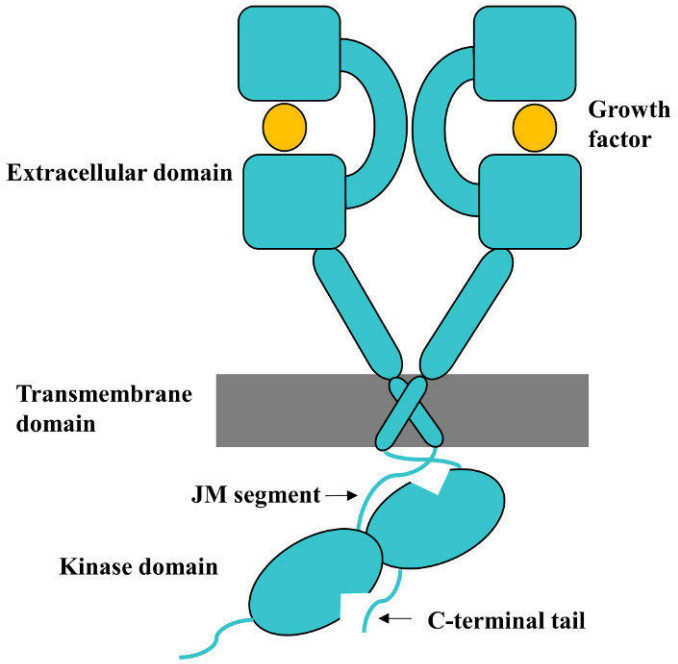


Figure 2. The ERBB receptor architecture. The different domains that make up the ERBB tyrosine kinases are highlighted. Figure adapted from publication I.

A plethora of X-ray and NMR structural data now exist for the ERBB receptors in the Protein Data Bank (PDB) (Berman et al. 2000), with EGFR having the highest share of experimentally resolved structures within the family. A wide array of crystal structures is available for the intracellular kinase domain, including those of the wild-type, some mutants, active, inactive, apo, various ligand-bound examples, and monomeric and dimeric forms. The first ERBB structures were deposited in the PDB by Stamos et al. and describe the active conformation of the apo (PDB ID: 1M14; 2.6 Å resolution (Stamos et al. 2002)) and erlotinib-bound (PDB ID: 1M17; 2.6 Å (Stamos et al. 2002)) EGFR kinase domains. The 2.6 Å resolution crystal structure of the ERBB3 ectodomain in the inactive state (PDB ID: 1M6B (Cho and Leahy 2002)) was resolved by Leahy, D.J., Cho, H.-S. around the same time.

2.3.1 The extracellular domain

The ERBB extracellular domain is glycosylated and is divided into four domains (I, II, III and VI). Domains I and III include several leucine residues and both have a β-helix LRR-like “solenoid” domain that is

involved in ligand binding. Domains II and IV are cysteine-rich, which form multiple disulphide bridges. Prior to growth factor binding, the ectodomains of EGFR, ERBB3 and ERBB4 attain an autoinhibited and "tethered" conformation (Figure 3A). In this state, intramolecular interactions between domains II and IV lock the hairpin-loop dimerization arm of domain II, and the ligand binding domains I and III are separated (Linggi and Carpenter 2006).

Upon growth factor binding, an extremely large conformational change occurs that results in an open and "extended" ectodomain conformation (Figure 3B). This structural rearrangement brings domains I and III together and releases the dimerization arm of domain II, which facilitates nearly symmetrical dimerization via monomer-monomer interactions (Dawson et al. 2007). Although the release of the dimerization arm of domain II is central to ERBB dimerization, other restraining intramolecular interactions are also critical. This notion is supported by mutagenesis studies that showed mutations that expose the dimerization arm of domain II increased ligand binding but lacked the ability to trigger receptor dimerization (Ferguson et al. 2003).

EGFR binds to seven growth factors, including EGF, EPG, TFGα, AR, BTC, HB-EGF and EPR. ERBB3 binds to NRG-1 and NRG-2 growth factors. Similar to EGFR, ERBB4 binds to seven growth factors, namely, NRG-1, NRG-2, NRG-3, NRG-4, BTC, HB-EGF and EPR (Fuller et al. 2008). Unlike the rest of the ERBBs, ERBB2 lacks growth factor ligands, however the unliganded form adopts a conformation in which the dimerization arm of domain II is exposed to the surface, making it dimerization competent (Cho et al. 2003; Garrett et al. 2003).

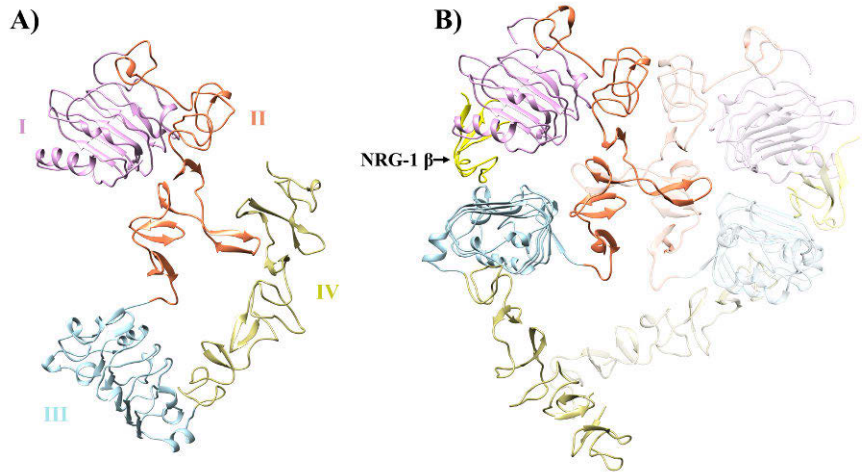


Figure 3. The ERBB4 extracellular domain. A) The ERBB4 ectodomain monomer structure in the inactive "tethered" conformation (PDB ID: 2AHX (Bouyain et al. 2005)). B) The ERBB4 ectodomain dimer in the active "extended"

conformation with bound NRG-1 β growth factor (yellow) (PDB ID: 3U7U (Liu et al. 2012)). Domains I to IV of the ectodomain are indicated.

2.3.2 The transmembrane domain and the cytoplasmic juxtamembrane segment

The ERBBs have a single-pass helical TM domain that is largely formed from hydrophobic residues as would be expected. The ERBB TM domain contains the GG4-like (glycine-X-X-X-glycine-like) motif, two small amino acids (glycine and/or alanine, serine and threonine) separated by three other residues, which is key for structural stability and dimerization. Two GG4-like motifs are present in the TM domains of EGFR, ERBB2 and ERBB4, whereas ERBB3 possesses only one GG4-like motif (MacKenzie 2006; Cymer and Schneider 2010). In the active state, the TM domains dimerize in such a way that the monomers are tilted with respect to each other and the TM domians cross each other at their N-terminus, and the C-terminal ends are splayed apart (Figure 4). In the inactive ERBB state, this helical packing takes place at the C-terminus, aiding inhibition of receptor activity (Lu et al. 2012; Bocharov et al. 2016). Multiple solution NMR structures of the TM domains of ERBBs have been deposited in the PDB.

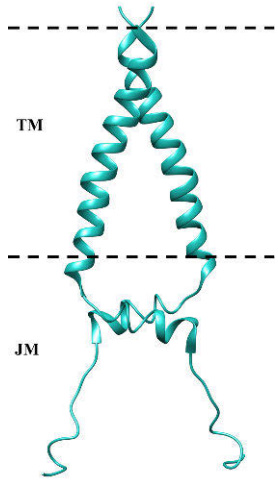


Figure 4. The EGFR TM domain and JM segment (PDB ID: 2M20). The extracellular side of the membrane is located at the top of the figure.

The cytoplasmic JM segment, which is largely composed of loops and is hence flexible, extends immediately from the TM domain on the cytoplasmic side of the membrane (Figure 4). The JM segment is helical at its N-terminus, which dimerizes in an antiparallel fashion during receptor activation (Jura et al. 2009a; Cymer and Schneider 2010). The C-terminal end of the segment, the JM latch, plays a critical role in the asymmetric

dimerization of the intracellular kinase domains by forming part of the dimerization interface and allowing for flexibility leading up to the tyrosine kinase domain and supporting conformational changes (Jura et al. 2009a). NMR structures of TM domain-intact JM segments of EGFR (PDB ID: 2M20 (Endres et al. 2013)) and ERBB2 (2N2A (Bragin et al. 2016)) are available in the PDB.

2.3.3 The intracellular kinase domain

The ultimate consequence of growth factor binding to the ERBB ectodomain is activation of the enzymatic "business-end" of the receptors, the cytoplasmic tyrosine kinase domain, via asymmetric dimerization and contrary to the typical means of protein kinase activation via phosphorylation (Zhang et al. 2006; Beenstock et al. 2016). The asymmetric ERBB kinase dimers mirror the CDK2-cyclin A heterodimer, where one of the kinase monomers ("activator" - as cyclin A) activates the other monomer ("receiver" - as CDK2) (Zhang et al. 2006; Qiu et al. 2008) (Figure 5A).

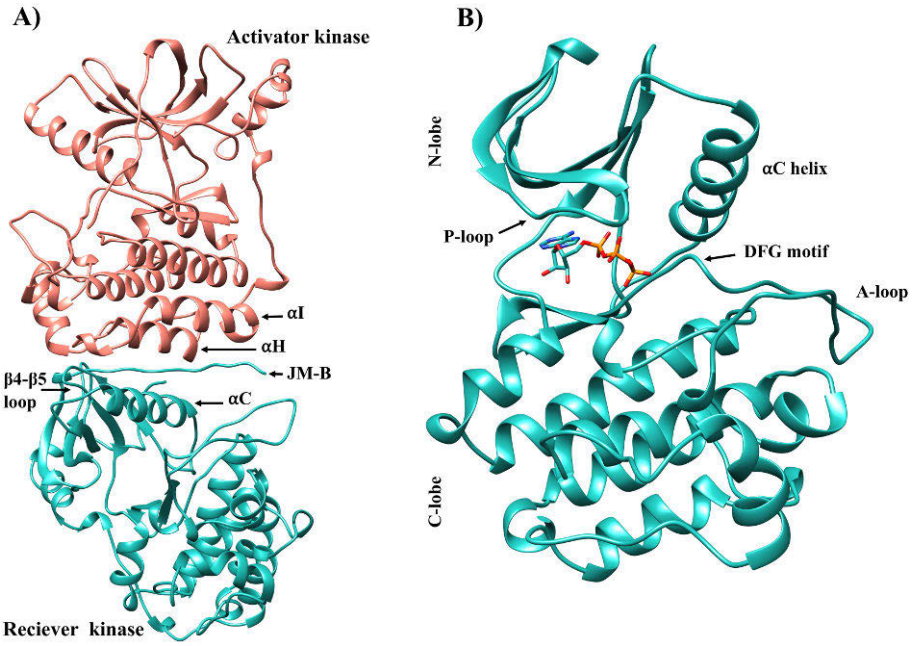


Figure 5. The EGFR kinase domain (PDB IDs: 2GS2 (Zhang et al. 2006), 2ITX (Yun et al. 2007)). A) The asymmetric dimer of the EGFR kinase domain highlighting the structural elements forming the dimer interface B) The ATP-bound EGFR kinase domain showing the locations of key structural units. Figure adapted from publication II.

The structure of the kinase domain is highly conserved among protein kinases: the kinase domain is divided into two lobes, the N- and C-lobes, with the former predominantly composed of β-strands and the latter largely formed of helices (Figure 5B). The ATP binding pocket in the kinase domain lies between the two lobes that are connected by a flexible hinge region. The smaller N-lobe includes key structural elements such as the P-loop that shields the binding pocket and the αC helix that governs the transition between active and inactive states of the kinase domain. The larger C-lobe on the other hand is the residence for the activation loop (A) and catalytic loop that are important for the activity of the kinase domain (Huse and Kuriyan 2002).

The ERBB kinase domain attains at least two distinct conformational states: the active and inactive states (Figure 6A). In the active kinase state, the αC helix is placed near the ATP binding pocket adopting the "αC-in" conformation, which allows the formation of the functionally critical and conserved ion-pair between Glu762 of the αC helix and Lys745 of the β3 strand (numbering based on EGFR). Furthermore, the activation loop that contains the aspartate-phenylalanine-glycine ("DFG") motif attains an extended and open conformation. The aspartate of the DFG motif is oriented towards the binding pocket, attaining the "DFG-in" state, where it is involved in catalysis by coordinating a magnesium ion. The phenylalanine of the DFG motif is embedded near the αC helix with the sidechain facing towards the binding pocket (Roskoski 2014a, 2014b).

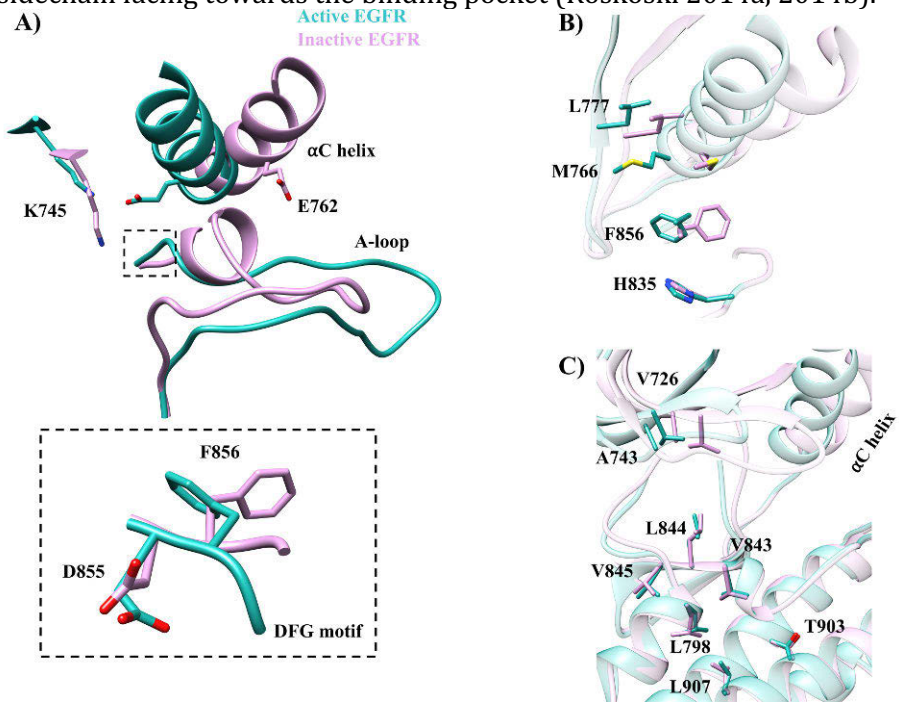


Figure 6. Superimposed structures of the active and inactive EGFR kinase domains (PDB IDs: 2GS2, 2GS7 (Zhang et al. 2006)). The αC helix, A-loop, DFG motif, Lys745-Glu762 salt bridge A), R-spine B) and C-spine C) structural elements in the active and inactive EGFR states. Figure adapted from publication II.

In the inactive kinase state (Figure 6A), the αC helix moves away from the binding pocket and assumes the "αC-out" conformation, which consequently breaks the Lys745-Glu762 salt bridge. Moreover, the A-loop holds a closed conformation with a small helical turn being formed at its N-terminus. Following the positional change of the αC helix, the phenylalanine of the DFG motif also reorients with the sidechain moving towards the αC helix. This conformation of the inactive state resembles the conformation seen in the crystal structure of the Src kinase, hence this conformation is called the Src-like inactive state (Jura et al. 20011; Zhang et al. 2006; Qiu et al. 2008). Although the ERBBs largely adopt the Src-like form of the inactive conformation, EGFR has been shown to exist in another inactive conformation known as the "DFG-out" inactive state (Gajiwala et al. 2013). In the "DFG-out" conformation the αC helix moves away from the binding pocket in a similar fashion to the Src-like inactive conformation, however, the A-loop is in the extended conformation as in the active state but the sidechains of aspartate and phenylalanine of the DFG motif flip "sides", abolishing the catalytic role of the aspartate.

A skeleton of non-consecutive hydrophobic residues from both the N- and C-lobes of the kinase domain make up two key structural units known as the regulatory (R) and catalytic (C) spines (Kornev et al. 2006; Taylor and Kornev 2011). The R-spine is composed of four hydrophobic residues located at the αC helix, β4 strand, A-loop and catalytic loop (Figure 6B). The linear arrangement of these residues is crucial for maintaining the active kinase conformation (Kornev et al. 2006; Roskoski 2014b). In the inactive kinase conformation, the spatial assembly of these residues is disrupted, with the residue from the αC helix (Met766/Met774/Met772 in EGFR/ERBB2/ERBB4) and A-loop (Phe856/Phe864/Phe862 in EGFR/ERBB2/ERBB4) changing their orientation due to the "αC-out" conformation of the αC helix. The C-spine of the kinase domain constitutes eight amino acids from the β2 strand, β3 strand, β7 strand, αD helix and αF helix (Figure 6C). The C-spine affects catalysis by mediating binding of ATP (Taylor and Kornev 2011; Roskoski 2014b).

2.3.4 The cytoplasmic C-terminal tail

The C-terminal tail of the ERBBs accommodates (Figure 7) multiple tyrosine residues that undergo autophosphorylation upon tyrosine kinase domain activation. The phosphotyrosines serve as docking sites for

molecules that initiate the downstream signaling of ERBBs (Leahy 2004; Fuller et al. 2008). Only the first 40 residues that follow the kinase domain are often present in resolved ERBB crystal structures. Most of these structures show a discontinuous C-terminal tail, which implies a high structural flexibility and intrinsic disorder for the region. An EGFR structure in the inactive state (PDB ID: 3W2S (Sogabe et al. 2013)), however, shows an uninterrupted C-terminal tail for the proximal segment with an intact AP2 helix. Another inactive EGFR structure (PDB ID: 3GT8 (Jura et al. 2009a)) shows the AP2 helix in a slighly different orientation packing against the N-lobe of the dimeric partner kinase domain. The N-terminal part of the EGFR tail has been described as having a role in receptor autoinhibition (Gajiwala 2013). This is supported by two oncogenic mutations where the proximal part of the tail is deleted, resulting in constitutive dimerization and increased kinase activity (Pines et al. 2010; Kovacs et al. 2015).

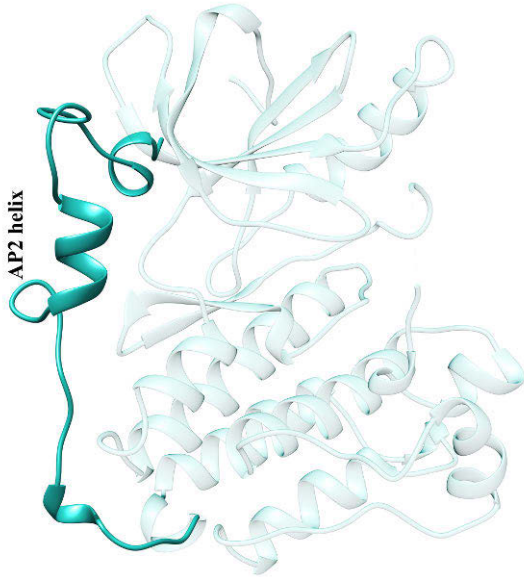


Figure 7. The C-terminal tail of EGFR (solid color; PDB ID:3W2S).

2.3.5 ERBB receptor dimerization

As mentioned, activation of the ERBB receptors is induced via dimerization of ERBB monomers. Binding of growth factor (with the notable exception of ERBB2) to the ectodomain triggers dimerization of both the ectodomain and the TM domain, propagating to the intracellular kinase domain, where two kinase domain monomers (activator kinase and receiver kinase) associate in an asymmetric manner prompting allosteric receptor activation. ERBB2 does not bind a growth factor but dimer formation can occur through heterodimer formation with a growth-

factor binding family member, leading to activation of multiple signaling pathways.

Interactions between residues of the C-lobe of the activator kinase and residues of the N-lobe of the receiver kinase are fundamental for the kinase domain dimerization (Zhang et al. 2006). In particular, the amino acids at the αI and αH helices of the activator kinase and the residues of the JM-B segment, β4-β5 loop and αC helix of the receiver kinase play a critical role (Figure 5A). The EGFR kinase domain in the inactive state has been resolved with two monomers complexed in a symmetric fashion, with the N-lobe of one monomer and the C-terminal tail of the other monomer making up the interaction interface (Zhang et al. 2006; Jura et al. 2009a).

The dimeric ERBB receptor complexes can be homodimers or heterodimers, with 28 possible combinations. Consideration of the 11 EGF-like ERBB ligands raises the potential combinations to 614 (Roskoski 2014b). EGFR and ERBB4 can form functional homodimers, unlike ERBB2 and ERBB3 (Leahy 2004). ERBB2 forms an active homodimer only when overexpressed (Harari and Yarden 2000), whereas ERBB3 has an impaired kinase activity, having EGFR's $1/1000^{th}$ activity, due to mutation of residues key for kinase domain activity (Leahy 2004; Jura et al. 2009b; Shi et al. 2010). Hence, the ERBB3-ERBB3 homodimer is nearly non-functional. ERBB2 is the choice of heterodimer partner among the ERBBs, with the ERBB2-ERBB3 complex exhibiting the strongest proliferative signal (Citri et al. 2003). While EGFR prefers to serve as the receiver kinase when paired with the rest of the ERBBs, ERBB2 takes the place of the receiver in the ERBB2-ERBB2, ERBB2-ERBB3 and ERBB2-ERBB4 dimers. ERBB4 functions as the receiver kinase only when complexed with itself or with ERBB3 (Ward and Leahy 2015).

2.4 The ERBB signaling pathways

The ERBB receptor kinases trigger a network of downstream signaling pathways following autophosphorylation at their C-terminal tails. The ERBBs recruit signal transducers, known as adaptor proteins, that possess the protein tyrosine binding (PTB) domain and/or Src homology 2 (SH2) domain, which bind to phosphotyrosines along the C-terminal tails of the receptors (Citri and Yarden 2006). The signaling molecules can also be recruited indirectly by binding to other tyrosine kinase-phosphorylated proteins including Gab1 (Grb2-associated binder) and IRS1 (insulin receptor substrate 1), which serve as their docking sites (Schlessinger 2000). These protein-protein interactions are integral to coordinating ERBB kinase signaling, which originates at the cell surface on binding a growth factor and propagates to the nucleus. Multiple signaling pathways are associated with ERBB kinases, which include the RAS-RAF-MEK-

ERK/MAPK, phospholipase C-gamma1 (PLC-γ1), phosphatidylinostiol 3-kinase (PI3K-AKT), Janus kinase-signal transducer and activator of transcription (JAK-STAT) and c-Jun N-terminal kinase (JNK) pathways (Yarden and Sliwkowski 2001).

The RAS-RAF-MEK-ERK/MAPK pathway promotes cell survival, proliferation, differentiation, migration and inhibition of apoptosis (Guo et al. 2020). ERBB-adaptor protein activated RAS interacts with and activates RAF. RAF in turn phosphorylates and activates MEK, which subsequently activates ERK. ERK then moves from the cytoplasm to the nucleus, where it phosphorylates different transcription factors, activating various genes that regulate the cell cycle (Wee and Wang 2017). Activating mutations occurring on proteins of this pathway are associated with the development of cancer. Consequently, the RAS-RAF-MEK-ERK/MAPK pathway is often targeted as a therapeutic intervention (Degirmenci et al. 2020).

The PLC-γ1 pathway exerts its signaling by direct binding of PLC-γ1 to C-terminal phosphotyrosines of EGFR (Wee and Wang 2017). Phosphorylated PLC-γ1 then translocates from the cytoplasm to the plasma membrane where it hydrolyzes phosphatidylinositol 4,5-bisphosphate (PIP2) to inositol 1,4,5-trisphosphate (IP3). This in turn triggers intracellular calcium release and consequent activation of protein kinase C that stimulates gene expression. Phosphorylated PLC-γ1 can also activate RAS that regulates cell growth.

The PI3K-AKT pathway is important in cell growth, metabolism, migration and vesicular trafficking. PI3K binds directly to ERBB3 and ERBB4 via its SH2 domain of the p85 subunit. Binding to EGFR and ERBB2 occurs indirectly via the GAB1 adaptor protein (Wee and Wang 2017). Following PI3K activation, phospholipid products of PI3K, such as PIP3, induce the translocation of AKT to the plasma membrane by binding to its pleckstrin homology (PH) domain. In the plasma membrane the kinase domain of AKT is phosphorylated and activated by phosphoinositide-dependent kinase-1 (PDK-1), which translocates from the cytoplasm to co-localize in the plasma membrane with AKT. Subsequently, AKT induces various biological effects, including cell survival, growth, motility and proliferation (Arcaro and Guerreiro 2007; Ferreira and Pessoa 2017).

The JAK-STAT pathway consists of four JAK proteins and seven members of the STAT proteins that are involved in inter-pathway talks. JAKs are phosphorylated after binding to ERBB C-terminus. Activated JAKs then phosphorylate STATs resulting in the formation of STAT homo- and heterodimers. STAT dimers subsequently translocate to the nucleus where they regulate the expression of genes that affect cell proliferation, migration and survival (Ferreira and Pessoa 2017; Bousoik and Montazeri Aliabadi 2018).

In addition to the well-established cell membrane based signaling, the ERBBs also act by translocating to the nucleus with the aid of the importin-A/B nuclear transport proteins (Giri et al. 2005; Lo et al. 2006). Nuclear ERBBs regulate the cell cycle, DNA replication, DNA damage repair and transcription (Wang and Hung 2009). Nuclear ERBBs are known to enhance expression of several genes that are critical in cancer biology. Both membrane and nuclear-based ERBB signaling are intricate and encompass multiple molecules and inter-linked pathways that elicit a wide range of cellular functions.

2.5 Role of ERBB receptor kinases in disease

The ERBBs are essential for the regulation of a diverse range of biological processes. Deviation from their optimal activity however associates them with a variety of human diseases. For instance, the ERBBs have been reported to play a role in the pathogenesis of neuropsychiatric diseases such as schizophrenia, bipolar disorder and major depression (Mei and Nave 2014). Anomalous activation of EGFR has been associated with the development of multiple dermatological diseases including psoriasis, atopic dermatitis and wound healing defects (Wang et al. 2019). Additionally, as membrane receptors, the ERBBs serve as key molecules for pathogens to enter host cell, and hence are linked to several infectious diseases caused by *Staphylococcus aureus*, *Neisseria gonorrhoeae*, Epstein–Barr virus, *Helicobacter pylori* and hepatitis C virus, among others (Ho et al. 2017).

The most common disease related to the ERBBs is cancer, which may not be surprising given the receptors' prominent role in pro-oncogenic cellular processes such as cell proliferation, differentiation, survival, motility and inhibition of apoptosis. Aberrant signaling of ERBB receptors, due to gene amplification, overexpression or activating mutations, is associated with several human cancers (Ding et al. 2008; Roskoski 2014b; Mishra et al. 2017; Del Re et al. 2020). Activating somatic mutations of the ERBBs have widely been observed in patients with non-small cell lung cancer (NSCLC) (Mishra et al. 2017), which accounts for 80% of all lung cancer. In 2004, three insightful studies showed the role of EGFR in tumorigenesis by reporting numerous EGFR mutations observed in NSCLC patients, and their selective response to TKIs (Lynch et al. 2004; Paez et al. 2004; Pao et al. 2004). Amplification of the *ERBB2* gene and/or mutations at the kinase domain of ERBB2, notably the exon 20 insertion mutations, are also oncogenic drivers in NSCLC (Singh et al. 2020). ERBB3 is linked to lung adenocarcinoma, a subtype of NSCLC, via induction of ERBB2-ERBB3 heterodimerization, which activates pathways involved in tumor progression. Additionally, somatic mutations in ERBB3 kinase domain have been observed in NSCLC patients (Del Re et al. 2020).

Similarly, ERBB4 somatic mutations have been identified in lung adenocarcinoma samples (Ding et al. 2008).

Other types of cancers frequently associated with the ERBB family proteins include breast, skin, gastric, pancreatic and colorectal cancers. ERBB2 overexpression and amplification are observed in 20-30% of breast cancers (Slamon et al. 1987; Yu and Hung 2000). ERBB4 mutations are commonly recorded in patients with melanoma (Prickett et al. 2009). Amplified or mutated EGFR is implicated in half of glioblastomas (Clark et al. 2012). The oncogenic role of the ERBB receptors consequently makes them ideal therapeutic targets in cancer treatment.

2.6 Activating somatic mutations of the ERBB receptors

Thousands of somatic mutations occurring in the ERBB family proteins have been reported in cancer tissues. These mutations can be 'drivers' that contribute to cancer development or 'passengers' that lack proliferative roles (Vogelstein et al. 2013). In general, the majority of somatic mutations observed in cancer cells are passenger mutations (Pon and Marra 2015) and only a portion of the ERBB somatic mutations have been functionally characterized. Mutations that increase ERBB kinase activity have high clinical relevance as they drive tumor progression. Although less common, inactivating ERBB mutations have also been observed in cancer samples. For instance, the cancer-associated G802dup and D861Y ERBB4 kinase domain mutations (Soung et al. 2006) were shown to be kinase-dead by interfering with ATP binding (Tvorogov et al. 2009). However, selective ERBB4 signaling was still possible by the mutants via heterodimeric activation of ERBB2.

Activating somatic mutations of the ERBBs constitute point mutations, short in-frame insertions, deletions and truncation of large segments. These mutations reside in different parts of the proteins and affect receptor signaling. Activating mutations occurring on the extracellular domain have been reported to alter ligand binding or affect the conformational change between the active and inactive states of the ectodomain (Riese et al. 2007). Among those affecting ligand binding is the EGFR S442F mutation that is located at the ligand binding domain III of the ectodomain. S442F stimulates EGFR phosphorylation by increasing the binding affinity of EGFR to NRG2B, which normally is a weak binder (Gilmore et al. 2006). Another activating mutation of the extracellular domain includes the EGFR variant III (EGFRvIII) mutation, which has similarities to the avian erythroblastosis virus oncogene – v-ERBB, is common in human glioblastomas. EGFRvIII results in a truncated form of the receptor and affects the shift in conformational equilibrium (Pedersen

et al. 2001). The Y285C (changing the position of a disulfide bond) and D595V (making the dimer arm interactions more hydrophobic) ERBB4 activating mutations have been identified in lung adenocarcinoma samples (Ding et al. 2008) and were demonstrated to enhance receptor dimerization, resulting in a longer life of the activated dimer (Kurppa et al. 2016). Several activating mutations also occur on the ERBB3 ectodomain, with domain II being the hotspot (Jaiswal et al. 2013).

Studies have described activating ERBB mutations in the JM region and the TM domain (Burke and Stern 1998; Penington et al. 2002). Individual replacement of various JM residues to cysteine in ERBB2 and ERBB4 has resulted in ligand-independent homodimerization and phosphorylation, likely due to the formation of intermolecular disulfide bridges between the substituted cysteine residues (Burke and Stern 1998). The ERBB2 V664E and I655E mutations in the transmembrane domain result in increased receptor activation, with the latter mutation implicated in breast cancer (Cymer and Schneider 2010).

The ERBB intracellular kinase domain accommodates an array of activating mutations. Most of these mutations occur near the ATP binding pocket of the domain. Gain-of-function mutations of the EGFR kinase domain have been identified in 10-30% of NSCLC patients (Shigematsu and Gazdar 2006; Collisson et al. 2014). These mutations include the exon 21 L858R mutation, a number of exon 19 in-frame deletions and exon 20 in-frame insertions (Figure 8). Analogous mutations are present in ERBB2, with the in-frame insertions being the most prevalent (Herter-Sprie et al. 2013; Wen et al. 2015). Activating mutations in the ERBB4 kinase domain have been observed in melanoma and lung adenocarcinoma samples, which include E836K, E872K, D931Y and K935I (Ding et al. 2008; Prickett et al. 2009; Kurppa et al. 2016). Activating mutations of the kinase-dead ERBB3 domain are relatively rare but a few mutations have been observed at the C-lobe that may activate the partner kinase domain, and which are implicated in lung and breast cancers (Jaiswal et al. 2013).

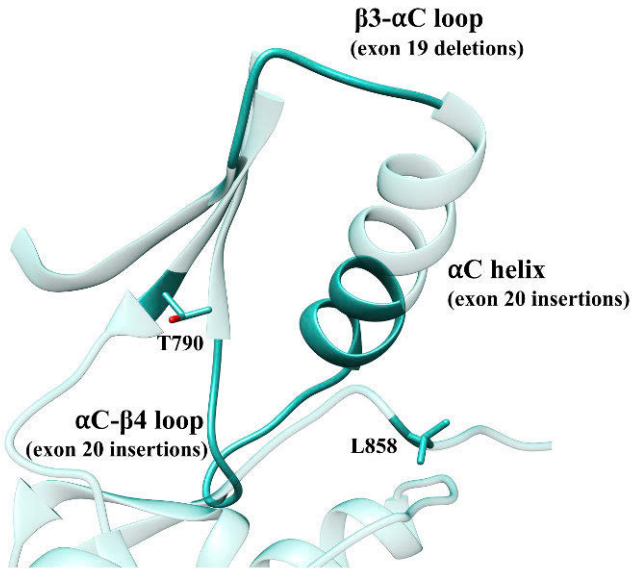


Figure 8. Locations of the common EGFR mutations in NSCLC. The positions of the L858R, T790M, exon 19 deletion and exon 20 insertion mutations in the EGFR kinase domain are shown.

The EGFR L858R and exon 19 deletion kinase domain mutations, also known as classical mutations, account for nearly 90% of all EGFR mutations observed in NSCLC (Lynch et al. 2004; Paez et al. 2004; Murray et al. 2008; Gazdar 2009). L858R is suggested to exert its activating effect by promoting a conformational change from the inactive towards the active kinase state by eliminating a hydrophobic interaction at the A-loop, which is key to maintaining a helical turn present only in the inactive conformation (Zhang et al. 2006). Exon 19 deletions comprise various in-frame deletions that differ in the number and type of amino acids deleted, among which the 746ELREA750 deletion is the most common (Gazdar 2009; Foster et al. 2016; Su et al. 2017). Exon 19 deletions are located at the β3-αC loop adjacent to the αC helix that plays a critical role in regulating kinase domain activation (Foster et al. 2016). A few cases of similar ERBB2 in-frame deletions have been observed in breast cancer (Lee et al. 2006; Shah et al. 2009; Ellis et al. 2012; Bose et al. 2013). NSCLC patients with the EGFR L858R mutation and exon 19 deletion mutations are treated with first-generation TKIs, which inevitably develop resistance due to the acquired T790M mutation (Pao et al. 2005). T790M, a "gate-keeper" mutation, is located at the ATP binding pocket of the EGFR kinase domain (Figure 8). Replacement of the smaller threonine to the bulkier methionine results in a steric hindrance against TKIs, hence the observed drug resistance (Eck and Yun 2010). An increased affinity to ATP is suggested as an additional mechanism of resistance resulting from the

mutation. Similarly, the gate-keeper mutation, T798I, located at the equivalent location in ERBB2, was observed to induce resistance to the EGFR/ERBB2 TKI neratinib in a breast cancer patient (Hanker et al. 2017).

In NSCLC, exon 20 insertion mutations make up around 10-12% of mutations in EGFR (Yasuda et al. 2012; Arcila et al. 2013; Oxnard et al. 2013) and 90% of the ERBB2 mutations (Arcila et al. 2012; Mazières et al. 2013). Collectively, EGFR and ERBB2 exon 20 insertions are observed in 4% of NSCLC patients (Arcila et al. 2012). Exon 20 insertion mutations exist in two regions of the kinase domain, the αC helix and the succeeding αC-β4 loop. The mutations are suggested to affect kinase activity by preventing a transition from the active kinase conformation towards the inactive kinase conformation (Yasuda et al. 2013). Contrary to the classical EGFR mutations, the exon 20 insertion in EGFR and ERBB2 are insensitive to first- and second-generation TKIs (Yasuda et al. 2013; Robichaux et al. 2018).

2.7 ERBB-targeted therapeutics

2.7.1 Monoclonal antibodies

Understanding the intricate mechanism of action of the ERBB receptors has paved the way for the development of a new generation of biological agents employed in cancer treatment. One class of agents targeting the ERBBs are monoclonal antibodies (mABs). The first clinically used mAB is trastuzumab that is commonly used in the treatment of ERBB2-amplified breast cancer (Molina et al. 2001; Arteaga and Engelman 2014). Trastuzumab binds near domain IV of the ectodomain and results in the inhibition of ectodomain cleavage and uncoupling of ERBB2 dimers (Molina et al. 2001; Ghosh et al. 2011; Arteaga and Engelman 2014). Pertuzumab was the second mAB developed that targets ERBB2 (Agus et al. 2002). Pertuzumab binds to the dimerization interface of the ectodomain and acts by interfering with receptor dimerization, which leads to inhibition of signaling (Franklin et al. 2004). Other developed mABs include cetuximab and panitumumab, both of which bind to the ectodomain of EGFR (Goldstein et al. 1995; Yang et al. 2001). These two mABs are used clinically in the treatment of colorectal cancer (Van Cutsem et al. 2007, 2009).

2.7.2 Tyrosine kinase inhibitors

Another class of drugs widely employed in the treatment of ERBB-related cancers is the small molecule TKIs. The critical role of the intracellular tyrosine kinase domain in ERBB signaling makes it an ideal target for inhibition of receptor activity. The high sequence and structure

conservation of the ERBB kinase domain with other protein kinases however poses a challenge to develop selective ERBB kinase inhibitors (Wieduwilt and Moasser 2008). Currently there are three groups of TKIs that target the ERBBs, namely, first-, second- and third-generation TKIs.

2.7.2.1 First-generation TKIs

The first-generation TKIs, which include gefitinib, erlotinib, lapatinib and icotinib, target the EGFR kinase domain (Fry et al. 1994; Ward et al. 1994; Johnston and Leary 2006). Lapatinib is a dual EGFR/ERBB2 inhibitor. These molecules are quinazoline-based reversible inhibitors that bind to the ATP binding site of the EGFR kinase domain (and ERBB2 for lapatinib). Crystal structures of gefitinib and erlotinib-bound EGFR are available in the PDB both in the active (PDB IDs 4WKQ, 1M17 (Stamos et al. 2002)) and inactive (PDB IDs 4I22 (Gajiwala et al. 2013), 4HJO (Park et al. 2012)) kinase conformations. A lapatinib-bound structure of EGFR in the inactive state has also been resolved (PDB ID 1XKK (Wood et al. 2004)). Gefitinib is a first-line treatment for NSCLC harboring the EGFR L858R and exon19 deletion mutations, whereas erlotinib is indicated as a first- and subsequent-line therapy for EGFR mutation-positive NSCLC (Gridelli et al. 2010; Reguart et al. 2010). Patients however often develop resistance against these agents, principally due to the T790M "gatekeeper" mutation near the ATP binding pocket (Pao et al. 2005).

2.7.2.2 Second-generation TKIs

In order to circumvent the acquired resistance limitation of the first-generation small molecules, second-generation TKIs were developed. The second-generation TKIs include afatinib, dacomitinib and neratinib that have a pan-ERBB activity (Engelman et al. 2007; Solca et al. 2012). These molecules are ATP-competitive and irreversibly bind to the ATP binding pocket by forming a covalent interaction with Cys797 (EGFR numbering). Afatinib is a first-line therapy for NSCLC patients with EGFR sensitizing (exon 19 deletions, L858R) and resistant mutations (T790M). Experimental assays have revealed second-generation TKIs to be more potent than first generation TKIs in cancer cell lines (Engelman et al. 2007; Li et al. 2008). Furthermore, two clinical trials have revealed afatinib to have a higher response rate than gefitinib (Park et al. 2016; Wu et al. 2017). Nonetheless, the second-generation TKIs have a more pronounced adverse effect as compared to the first generation TKIs, likely due to their irreversible and broad range inhibition, posing a hurdle in their effective use. Furthermore, and similar to the first generation TKIs, patients treated with these drugs eventually develop resistance (Tanaka et al. 2017; Wang and Li 2019).

2.7.2.3 Third-generation TKIs

The third-generation TKIs include osimertinib, olmutinib and rociletinib, with osimertinib being the only clinically approved drug. Osimertinib is indicated in both EGFR sensitizing and T790M resistance mutations in NSCLC (Tan et al. 2018). Osimertinib is an irreversible inhibitor that covalently interacts with Cys797 at the EGFR binding pocket of the tyrosine kinase domain. Preclinical data shows that osimertinib is more potent in inhibiting mutant EGFR cell lines (L858R/T790M) compared to the wild-type EGFR (Cross et al. 2014). Clinical studies have also demonstrated the efficacy of osimertinib in mutation-positive NSCLC patients (Goss et al. 2016). The impressive outcome of osimertinib is, however, challenged by the C797S resistant mutation, which alters the C797 residue that is critical for osimertinib binding (Niederst et al. 2015).

3 Aims of the study

The objective of this thesis was to probe the dynamic structural changes that take place on the ERBB proteins due to several activating somatic mutations, the majority of which are linked to NSCLC. The mutations of interest include:

1. 746ΔELREA750 (ΔELREA) – an EGFR exon 19 deletion mutation (I, IV)

2. V769insASV and D770insNPG – EGFR exon 20 insertion mutations (II)

3. E936K – an ERBB2 point mutation (III)

The study aimed to investigate the impact of these mutations on local structure, domain conformation, and interactions key to the state of the ERBB kinase activity. The study further sought to provide a mechanistic explanation on how the mutations increase ERBB kinase activity from a structural point of view.

4 Materials and methods

The materials and methods described here are thoroughly discussed in publications I-III. Furthermore, publication IV exhaustively dissects the methodology used to assess the structural consequences of the activating ERBB mutations. Publication IV additionally has a complimentary video (https://www.jove.com/v/61125/deciphering-structural-effects-activating-egfr-somatic-mutations-with) from which readers can visualize the procedures taken.

4.1 Structure and sequence retrieval

Three-dimensional (3D) X-ray structures of the ERBB proteins were accessed from the PDB (Berman et al. 2000). The sequences of the proteins were obtained from the Uniprot database (Bateman 2019). The sequences were particularly necessary when modeling an ERBB protein for which a resolved experimental structure did not exist.

Table 1. Principal ERBB structures accessed from the PDB.

Protein (domain)	State	PDB ID	Resolution (Å)
EGFR (kinase domain)	Apo, active	2GS2 (Zhang et al. 2006)	2.80
EGFR (kinase domain)	ANP-bound, active	2ITX (Yun et al. 2007)	2.98
EGFR (kinase domain)	Apo, inactive	2GS7 (Zhang et al. 2006)	2.60
ERBB2 (kinase domain)	Apo, active	3PP0 (Aertgeerts et al. 2011)	2.25
EGFR D770insNPG (kinase domain)	PD168393-bound, active	4LRM (Yasuda et al. 2013)	3.53
ERBB4 (kinase domain)	Apo, active	3BCE (Qiu et al. 2008)	2.50

4.2 Homology modeling and structure preparation

For the EGFR ΔELREA deletion mutation study, the monomeric form of the wild-type EGFR kinase domain in the apo-active (PDB ID: 2GS2), apo-inactive (PDB ID: 2GS7) and ANP-bound active (PDB ID: 2ITX) states were initially accessed from the PDB. The ANP ligand in 2ITX was converted to ATP by changing the N3 atom to O3. The ΔELREA deletion mutant forms were respectively modeled based on the prepared wild-type EGFR structures using Modeller (Šali and Blundell 1993).

For the EGFR V769insASV and D770insNPG insertion mutation study both the active state wild-type (PDB ID: 2GS2) and D770insNPG (PDB ID: 4LRM) EGFR kinase structures, and the inactive state wild-type (PDB ID: 2GS7) EGFR kinase domain structure were obtained from the PDB. The active state V769insASV kinase domain structure was modeled with the Modeller program using the wild-type EGFR structure as a template. The inactive states of V769insASV and D770insNPG EGFRs were similarly modeled based on the inactive wild-type EGFR kinase domain structure (PDB ID: 2GS7).

The analysis of the ERBB2 E936K mutation was initiated from the wild-type ERBB2-ERBB2 asymmetric homodimer structure (PDB ID: 3PP0). During this study the ERBB2-EGFR and ERBB2-ERBB4 heterodimers were also probed. These two heterodimer structures were built based on the EGFR-EGFR and ERBB4-ERBB4 homodimeric structures, which were generated from their respective monomeric structures (PDB ID: 2GS2 and 3BCE) by applying symmetry operations in the Chimera program (Pettersen et al. 2004). The ERBB2 activator kinase structure was then superimposed on the activator kinase of the EGFR-EGFR and ERBB4-ERBB4 dimers. The EGFR and ERBB4 activator kinases were then replaced by ERBB2, giving rise to the ERBB2-EGFR and ERBB2-ERBB4 heterodimers. The E936K ERBB2 mutant forms of the ERBB2-ERBB2, ERBB2-EGFR and ERBB2-ERBB4 dimers were built by replacing Glu936 in the ERBB2 kinase domain with lysine.

Missing loops in the above wild-type and mutant ERBB structures were built either from other ERBB structures that contain these segments or modelled with Modeller. These structures were then prepared using the protein preparation wizard in Maestro (Schrödinger Release 2018-3: Maestro 2018). The preparation process involved addition of hydrogen atoms, removal of unwanted water and complexed molecules, determination of protonation states for ionizable side chains at pH 7.0, optimization of hydrogen bonds and energy minimization of the structures.

4.3 Molecular dynamics simulation (MDS)

MDS is a central computational technique in this research because it is one of the few ways in which the time-dependent dynamic motions of biological systems, such as proteins, nucleic acids and small molecules, can be examined. Structural data from experimental methods, such as X-ray crystallography, are indispensable, however, such methods provide a snapshot of a conformation and may fall short in capturing the dynamic nature of the molecules. The combination of the experimental structures with MDS provides a powerful means to examine the dynamic structural properties of biological systems.

The basic principle behind MDS is molecular mechanics (MM), in which molecular systems are modeled using potential energy functions (Leach 2001). In MM, each atom in a system is defined solely based on its nuclear position, excluding information regarding electrons, which are considered in the rather computationally expensive quantum mechanics. During MDS, atoms in a system interact with each other. Their interaction is coupled with movements, which are computed using Newton's equation of motion (1.1) (Leach 2001).

$$F = ma \quad 1.1$$

$$F = -\nabla U \quad 1.2$$

F - force, m - mass, a - acceleration, ∇U - gradient of potential energy

The force exerted on an atom during MDS can be calculated from the gradient of the potential energy (1.2). The potential energy is the sum of energies of bonded and non-bonded interactions taking place between atoms in a system. Once the total force to be exerted on an atom at a particular time is known, the acceleration of the atom can be derived using Newton's equation of motion (1.1). From this calculation the new position and velocity of the atom can be determined. Repetition of this process for a given number of cycles using a short time step provides the subsequent velocities and positions of atoms. The ensemble of the recorded positions for all the atoms in a system with changing time results in a trajectory, which describes the time evolution of the system.

In this thesis, the wild-type and mutant ERBB kinase domain structures were simulated in order to examine the atomic-level dynamic motions taking place in the proteins, which aids in pinpointing possible structural changes resulting from the activating mutations and suggesting how those changes led to the observed enzymatic activation.

4.3.1 Simulation system setup

Simulation systems were built for the prepared wild-type and mutant ERBB structures with the program leap in AMBER (Case et al. 2018). The AMBER ff14SB force field (Maier et al. 2015) and parameters for ATP (Meagher et al. 2003) were used. The proteins and protein-ATP complexes were solvated with TIP3P water molecules (Jorgensen et al. 1983) in an octahedral box, setting a distance of 10 Å between solute surface atoms and the box periphery. All solvated systems were neutralized by adding necessary counter ions, with extra Na^+/Cl^- ions added to bring the salt concentration to 150 mM.

4.3.2 Simulation runs

The simulation protocol, which involves four stages, was performed using the AMBER program (version 2018) (Case et al. 2018). (1) The systems were initially energy minimized for 5000 cycles using the steepest descent and conjugate gradient methods to circumvent unfavorable configurations. The minimizations were started by employing a 25 kcal/mol$Å^2$ restraint force on solute atoms, which was gradually lowered to conclude with a restraint-free minimization. (2) The systems were then heated to 300 K for 100 ps with a 10 kcal/mol$Å^2$ solute atom restraint. (3) Subsequently, the systems were equilibrated at constant pressure for 900 ps, while systematically reducing the solute atom constraint from 10 to 0.1 kcal/mol$Å^2$. A 5 ns unrestrained simulation was conducted to finalize the equilibration. (4) The production simulations were performed at constant temperature (300 K) and pressure (1.0 bar). Periodic boundary conditions were employed and a 9 Å cut-off was set for non-bonded interactions and the particle mesh Ewald method (Essmann et al. 1995) was used for long-range electrostatic interactions. The production runs were carried out for 100-600 ns with a 2 fs integration time step and coordinates were saved every 10-20 ps. In order to sample a broad conformational space, the simulations were often performed in duplicates or triplicates by varying the initial velocities in each run.

4.4 Analyses

The simulation trajectories were initially processed with CPPTRAJ (Roe and Cheatham 2013) and the overall dynamics of the different systems were visually inspected using the VMD program (Humphrey et al. 1996). The stability of the proteins was assessed with backbone-atom root-mean-square deviation (RMSD) and Cα-atom root-mean-square fluctuation (RMSF) calculations. Inter- and intramolecular hydrogen bond interactions were identified using CPPTRAJ, requiring a bond angle ≥135° and a bond distance ≤3.5 Å. Free energy of binding calculations and energy

decompositions were carried out with the molecular mechanics generalized Born surface area (MM-GBSA) method using the MMPBSA.py script (Miller et al. 2012) in AMBER. Principal component analysis (PCA) was conducted to probe the dominant pattern of movements observed during the simulations. The DSSP method (Kabsch and Sander 1983) was used to assign the secondary structure. The correlated motions between residues of interest were examined with a cross-correlation analysis in CPPTRAJ. Structural depictions included in this thesis were rendered with the Chimera program.

4.5 Experimental work

The experimental findings that served as a basis for the computational studies discussed in this thesis were carried out either by our collaborators in Prof. Klaus Elenius' laboratory (University of Turku) or by other research groups referred to in the thesis.

5 Results

5.1 The EGFR 746ΔELREA750 exon 19 deletion mutation (I, IV)

The ΔELREA mutation is the most commonly observed EGFR exon 19 deletion mutation in NSCLC. ΔELREA and the other types of EGFR exon 19 deletions increase EGFR autophosphorylation and activate the AKT and STAT pathways, enhancing cell survival (Paez et al. 2004; Sordella et al. 2004; de Gunst et al. 2007). The deletion of the amino acids 746ELREA750 takes place at the β3-αC loop of the kinase domain, which is located next to the functionally important αC helix. As a result of the deletion, the length of the β3-αC loop is shortened, which could affect the structural integrity of the αC helix and functional aspects related to it. Indeed, the structural model of the ΔELREA EGFR shows deformation of the N-terminal part of the αC helix, which is pulled towards the remaining β3-αC loop (Figure 9). Supporting this observation is the crystal structure of the BRAF serine/threonine kinase with a β3-αC loop deletion mutation, which exhibits a similar local structural change (Foster et al. 2016).

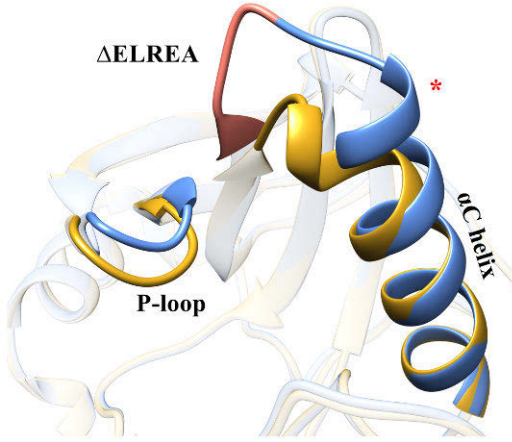


Figure 9. Superimposed structures of the wild-type (blue) and ΔELREA (gold) EGFR kinases. The 746ELREA759 sequence in the wild-type is colored orange. The N-terminal part of the αC helix (indicated by an asterix) in the mutant EGFR model is deformed and moves toward the shortened β3-αC loop, bringing it closer to the P-loop. Figure from publication I.

5.1.1 Effect of the ΔELREA mutation on the active state αC helix

The simulation of the ΔELREA EGFR kinase showed that the N-terminal end of the αC helix that is located next to the deleted amino acids uncoils, losing its helical composition (Figure 10A). The mutant simulation

additionally revealed that the N-terminus of the αC helix is placed closer to the P-loop, which encloses the ATP binding pocket and interacts with the nucleotide, as compared to the wild-type EGFR. The closer proximity of the αC helix to the P-loop in the ΔELREA EGFR is reflected by the shorter distance between Phe723 of the P-loop and Ile759 of the αC helix (average distance of 10.1 ± 1.0 Å). Consequently, in the mutant the P-loop tightly packs with the N-terminus of αC helix, imparting mutual stabilization to both structural units. In contrast, in the wild-type EGFR, the distance between Phe723 and Ile759 is 1.5 Å longer than seen for ΔELREA EGFR, with an average distance of 11.6 ± 1.0 Å.

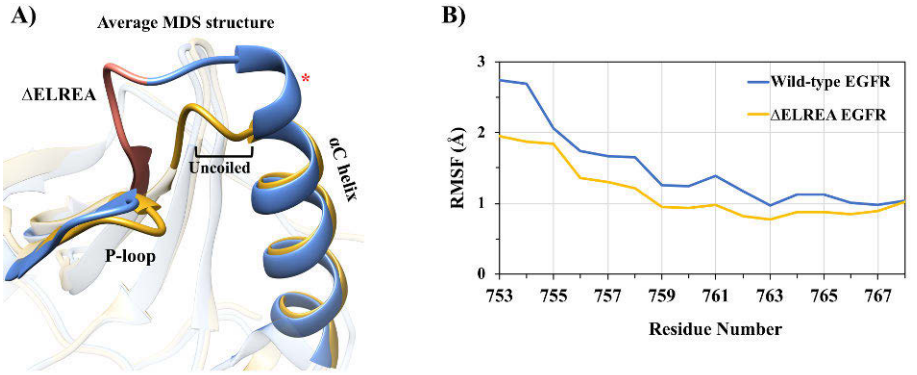


Figure 10. Structural dynamics of the αC helix. A) Superimposed average structures from the MDS of the wild-type (blue) and ΔELREA (gold) EGFR kinases. The N-terminal end of the αC helix in the mutant EGFR uncoiled during MDS, while the wild-type preserved the helical structure. B) Cα-atom RMSF of the αC helix during the wild-type and mutant EGFR simulations. Figure from publication I.

The Cα-atom RMSF calculation for residues of the αC helix (753-768) relative to the average structure showed that the amino acids fluctuate less in the active state ΔELREA EGFR (average RMSF of 1.1 ± 0.4 Å) as compared to the wild-type (1.5 ± 0.57 Å) (Figure 10B). This observation is reflective of the structural stability imparted on the αC helix, as a result of the deletion mutation, which shortens the length of the flexible β3-αC loop. Indeed, the β3-αC loop is often not resolved in crystal structures of the EGFR kinase domain, signifying the flexibility of the loop. The five-residue deletion would hence limit the flexibility of the remaining part of the β3-αC loop, consequently constraining the movement of the adjacent αC helix, which would provide the helix some degree of structural and positional stability. As a result, the αC helix would be skewed to remain in the activated "αC-in" conformation, promoting kinase activity and receptor signaling.

The concerted motions of the αC helix and the β3-αC loop during the simulations of the wild-type and ΔELREA EGFR were assessed using

principal component analysis (PCA) (Figure 11). PCA was carried out on the backbone atoms, which demonstrated lesser movement for the αC helix and the β3-αC loop in the ΔELREA mutant as compared to the wild-type EGFR. The PCA result was described with a porcupine plot, where the cones represent the backbone atom direction and amplitude of motion. The top three principal components (PCs) show that the αC helix and the β3-αC loop exhibit greater motions for the wild-type EGFR as represented by the bigger size of the cones, unlike that of the ΔELREA EGFR, which was more stable as indicated by the smaller cones. Moreover, in the wild-type EGFR the αC helix and β3-αC loop attain an outwards motion from the ATP binding pocket, which could trigger a conformational shift from the "αC-in" to the "αC-out" state.

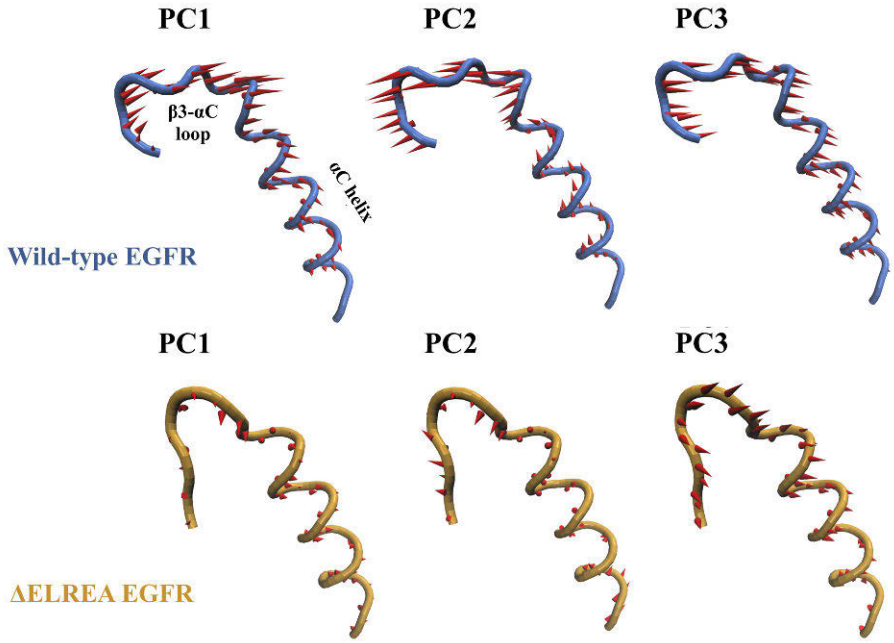


Figure 11. PCA of the αC helix and β3-αC loop of the wild-type and ΔELREA EGFRs. The three top PCs show that the αC helix and β3-αC loop of the wild-type EGFR are more mobile than the mutant, as illustrated by the bigger size of the cones. Figure from publication I.

The structural stability imparted on the αC helix by the deletion mutation might also affect the dimeric EGFR conformation, in addition to the monomeric state discussed above. The αC helix, the juxtamembrane B segment and the β4-β5 loop of the receiver kinase form part of the EGFR asymmetric dimer interface, accompanied by the αH and αI helices of the activator kinase domain. Even though the N-terminal part of the αC helix near the site of the deletion mutation is not directly involved in forming

the dimer interface, the ΔELREA mutation-induced structural stability on the entire αC helix could help stabilize the interactions taking place at the dimer interface, which could prolong the duration of the activated EGFR state.

5.1.2 The ΔELREA mutation stabilizes the Lys745-Glu762 salt bridge in the active state kinase domain

The conserved salt bridge of the kinase domain, between Lys745 of the β3 strand and Glu762 of the αC helix, is crucial for the catalytic activity of EGFR by optimally positioning lysine to interact with ATP. In the inactive state EGFR kinase, due to the displacement of the αC helix, the salt bridge is broken. The state of the Lys745-Glu762 salt bridge in the wild-type and ΔELREA EGFRs was assessed by measuring the distance between the Cδ atom of Glu762 and the Nζ atom of Lys745 (these atoms were selected in order to attain a central point of measurement) (Figure 12A). In the wild-type EGFR simulation the distance between Lys745 and Glu762 was longer and more variable (average distance of 4.0 ± 1.0 Å), with multiple sampled conformations between 20 and 70 ns of the simulation showing a distance exceeding 6 Å. In comparison, the ΔELREA EGFR maintained a shorter and more consistent distance (3.4 ± 0.2 Å) between the two residues. This observation indicates that the critical salt bridge was preserved in the ΔELREA EGFR, whereas it was broken in multiple conformations sampled during the wild-type simulation. Further supporting this notion is the frequency of hydrogen bonds between the side-chain carboxyl atoms of Glu762 (OE1, OE2) and the amine (NH3) nitrogen atom of Lys745 (Figure 12B). The hydrogen bonds were more frequently formed during the ΔELREA EGFR simulation as compared to the wild-type kinase domain. Collectively, the simulations revealed a better conservation of the functionally important Lys745-Glu762 salt bridge in the ΔELREA EGFR relative to the wild-type EGFR, likely owing to the positional stability of the αC helix, which accommodates Glu762.

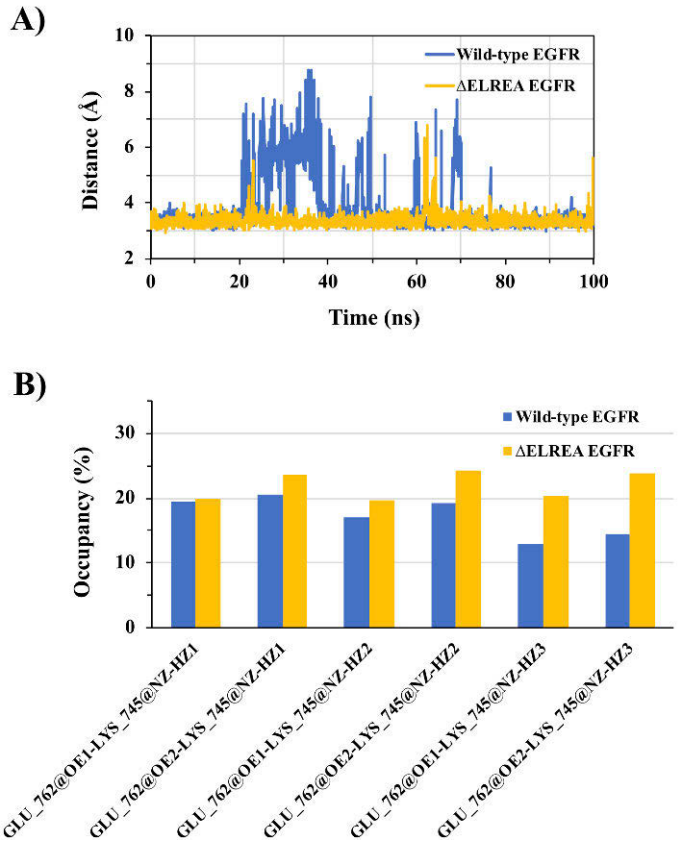


Figure 12. Dynamics of the Lys745-Glu762 salt bridge during the wild-type and ΔELREA EGFR kinase simulations. A) The distance between the side-chain polar atoms of Lys745 and Glu762. B) The hydrogen bond occupancy between the ε-amino and carboxyl side-chain atoms of Lys745 and Glu762 during the 100 ns simulation. Figure from publication I.

5.1.3 The ΔELREA mutation results in a stronger binding of ATP to active state EGFR

ATP binds at the active site of the EGFR kinase domain, where its γ-phosphate group is transferred to tyrosine residues located at the C-terminal tail. A magnesium ion coordinates with the phosphate groups of ATP, which aids to optimally position ATP for phosphate transfer. Simulations of the ATP-bound wild-type and ΔELREA EGFRs showed that Gln791 and Met793 at the hinge loop frequently interact with the adenosine moiety of ATP, which were formed during more than 90% of the simulation times of both the wild-type and mutant EGFRs (Figure 13A). Furthermore, the phosphate groups of ATP were observed

frequently interacting with Lys745 of the conserved salt bridge in the ΔELREA EGFR simulation than the wild-type.

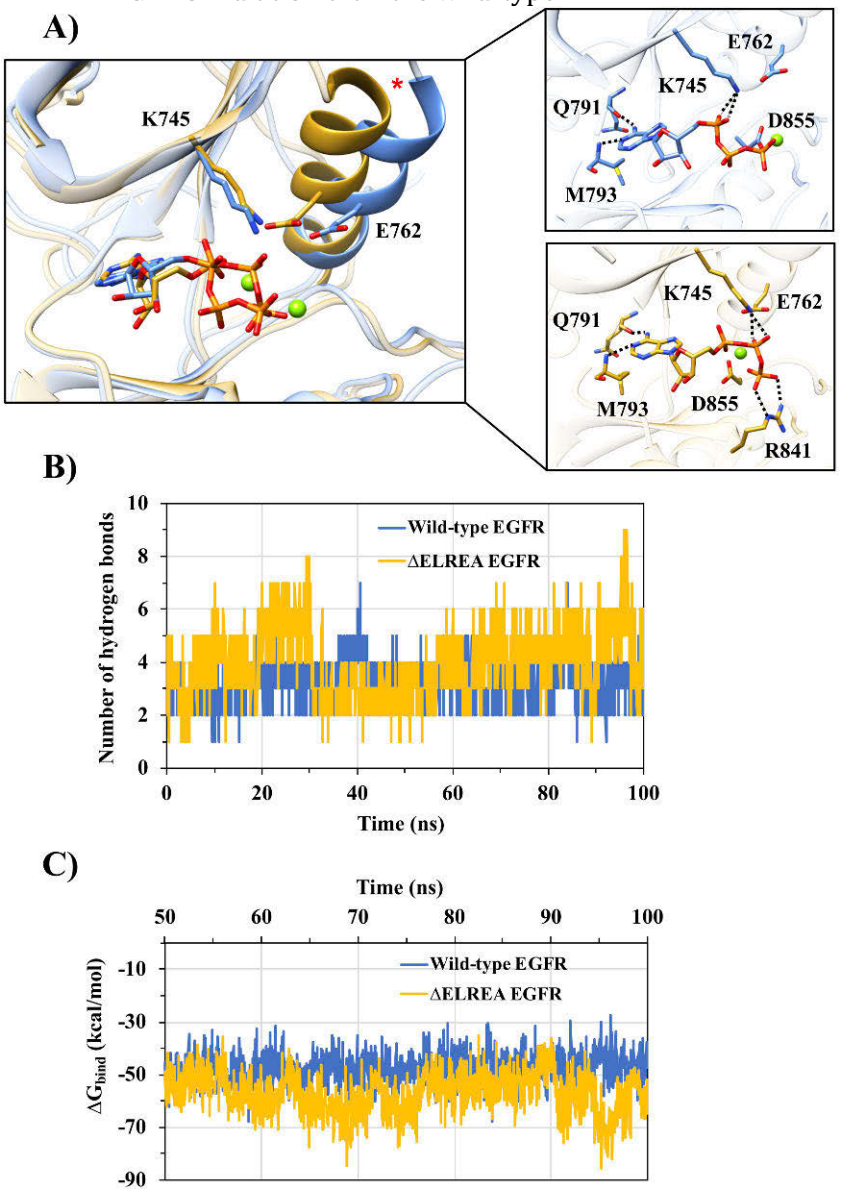


Figure 13. ATP binding to wild-type and ΔELREA EGFR kinases. A) superimposed MDS average structures of the ATP-bound wild-type (blue) and ΔELREA (gold) EGFRs (left). The magnesium ion that coordinates with ATP is depicted as a green sphere. Frequently formed hydrogen bond interactions between EGFR and ATP are highlighted in dotted lines (right). B) The number of hydrogen bonds between ATP and the two EGFRs during the 100 ns simulations. C) The free energy of binding of ATP to the wild-type and mutant EGFRs during the simulations. Figure from publication I.

The differential interaction between ATP and Lys745 in the two EGFR forms is linked to the stability of the αC helix, which varied between the mutant and wild-type simulations. In the wild-type simulation, the αC helix was less stable (average RMSF of 1.4 ± 0.7 Å) and was seen displaced from its initial position, adopting a similar conformation observed in the inactive EGFR state (Figure 13A). This conformational flexibility of the wild-type αC helix, which is fueled by the longer β3-αC loop, affects the formation of the Lys745-Glu762 salt bridge. Consequently, the optimal orientation of Lys745 is compromised, which in turn affects the interaction Lys745 makes with ATP. On the contrary, during the ΔELREA simulation, the αC helix was more stable (average RMSF 0.5 ± 0.17 Å) and preserved the "αC-in" conformation. This allowed for the strict formation of the Lys745-Glu762 salt bridge and the optimal orientation of Lys745 to form interactions with ATP. The mutant EGFR also formed more frequent interactions between ATP and Arg841, which was less often observed in the wild-type simulation. Indeed, the average number of hydrogen bonds between ATP and EGFR (Figure 13B) were higher in the ΔELREA EGFR simulation (4.0 ± 1.3) than the wild-type (3.2 ± 0.7), suggesting a stronger interaction between ATP and the mutant EGFR.

The relative binding affinity of ATP to the wild-type and ΔELREA EGFRs was examined by computing the free energy of binding utilizing the MM-GBSA method (Figure 13C). The analysis predicted that ATP has a higher binding affinity to the deletion mutant (average ΔG_{bind} -57 ± 7.9 kcal/mol) than to the wild-type EGFR (average ΔG_{bind} -48 ± 6.0 kcal/mol), as indicated by the smaller ΔG_{bind} value of the mutant. This observation is in line with the greater number of hydrogen bonds recorded between ATP and ΔELREA EGFR. Taken together, the simulations demonstrate that ATP likely binds more strongly to the ΔELREA EGFR than the wild-type, which might be one of the reasons behind the experimentally reported increase in kinase activity for the deletion mutant (Guha et al. 2008; Furuyama et al. 2013).

5.1.4 The ΔELREA mutation results in a conformational change on the inactive EGFR kinase

The inactive state ΔELREA EGFR kinase model structure exhibits an overall similar structure as the wild-type, except for the shortening of the β3-αC loop. During the simulations, however, additional interesting differences between the wild-type and ΔELREA EGFR structures were observed. The αC helix of the mutant moved towards the ATP binding pocket unlike in the wild-type EGFR, which largely maintained the initial location of the αC helix (Figure 14A). Moreover, the αC helix of the mutant exhibited a curved conformation, reflecting the pulling effect of the

shortened β3-αC loop. The small hydrophobic A-loop helix that packs against the αC helix also plays a role in forming the bent αC helix conformation by acting as a wedge and resisting a full inward motion. Although, the inwards movement of the αC helix in the ΔELREA EGFR also displaced the A-loop helix.

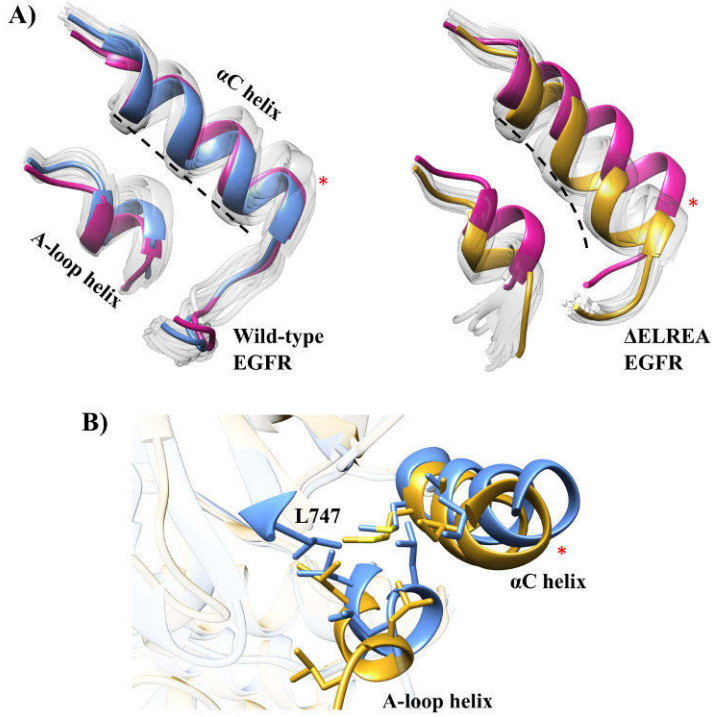


Figure 14. Structural differences observed between the wild-type and ΔELREA inactive EGFR kinases during the 100 ns simulations. A) Superimposed sampled conformations (faded white) from the wild-type (left) and mutant (right) EGFR simulations. The corresponding median structures are in blue (wild-type) and gold (mutant). The initial structures prior to energy minimization and MDS are in pink. The superimposed structures show that during the ΔELREA simulation the αC helix dislocated and attained a curved orientation, unlike the wild-type EGFR. B) Displacement of the αC helix and A-loop helices in the ΔELREA EGFR disrupts key hydrophobic interactions between the two helices, in addition to the deletion of L747, which is part of the hydrophobic network in the wild-type EGFR. Figure from publication I.

The hydrophobic residues of the A-loop helix interact with nearby hydrophobic amino acids located at the αC helix and Leu747 of the β3-αC loop (Figure 14B). However, due to the displacement of the A-loop helix in the mutant EGFR, the distance between the A-loop and the αC helices is widened as supported by the distance between the Cα atom of Leu863 of the A-loop helix and the Cα atom of Ile759 of the αC helix. The average distance between the two residues was longer in the ΔELREA EGFR (7.4 ±

0.49 Å) than the wild-type (5.8 ± 0.34 Å). Furthermore, due to the 746ELREA750 deletion, Leu747 of the β3-αC loop is missing in the mutant EGFR, hence diminishing the strength of the hydrophobic cluster important to maintaining the inactive EGFR conformation.

5.2 The EGFR V769insASV and D770insNPG exon 20 insertion mutations (II)

The V769insASV and D770insNPG activating mutations belong to the most predominant types of EGFR exon 20 insertions implicated in NSCLC (Figure 15A, 15B). The insertions increase the catalytic activity of EGFR and promote cell proliferation (Yasuda et al. 2013). The V769insASV mutation is located at the C-terminal end of the αC helix of the kinase domain, whereas the D770insNPG insertion is placed at the succeeding αC-β4 loop (Figure 15C). The crystal structure of the D770insNPG EGFR (PDB ID: 4LRM) shows that the insertion mutant adopts the active EGFR conformation. The three inserted residues, *NPG* (inserted residues are in italics), form a β-turn at the beginning of the αC-β4 loop (Figure 15C), which is maintained by a hydrogen bond between the main-chain nitrogen atom of the inserted *Gly773* and the main-chain oxygen atom of Asp770. Since the only available EGFR exon 20 insertion mutant structure is of the D770insNPG, the V769insASV mutant structure was modeled. The V769insASV EGFR model structure shows that the insertion of the *ASV* residues at the C-terminus of the αC helix results in an additional helix-turn (Figure 15C), with hydrogen bonds being formed between the main-chain oxygen atom of Val769 and the main-chain nitrogen of the inserted *Val772*.

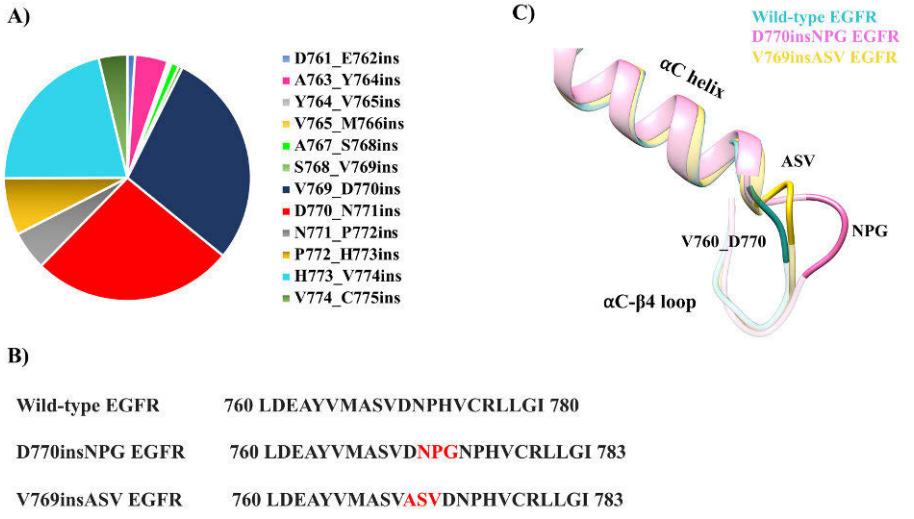


Figure 15. The V769insASV and D770insNPG exon 20 EGFR kinase insertion mutations. A) The frequency of EGFR exon 20 insertion mutations sampled in

NSCLC, according to the COSMIC database (v92). The top two insertion mutation categories are V769_D770ins and D770_N770ins, in which the V769insASV and D770insNPG EGFR mutations belong to. B) Sequences of the wild-type, V769insASV and D770insNPG EGFRs near the location of the mutations. The three inserted residues in the mutant EGFRs are in red. C) Superimposed structures of the wild-type and insertion mutant EGFRs, showing the structural changes at the site of the insertions relative to the wild-type EGFR. Figure from publication II.

5.2.1 The exon20 insertion mutations stabilize the active state αC helix

The conformation of the αC helix governs the active-inactive kinase states. Structural alterations on or near the helix may therefore affect the equilibrium between the two kinase states. Given the proximity of the V769insASV and D770insNPG mutations to the αC helix, it is essential to assess the impact of the insertions on the helix. The 600 ns simulations revealed that the αC helix of the D770insNPG and V769insASV EGFRs is more stable than the wild type, as supported by the lower Cα-atom RMSF difference values of the residues of the mutant helix (Figure 16A). The better stability of the αC helix in the mutants is likely attributed to the greater number of interactions present at the C-terminus of the helix. These interactions are formed between the inserted amino acids and residues of the αC helix, αE helix and αC-β4 loop, which aid to hold the αC helix in the "αC-in" active conformation. A per-residue secondary structure analysis for the αC helix of the wild-type and insertion mutants demonstrated that the mutants better preserved the helical content for the central and C-terminal part of the αC helix, compared to the wild-type EGFR, which exhibited turns and bends in multiple sampled conformations (Figure 16B). Both the wild-type and mutant EGFRs however had similar secondary structure representations for the N-terminal part of the αC helix, as the helical content was largely lost in all three EGFRs.

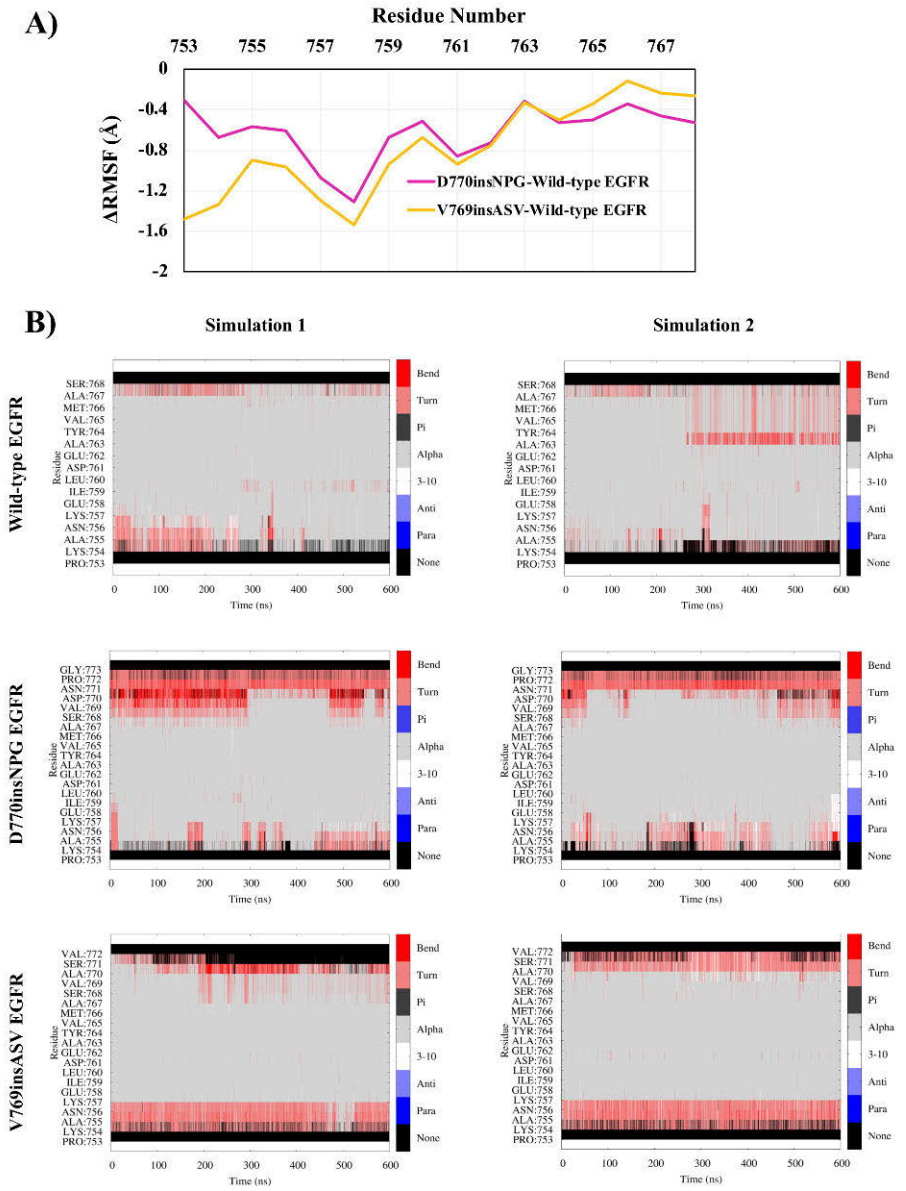


Figure 16. The αC helix during the wild-type, V769insASV and D770insASV EGFR kinase simulations. A) RMSF difference between the αC helix of the wild-type EGFR and the two insertion mutants. B) Secondary structure assignments for the residues of the αC helix during the duplication simulations of the wild-type and mutant EGFRs. The inserted residues in V769insASV and D770insASV EGFRs are included in the analysis. Figure from publication II.

The dominant concerted motions of the αC helix during the wild-type and insertion mutant simulations were probed with PCA (Figure 17A). Motions described in the first PC, represented in a porcupine plot, support

the observation made from the RMSF analysis, where the αC helix of the mutants is more stable than the wild type. The size of the cones on the porcupine plot signifies the magnitude of motion of the αC helix, which is smaller in the insertion mutants, in comparison to the wild type. Furthermore, the direction of motion of the cones indicates that the αC helix of the D770insNPG and V769insASV mutants makes an inwards movement - towards the binding pocket, which is consistent with the active state EGFR conformation. On the contrary, the wild-type EGFR adopts an outwards motion of the αC helix, similar to that observed in the inactive EGFR. These findings suggest that the D770insNPG and V769insASV insertion mutations would lead to increased kinase activity by conserving the active state of EGFR through the stabilization of the αC helix in the "αC-in" conformation.

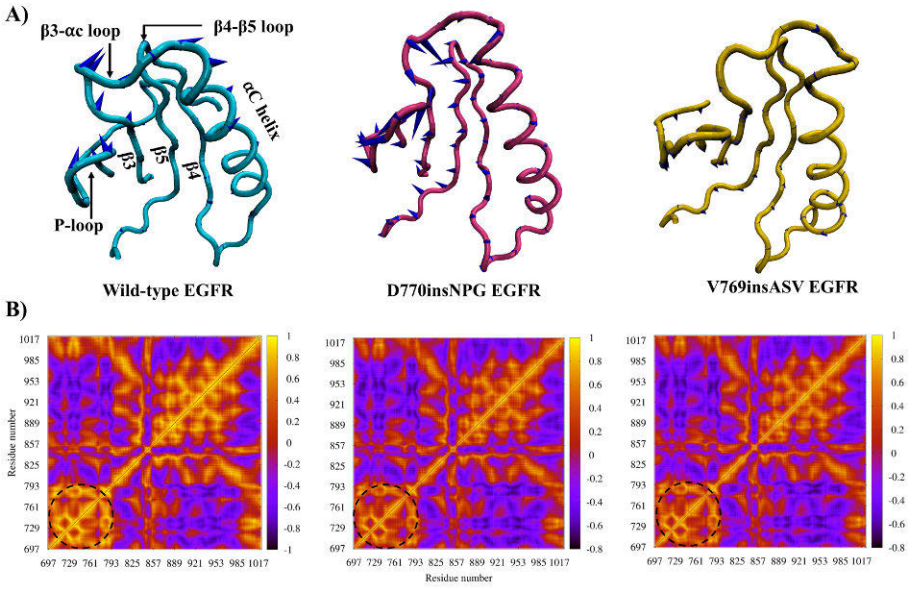


Figure 17. Movements of the αC helix and the N-lobe units of the kinase domain during the simulation of the wild-type, D770insNPG and V769insASV EGFRs. A) The first PC from the PCA analysis shows that the αC helix of the wild-type EGFR is more mobile than the insertion mutants. Furthermore, the majority of the N-lobe structural elements adopt an outwards motion (away from the binding pocket) as represented by the direction of the cones in the porcupine plot. In contrast, the N-lobe in the insertion mutants largely attains an inward motion. B) Cross-correlation analysis for the residues of the wild-type and mutant EGFRs. Movements from correlated to non-correlated to anti-correlated are colored in gradient from yellow to brown to black. The amino acids of the N-lobe of the kinase domain (encircled) exhibit highly correlated motion during the simulations of both the wild-type and mutant EGFRs. Figure from publication II.

PCA additionally showed that the movement of the αC helix is in sync with the N-lobe structural units, such as the β3-αC loop, β4-β5 loop, P-loop, β3 and β4 strands. In the wild-type EGFR these structural units attained an outwards motion, away from the binding pocket as seen for the αC helix. Whereas, in the insertion mutants these N-lobe structures mostly adopted an inward movement, mirroring the movement of the αC helix. The correlated motions among the structural elements of the N-lobe of the kinase domain in both the wild-type and insertion mutants is substantiated by a cross-correlation analysis, which showed higher correlated movements for this region during the MDS (Figure 17B).

5.2.2 The insertion mutations generate additional interactions at the mutation site of active state EGFR

The V769insASV and D770insNPG insertions, in addition to resulting structural changes at the site of the mutations, also alter the interactions being formed. As a result of the D770insNPG insertion, Asp770 slightly changes its location, with its side chain coming closer to Arg779 (Arg776 in the wild-type) of the αC-β4 loop (Figure 18A, left). During the simulation of the D770insNPG EGFR kinase, the interaction between Asp770 and Arg779 was indeed observed for 83% of the simulation time, while in the wild-type simulation the interaction was formed only for 25% of the simulation time. Furthermore, in the D770insNPG simulation the newly formed β-turn was maintained by two interactions: a hydrogen bond between the main-chain oxygen of Asp770 and the nitrogen atom of the inserted *Gly773* (49%), and a hydrogen bond between the main-chain oxygen of Ala767 and main-chain nitrogen of *Asn771* (10%) (Figure 18A, right). These interactions are lacking in the wild-type simulation, which on top of the less frequently observed Asp770-Arg776 interaction could impart more flexibility to the C-terminus of the αC helix, as compared to the D770insNPG EGFR.

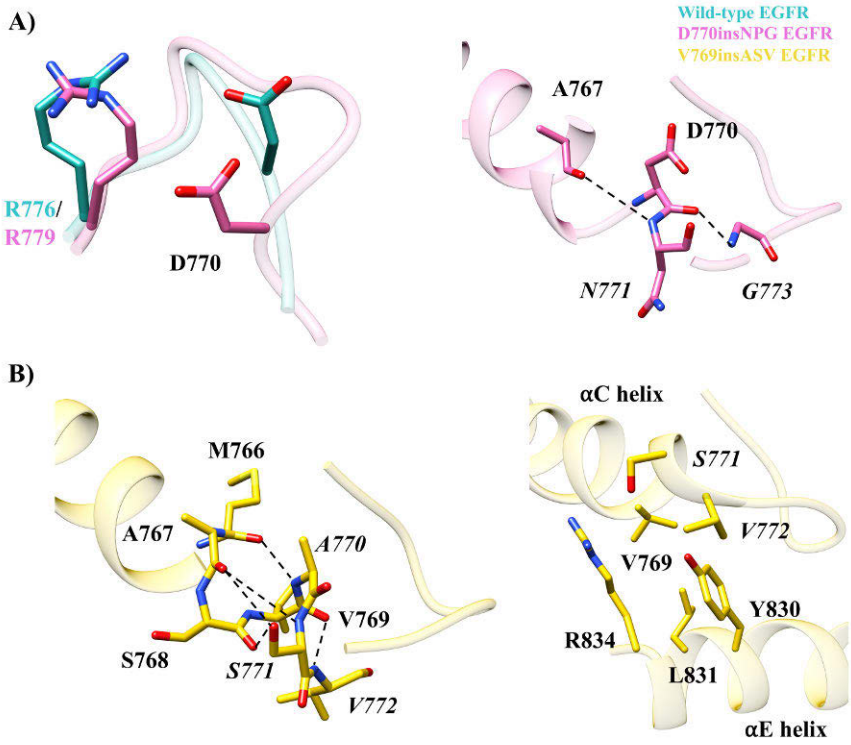


Figure 18. Interaction near the site of the D770insNPG and V769insASV mutations. A) The orientations of Arg776/Arg779 and Asp770 residues in the wild-type and D770insNPG EGFR kinase crystal structures (left). Hydrogen bonds (dotted line) observed during the D770insNPG EGFR simulation that stabilize the β-turn formed at the site of the insertion (right). B) Frequently formed hydrogen bond interactions near the site of mutation in the V769insASV EGFR simulation (left). Key interactions between αC and αE helices (right) in V769insASV EGFR. Figure from publication II.

During the V769insASV EGFR simulation, several hydrogen bonds were formed near the inserted residues: *Ala770* N-Met766 O (86%), *Ser771* OH-Ser768 O (11%), *Ser771* OH-Ala767 O (16%), *Ser771* N-Ala767 O (16%) and *Val772* N-Val769 O (9%) (Figure 16B, left). These interactions help maintain the newly formed helix-turn at C-terminus of the αC helix. Additionally, interactions between the inserted residues and amino acids of the adjacently placed αE helix (*Val772*-Tyr827, *Ser771*-Arg831) further strengthen the stability of the active state αC helix (Figure 18B, right). Furthermore, during the simulations of both V769insASV and D770insNPG mutants, Leu831 of the αE helix was placed closer to Val769 of the αC helix (average distance of 7.5 ± 0.4 Å for V769insASV and 8.0 ± 0.5 Å for D770insNPG), relative to the wild-type EGFR (9.5 ± 0.7 Å). Consequently, the mutants could experience a stronger hydrophobic interaction between Val769 and Leu831, which would play a role in

stabilizing the αC helix. In conclusion, the analysis of the simulations revealed that the D770insNPG and V769insASV mutations generate additional interactions involving the inserted residues, which could provide further stability to the αC helix.

5.2.3 The insertion mutations preserve the Lys745-Glu762 salt bridge in the active state EGFR

In order to examine whether the insertion mutations alter the state of the catalytically important Lys745-Glu762 salt bridge, the distance between the two residues were measured during the simulations of the mutant and wild-type EGFR kinases (Figure 19A). The distance between the Cδ atom of Glu762 and the Nζ atom of Lys745 was shorter and less variable in the D770insNPG (average distance of 3.5 ± 0.5 Å) and V769insASV (3.45 ± 0.46 Å) EGFR simulations. During the wild-type EGFR simulation the distance was longer and less consistent (6.0 ± 2.6 Å), in particular from 250 ns onwards of the simulation time, suggesting breakage of the salt bridge. In line with this observation is the frequency of hydrogen bonds formed between the side-chain polar atoms of Lys745 (NH3) and Glu762 (OE1, OE2) during the simulations; these interactions were formed more often in the insertion mutants than the in wild-type EGFR (Figure 19B). The better stability of the Lys745-Glu762 salt bridge in the exon 20 insertion mutants may be attributed to the higher structural and positional stability of the αC helix, where Glu762 resides in. In the wild-type EGFR, the higher flexibility observed for the αC helix results in the displacement of Glu762, hence preventing the formation of the salt bridge key for EGFR kinase activity.

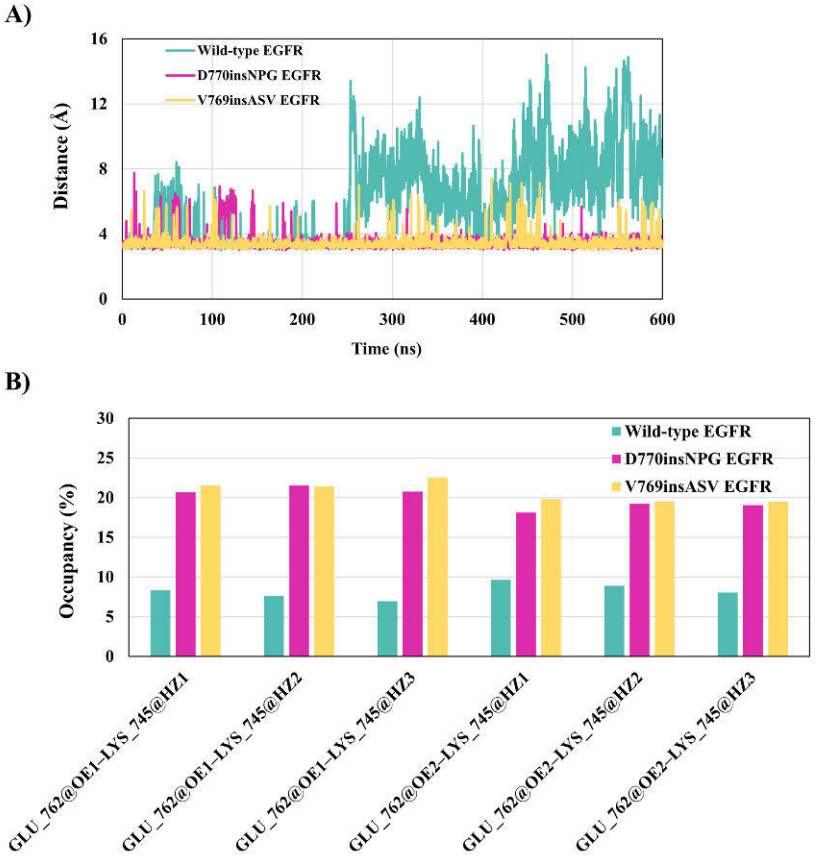


Figure 19. The Lys745-Glu762 salt bridge during the wild-type and insertion mutant simulations. A) Distance between the Nζ atom of Lys745 and Cδ atoms of Glu762. Multiple sampled conformations during the wild-type simulation indicate breakage of the Lys745-Glu762 bond, in particular from 250-600 ns of the simulation time. B) The frequency of the hydrogen bonds between the side-chain polar atoms of Lys745 and Glu762. The hydrogen bonds were more often formed in the mutant simulations in contrast to the wild-type EGFR kinase. Figure from publication II.

5.2.4 The insertion mutations stabilize the active state DFG motif and R-spine

The DFG motif is one of the structural elements of the kinase domain that exhibits alterations based on the active-inactive state of the EGFR kinase. In the Src-like EGFR inactive state, Phe856 of the DFG motif attains a different side chain orientation than in the active state, with the phenyl ring shifting towards the αC helix (Figure 6B). In order to assess the impact of the insertion mutations on the state of the DFG motif, the χ1 dihedral angle of the Phe856 side chain was monitored during the

duplicate simulations of the wild-type and mutant EGFRs (Figure 20A). In simulation 1, a χ1 dihedral angle similar to the one observed in the active state EGFR crystal structure (-63.8°) was recorded for D770insNPG (-65.2 ± 9.9°) and V769insASV (-59.2 ± 10.25°) EGFRs. In contrast, between 100 and 450 ns of simulation 1 of the wild-type EGFR, the χ1 dihedral angle (68.1 ± 10°) was similar to the Phe856 orientation seen in inactive EGFRs (50°–55°) – PDB ID: 3POZ and 3W2S. During 450 to 575 ns of the simulation, the χ1 angle (178.7 ± 10.4°) was in a similar range with the ones observed in other inactive EGFR crystal structures (163°–169°) – PDB ID: 2GS7 and 4HJO, where the phenyl ring of Phe856 leaves the hydrophobic groove at the C-terminus of the αC helix. In simulation 2, both the wild-type and mutant EGFRs had a χ1 dihedral angle similar to the active state EGFR.

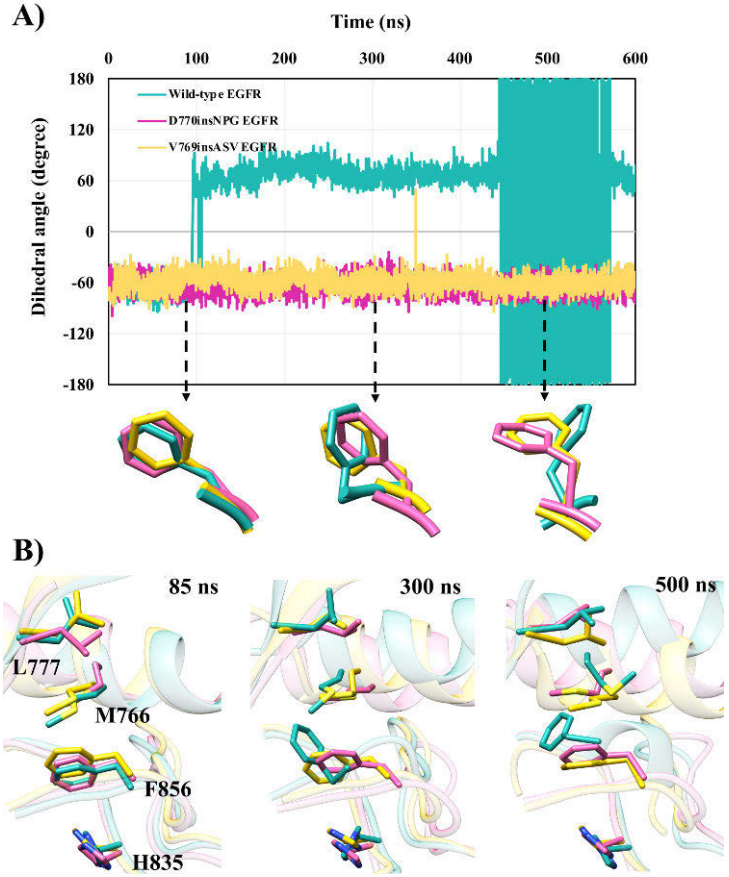


Figure 20. Phenylalanine of the DFG motif and R-spine of the EGFR kinase domain. A) χ1 dihedral angle of Phe856 of the DFG motif during the wild-type, V769insASV and D770inNPG EGFR kinase domain simulations. In the wild-type simulation, Phe856 attained a dihedral angle observed in inactive EGFR kinase structures, unlike the mutant EGFRs. The χ1 angle values between 450 to 575 ns

of the wild-type simulation deceivingly exhibit high variation in the graph, although the actual differences are small. This is the result of the dihedral angle cap (180° to -180°), hence, an angle of 182° would be represented as -178° in the graph, even if the angle difference is just 2°. B) Superimposed sampled conformations at 85, 300 and 500 ns of the simulation showing the orientations of the R-spine residues. At 500ns the side chain of both Met766 and Phe856 of the wild-type EGFR changed to the one observed in the inactive EGFR conformation. Figure from publication II.

The R-spine, which is key to stabilizing the active conformation of the kinase domain, is formed by the spatial arrangement of four non-consecutive hydrophobic residues including Phe856 of the DFG motif and Met766 of the αC helix. In simulation 1 of the wild-type EGFR, the side-chain orientation of Met766 altered towards the rotamer seen in the inactive state due to the positional change of the αC helix along with the reorientation of the Phe856 side chain (Figure 20B). As a result, the relative arrangement of the R-spine residues is disrupted. In comparison, in the D770insNPG and V769insASV EGFR simulations the R-spine assembly was maintained, which can be associated with the conserved sidechain orientation of Phe856 and the more stable αC helix. Taken together, the simulations revealed that unlike the wild-type EGFR, the D770insNPG and V769insASV mutations help maintain the integrity of the DFG motif and the R-spine, which are fundamental for the stability of the active conformation of EGFR.

5.2.5 Effect of the insertion mutations on key units of the inactive state kinase domain

The simulations of the wild-type and insertion mutant inactive EGFR kinases suggest the maintenance of the inactive kinase domain conformation, as represented by superimposed sampled conformations from the MDS (Figure 21A). The structures adopted an αC helix with the "αC-out" conformation and an A-loop with an intact one-turn helix. Furthermore, the salt bridge between Lys745 and Glu762 was broken in all EGFR simulations, with an average distance of 15 Å between the side-chain polar atoms of the two residues (Figure 21B). The orientation of Phe856 of the DFG motif was maintained as the inactive state rotamer in simulation 2 of the wild-type and mutant simulations. However, in simulation 1 of the V769insASV EGFR (between 400 and 450 ns), the χ1 dihedral angle of Phe856 was similar to the EGFR active conformation (Figure 21C), unlike the wild-type and D770insNPG EGFRs, which attained a χ1 angle as the inactive EGFR. The V769insASV mutation hence might advocate structural changes towards the active EGFR conformation that may only be seen in very long simulations.

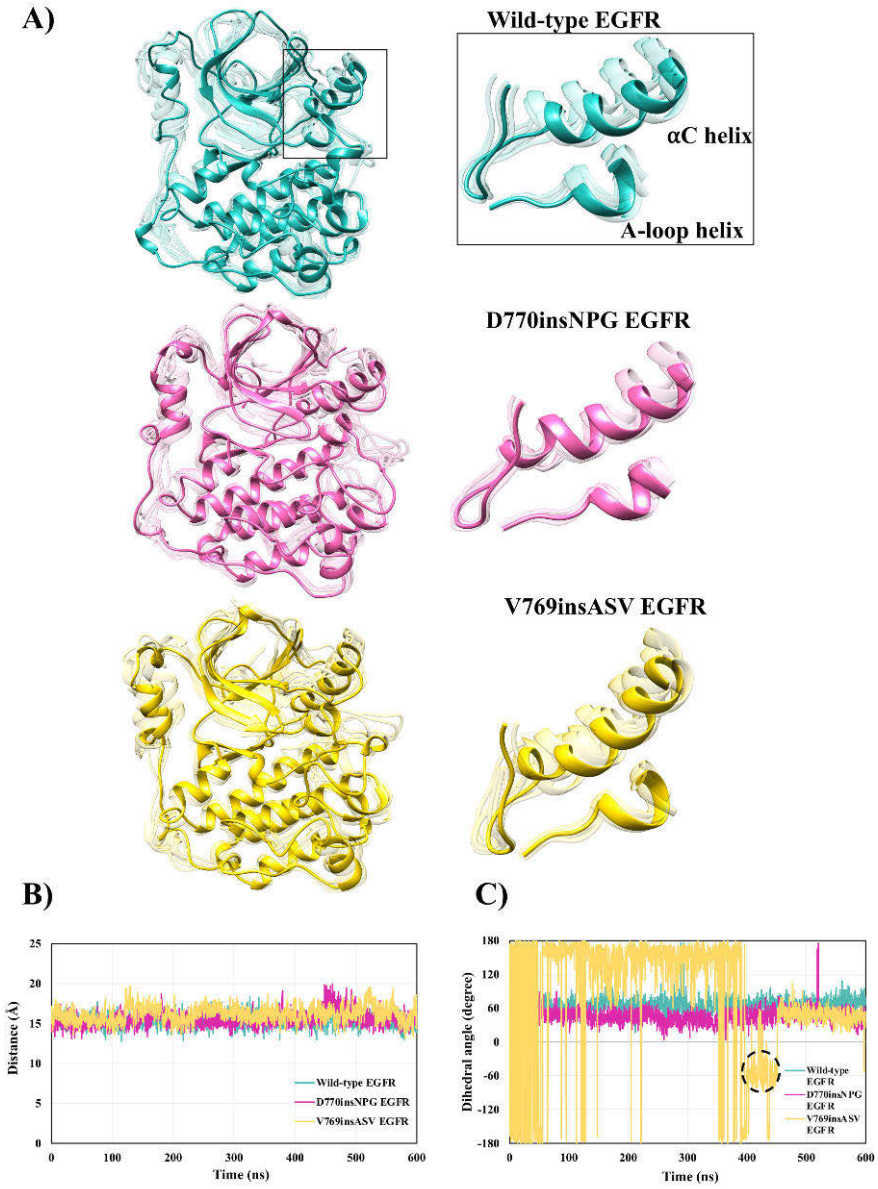


Figure 21. Dynamics of the wild-type, D770insNPG and V769insASV inactive EGFR kinase domains. A) Sampled conformations (faded color) during the wild-type and insertion mutant EGFR simulations are superimposed on their respective median structure (solid color). The structures maintain the inactive EGFR conformation in both the wild-type and mutant simulations, with an αC helix in the "αC-out" state and a preserved A-loop helix. B) The distance between Lys745 (Nζ) and Glu762 (Cδ) during the wild-type and insertion mutant simulations. C) χ1 dihedral angle of Phe856 of the DFG motif during the wild-type, V769insASV and D770inNPG inactive EGFR simulations. The χ1 dihedral angle of the wild-type and D770inNPG EGFRs maintain the side-chain orientation of

Phe856 in the inactive state EGFR, whereas, the V769insASV EGFR attains a dihedral angle observed in the EGFR active state between 400 and 450 ns of the simulation (encircled). Figure from publication II.

5.2.6 The insertion mutants alter the Ala767-Arg776 autoinhibitory interaction in the inactive EGFR

In the inactive conformation of the wild-type EGFR kinase structure, an autoinhibitory interaction between the main-chain oxygen atom of Ala767 located at the C-terminus of the αC helix and the guanidinium group of Arg776 of the αC-β4 loop helps to hold the αC helix in the "αC-out" conformation (Ruan and Kannan 2015). In the mutant EGFRs, this key interaction is seen compromised due to the introduction of the inserted residues (Figure 22A). In the D770insNPG model structure, Asp770 is placed in close proximity to Arg779 (Arg776 in wild-type EGFR), posing a steric hindrance, which could alter the orientation of Arg779 and interfere with the interaction being formed between Arg779 and Ala767. Similarly, in the V769insASV EGFR structure, the inserted *Ala770* poses the same challenge as Asp770 in D770insNPG EGFR, as *Ala770* is situated close to Arg779.

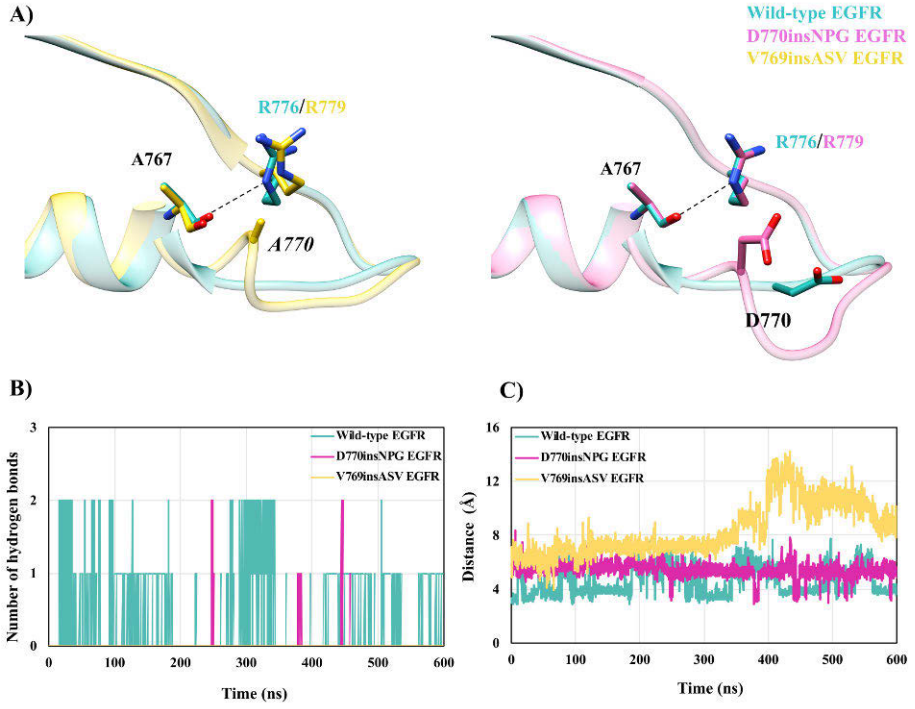


Figure 22. The autoinhibitory Ala767-Arg776/Arg779 interaction in the inactive EGFR kinase domain. A) The hydrogen bond between Ala767 and Arg776 (dotted line) present in the wild-type EGFR is disturbed by the mutation-

induced positioning of *Ala770* in V769insASV EGFR and Asp770 in D770insNPG EGFR. B) The number of hydrogen bonds between Ala767 and Arg776/Arg779 during the wild-type and mutant 600 ns simulations. C) The distance between Ala767 (O) and Arg776/Arg779 (Cζ) during MDS of the wild-type and insertion mutant inactive EGFRs. Figure from publication II.

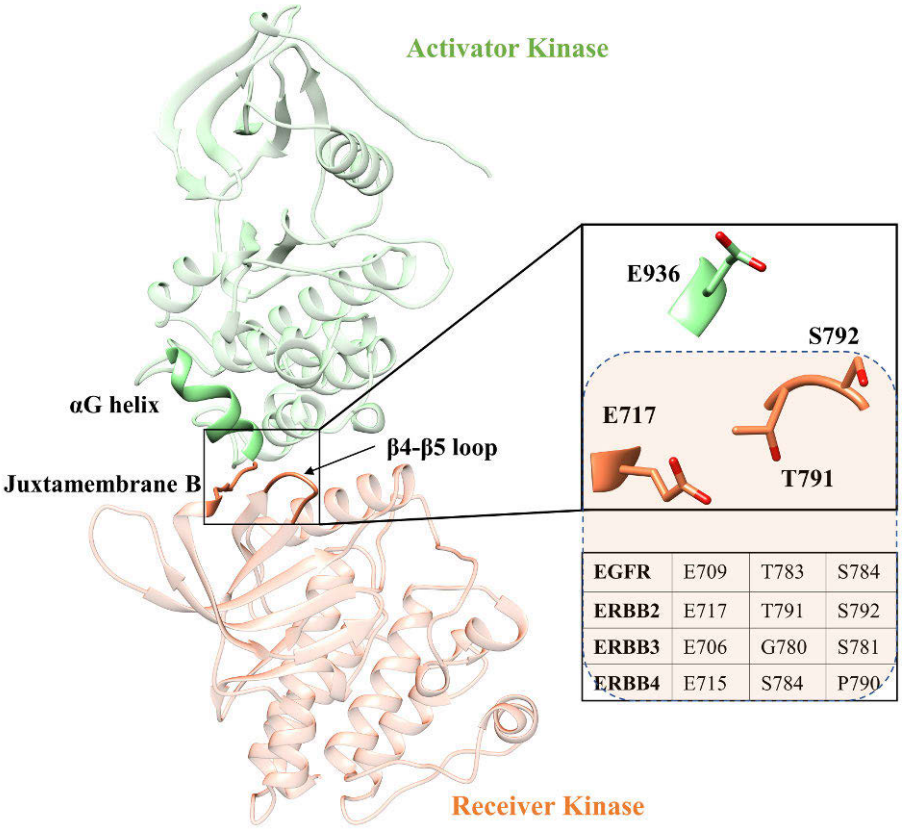
These observations are consistent with the dynamics data, which showed that the hydrogen bond between Arg779 and Ala767 is rarely formed during the D770insNPG (1.33%) and V769insASV (0%) EGFR simulations (Figure 22B). In contrast, this interaction was preserved for 53% of the wild-type simulation. During the D770insNPG EGFR simulation, the positioning of Asp770 indeed resulted in the reorientation of the side chain of Arg779, which prevented the formation of the Ala767-Arg779 interaction. Similarly, in the V769insASV simulation, *Ala770*, which is jammed between Ala767 and Arg779, obstructed the formation of the autoinhibitory hydrogen bond. Indeed, the distance between Ala767 (O) and Arg779/Arg776 (Cζ) was longer in the V769insASV (average distance of 8.4 ± 1.9 Å) and D770insNPG (5.4 ± 0.5 Å) EGFR simulations, as compared to the wild-type EGFR (4.5 ± 1.0 Å) (Figure 22C).

5.3 The ERBB2 E936K kinase domain mutation (III)

The E936K somatic mutation in the ERBB2 kinase domain was observed in a human cancer cell line derived from a leukemia patient (Barretina et al. 2012). As demonstrated in experiments carried out by our collaborators in publication III, the mutation increases ERBB2 activity by enhancing tyrosine phosphorylation. The E936K mutation promoted ERBB2 phosphorylation in ERBB2-ERBB2 homodimers and ERBB2-EGFR heterodimers, with the latter exhibiting twice higher phosphorylation as compared to the wild-type ERBB2-EGFR dimer. The E936K mutation however did not alter kinase activity in the context of ERBB2-ERBB3 heterodimers. In order to uncover the mechanism by which the mutation exerts its activating effect, the structure of the mutant was investigated and compared with the wild-type ERBB2 structure.

In the X-ray crystal structure of the wild-type ERBB2 homodimer (PDB ID: 3PP0), Glu936 is situated at the αG helix of the kinase domain C-lobe (Figure 23). In the homodimeric assembly, Glu936 of the activator kinase is part of the asymmetric dimer interface, which is comprised of αI and αH helices of the activator kinase, and JM-B segment, β4-β5 loop and αC helix of the receiver kinase. The crystal structure reveals that Glu936 of the activator kinase is capable of making interactions with amino acids in the β4-β5 loop (Thr791, Ser792) and the C-terminus of the JM-B segment (Glu717) of the receiver kinase. Glutamate 936 is conserved among the ERBB kinases and the residues at the N-lobe of the receiver kinase

interacting with Glu936 are also conserved in ERBB2 and EGFR, whereas different residues are present at the β4-β5 loop of ERBB3 (Gly780, Ser781) and ERBB4 (Ser784, Pro790).



EGFR	E709	T783	S784
ERBB2	E717	T791	S792
ERBB3	E706	G780	S781
ERBB4	E715	S784	P790

Figure 23. The ERBB2 asymmetric kinase homodimer (PDB ID: 3PP0). Glutamate 936 from the activator kinase is located near the ERBB-conserved Glu717 and residues at the JM-B segment and β4-β5 loop of the receiver kinase. Figure from publication III.

Analysis of the wild-type ERBB2-ERBB2 homodimer and ERBB2-EGFR heterodimer structures shows that the side-chain carboxyl group of Glu936 of the activator kinase can make hydrogen bonds with both the backbone and side-chain atoms of Ser792 (ERBB2)/Ser784 (EGFR) of the receive kinase (Figure 24A, left). Although the E936K mutant can still form similar interactions with Ser792/Ser784, a unique and stronger ionic interaction is formed between Lys936 and Glu717/Glu709 of the receiver kinase, which is lacking in the wild-type dimers (Figure 24A, right). The Lys936-Glu717/Glu709 interaction would consequently enhance the activator-receiver kinase binding in the mutant ERBB2-ERBB2 homodimer and ERBB2-EGFR heterodimer, which could extend the timespan of the activated ERBB2 and EGFR states. In the context of the

ERBB2-ERBB4 heterodimer, the Lys936-Ser784 and Lys936-Glu715 interaction can conceptually take place (Figure 24B). However, ERBB4 prefers to take the place of the activator kinase when complexed with ERBB2, therefore diminishing the possibility of the above interactions being formed. Similarly, the mutation might not have an effect on the ERBB2-ERBB3 heterodimer, as the catalytically impaired ERBB3 largely functions as the activator kinase. As a result, the E936K mutation is unlikely to have a significant impact on the dimeric binding of the ERBB2-ERBB4 and ERBB2-ERBB3 heterodimers, unlike the ERBB2-ERBB2 and ERBB2-EGFR dimers.

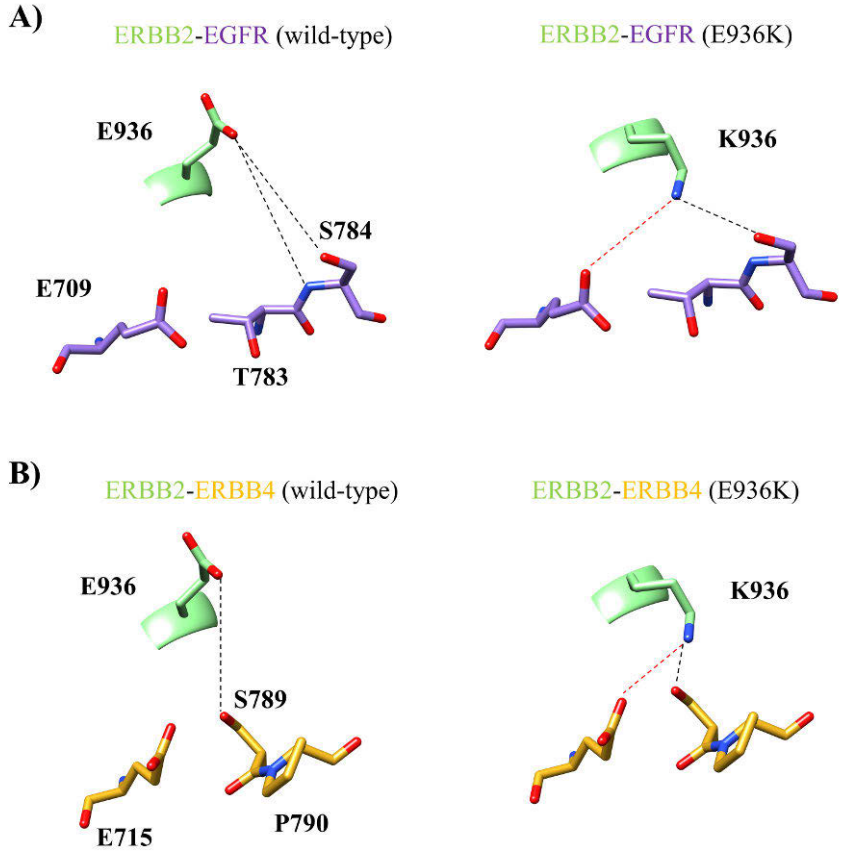


Figure 24. Dimer interface interactions between Glu936 (wild-type)/Lys936 (mutant) in ERBB2 and residues of EGFR and ERBB4 in ERBB2 heterodimers. Hydrogen bond (black dotted line) and ionic interactions (red dotted line) between Glu936/Lys936 in ERBB2 (activator kinase) and A) EGFR, and B) ERBB4 (receiver kinases). Figure from publication III.

MDS of the wild-type and E936K mutant ERBB2-EGFR structures revealed that the hydrogen bond between Ser784 of EGFR (receiver kinase) and Lys936 in the ERBB2 mutant (activator kinase) is formed

twice as often (9.2% of the simulation time) compared to Ser784-Glu936 in the wild-type heterodimer (4.2% of the simulation time). This could be linked to the better stability of the β4-β5 loop of the mutant EGFR that accommodates Ser784. In the mutant heterodimer the loop had an average RMSD of 0.26 ± 0.06 Å and in the wild-type simulation the RMSD of the loop was 0.40 ± 0.18 Å (Figure 25A). Additionally, interaction energy analysis showed that the Lys936-Glu709 interaction in the mutant heterodimer was more favored (ΔG = -0.4 ± 0.3 kcal/mol) in comparison to the Glu936-Glu709 interaction in the wild-type ERBB2-EGFR dimer (ΔG = -0.03 ± 0.15 kcal/mol). This would result in stronger dimeric binding for the ERBB2-EGFR heterodimer, a conclusion further supported by the free energy of binding values of the heterodimers (Figure 25B): The mutant ERBB2-EGFR dimer had a lower average ΔG of binding (-77.1 ± 8.2 kcal/mol) relative to the wild-type heterodimer (-66.6 ± 11.1 kcal/mol), consistent with stronger binding between ERBB2 and EGFR in the E936K mutant heterodimer.

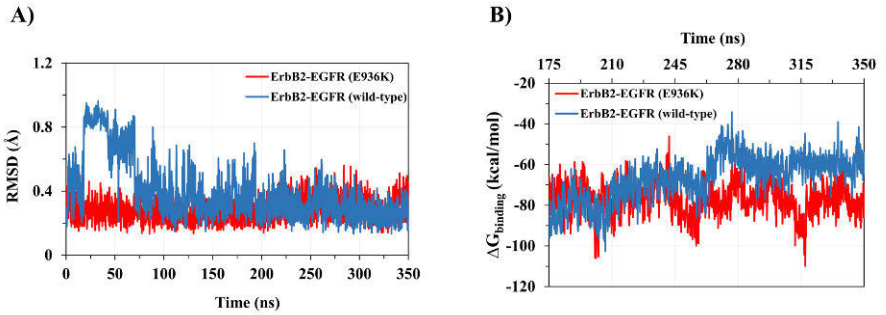


Figure 25. Dynamics of the wild-type and E936K ERBB2 mutant in the ERBB2-EGFR kinase heterodimers. A) Backbone atom RMSD of the β4-β5 loop in EGFR (receiver kinase) during the wild-type (blue) and E936K (red) mutant ERBB2-EGFR heterodimer simulations. B) The free energy of binding between EGFR and ERBB2 in wild-type and E936K ERBB2-EGFR simulations. Figure from publication III.

6 Discussion

Activating somatic mutations that occur on the ERBB tyrosine kinase proteins are implicated in numerous human cancers. In this thesis, the structural consequences of several activating ERBB mutations reported in NSCLC and leukemia patients were explored: EGFR-ΔELREA, EGFR-V769insASV, EGFR-D770insNPG and ERBB2-E936K. The studies further shed light on the possible mechanistic basis for the experimentally reported increased kinase activity by the mutations. ERBB receptors as a whole and their component domains are highly dynamic biological molecules, therefore, when examining the effects somatic mutations exert on ERBB structures, it is critical to consider their dynamic nature. With this in mind, in this thesis, all-atom MDS was principally employed on the wild-type and mutant ERBB structures in order to attain the goals of the study.

The simulations on the wild-type and ΔELREA EGFR kinase domains pointed out several key findings. The ΔELREA deletion mutation stabilized kinase domain structural features critical to maintaining the active EGFR state. This included the stabilization of the αC helix in the "αC-in" conformation and conservation of the Lys745-Glu762 salt bridge. Moreover, the mutation resulted in a stronger binding of EGFR with the natural substrate, ATP. Intriguingly, the mutation also resulted in a structural change on the inactive conformation, an inward movement of the αC helix, a motion required to transition to the active EGFR state. Hence, the deletion mutation could steer a conformational shift from the inactive towards the active EGFR. In order to observe a full transition between the active and inactive EGFR conformations, however, long-run simulations in the range of tens of microseconds are required (Shan et al. 2013), which were not fulfilled by the simulations carried out for this study.

The observations from the ΔELREA simulations were in line with the effect of TKIs on ΔELREA EGFR lung cancer cell lines (See Figure 10 in publication I). While TKIs such as gefitinib, erlotinib and afatinib that recognize the EGFR active conformation were more effective against the deletion mutant form as compared to the wild-type EGFR, the inactive state recognizing TKI, lapatinib, did not have a significantly different effect between the ΔELREA and wild-type EGFRs. These results are also supported by a simulation study that reported better binding of TKIs, such as gefitinib and erlotinib, to the active state ΔELREA mutant as a result of a compact ligand binding site arising from a positionally stable αC helix (Kannan et al. 2017).

The study on the V769insASV and D770insNPG EGFR exon 20 insertion mutations interestingly conveyed some similar findings as the ΔELREA EGFR kinase mutation. The insertions stabilized the active state αC helix

in the "αC-in" conformation, maintained the Lys745-Glu762 interaction and preserved the integrity of the DFG motif and the R-spine, features that were compromised in the wild-type EGFR kinase simulation. The insertion mutants also disrupted the formation of an autoinhibitory interaction between Arg767 of the αC helix and Arg776 of the αC-β4 loop in the inactive EGFR state, which may lead to an equilibrium shift from the inactive towards the active EGFR state.

The majority of the exon 20 insertion mutations, including V769insASV and D770insNPG, are insensitive to first- and second-generation TKIs, such as gefitinib and erlotinib. The structural reasons behind the reduced TKI sensitivity are yet to be exhaustively explored. The apo form wild-type and insertion mutant simulations presented in this thesis indicate a likely role of the N-lobe of the kinase domain in the reported TKI resistance by the insertions. The N-lobe, which includes structural elements such as the P-loop and αC helix is observed shifting with respect to the C-lobe of the kinase domain in the mutant EGFR simulations (Figure 26). Consequently, the P-loop moves closer to the A-loop of the C-lobe, obstructing the ATP binding pocket. Other structural units such as the β2, β3 and β4 strands and the αC helix also make a positional shift. In contrast, the wild-type EGFR maintains the crystal structure conformation of the N-lobe, in which the ATP binding pocket is exposed. In the insertion mutant EGFRs, the occlusion of the binding pocket due to the move of the P-loop might deny TKIs access to the binding pocket and/or result in a steric hindrance to flexible TKIs, such as gefitinib. Furthermore, the positional shift of the β3 and β4 strands might affect the interactions the drugs make with EGFR, subsequently affecting binding of the drugs. It is noteworthy that a thorough examination on the TKI-bound wild-type and mutant ERBBs would be relevant and more revealing to understand the underlying reasons behind the TKI insensitivity by exon 20 insertion mutations.

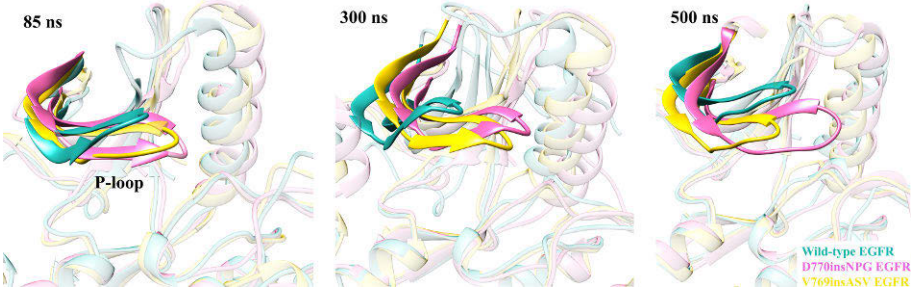


Figure 26. Structural change at the N-lobe of the kinase domain of wild-type, V769insASV and D770insNPG EGFRs. Superimposed sampled conformations from the wild-type and mutant EGFR simulations show a positional shift for the N-lobe structural units, most notably the P-loop, in the mutant forms. The P-loop of the insertion mutants is placed closer to the A-loop of the C-lobe. Figure from publication II.

The investigation on the E936K ERBB2 kinase domain mutation gave insights into how peripheral mutations are able to alter the kinase activity of ERBBs. The E936K mutation, located at C-lobe of the kinase domain led to stronger interactions at the dimer interface both between ERBB2-ERBB2 and ERBB2-EGFR dimers. As a result, the binding between the monomers is enhanced, lengthening the duration of the activated kinase state. Although the ERBB2 E936K mutation is the only peripheral mutation discussed in this thesis, other peripheral activating mutations located at the JM-B segment of the ERBBs have been examined in other studies listed in the 'additional publications' section of this thesis. These mutations include the EGFR A702V and the ERBB4 E715K mutations, which have been respectively observed in NSCLC (Reckamp et al. 2006) and skin cancer (Pickering et al. 2014) samples. Similar to the ERBB2 E936K mutation, these mutations also strengthened the interactions at the ERBB asymmetric dimer interface, which would extend the time of the activated kinase state. The identification of several of such peripheral activating mutations perhaps calls for the spotlight to also be directed to dimer interface ERBB mutations. Furthermore, these mutations could present an alternative approach to target ERBB activity in cancer. For instance, peptide inhibitors that bind to a mutant dimer interface might be ideal candidates to interfere with the asymmetric association of ERBB monomers and subsequent receptor signaling.

In the above studies, MDS was an integral tool to uncover the dynamic structural consequences of the activating ERBB mutations. Before performing MDS on these proteins, however, it is of great importance to use relevant starting structures in order to get a correct interpretation of what might be taking place in the biological system with respect to the mutations of interest. For instance, a mutation at the dimeric interface of ERBBs would be best examined by using the dimeric structure, not the monomeric, which would fail to capture the realistic structural impact of the mutation. Therefore, it is essential to first check the location of a mutation using available 3D structures and probe its likely effect relative to nearby structural elements. For instance, the ERBB2 E936K mutation discussed above is located at the C-terminal lobe of the kinase domain, which is at the interface between the activator and receiver kinases. Hence, the ERBB2 dimer structure was used in the study, as opposed to the monomer form.

The use of biologically relevant structures is also fundamental, which is not always properly represented by the 3D coordinate files accessed from the PDB. For example, the ERBB4 dimer structure (PDB ID: 3BCE) shows an ERBB4 trimer placed near each other and making few contacts, which might be mistaken for the asymmetrically arranged dimer. By using the assembly information with in the PDB file, the biologically significant

asymmetric ERBB4 dimer can be generated with visualization programs such as Chimera.

The carefully selected structures need to be properly prepared, minimized, heated and equilibrated before running the production simulations. The input parameters used in these processes can be adjusted as needed, such as the number of minimization cycles, the temperature set for the systems and the length of equilibration and simulation runs. The time period for a simulation can be set according to the objective of the work. For instance, in order to probe the dynamics of intramolecular and intermolecular interactions, it may be sufficient to simulate a system for tens of nanoseconds. However, simulations in the range of micro- to milliseconds may be needed to observe large conformational changes, such as a transition between active-inactive ERBB kinase states.

Another essential practice with MDS is employing independent replicate simulations, which would aid to explore a wide conformational space. The duplicate simulations of the wild-type and insertion mutant EGFRs are good examples to demonstrate the importance of performing replicate experiments. As discussed before, in the wild-type EGFR simulation the change in the $\chi 1$ dihedral angle of Phe856 of the DFG motif from the active to the inactive kinase orientation was sampled only in one of the duplicate simulations. Hence, conducting a single simulation would have minimized the possibility of observing this structural change and altered the conclusions made from it. Although MDS can be a powerful method to examine molecular systems – in this case the structural effects of mutations – it is beneficial to carefully evaluate the computational results and whenever possible to combine the computational studies with experimental studies for better interpretation of the findings and to benefit from the synergy of the combined approaches.

7 Conclusion

This thesis aimed to elucidate the structural effects of four activating ERBB mutations observed frequently in cancer patients and identify the mechanisms by which the mutations increase kinase activity. The mutations include EGFR-ΔELREA, EGFR-V769insASV, EGFR-D770insNPG and ERBB2-E936K, the former three being well-established driver mutations in NSCLC. Based on the study on the above four ERBB mutations these conclusions can be drawn:

1. The ΔELREA EGFR deletion mutation shortened the length of the β3-αC loop, which consequently constrains the neighboring αC helix. In the active EGFR state, the mutation stabilized the residues of the αC helix and ensured the formation of the Lys745-Glu762 interaction that is fundamental to kinase catalytic activity. The deletion also resulted in frequent interactions between ATP and EGFR, which led to stronger binding between the nucleotide and EGFR. Moreover, in the inactive EGFR, ΔELREA resulted in an inwards motion of the αC helix and displacement of the A-loop helix. These observations collectively suggest that the ΔELREA deletion mutation exerts its activating role by stabilizing the active EGFR conformation and by supporting a conformational shift from the inactive towards the active EGFR state.

2. The V769insASV and D770insNPG EGFR insertion mutations altered the EGFR structure at the C-terminus of the αC helix and beginning of the αC-β4 loop. In the active EGFR state, the insertions imparted structural and positional stability to the αC helix. The mutations maintained the Lys745-Glu762 salt bridge, the DFG motif and the R-spine. In the inactive EGFR conformation, the mutations compromised the formation of an autoinhibitory interaction between Ala767 and Arg776, which is key for the stability of the inactive state of EGFR. These findings propose a dual mechanism for the activating effect of the two exon 20 insertion mutations: by stabilizing structural features key for the active state EGFR and by advocating a shift in equilibrium from the inactive towards the active EGFR conformation.

3. The E936K ERBB2 mutation led to more frequent hydrogen bonding and newly formed ionic interactions at the ERBB2-EGFR heterodimer. Due to the enhanced interactions at the dimer interface, the E936K mutation resulted in a strong binding between the activator (ERBB2) and receiver (EGFR) kinases, which prolongs the time of the activated ERBB dimer. This

observation is likely to be reflected for the ERBB2-ERBB2 homodimer.

Overall, the studies have unveiled that concerted atomic and structural alterations govern the functional changes by the oncogenic ERBB mutations. The findings obtained from the structural characterization of activating ERBB mutations would not only advance our understanding of the nature of mutations that play a crucial role in the development of cancer, but also could provide therapeutic tools aimed at controlling ERBB enzymatic activity and signaling. This may be attained for instance through the selection of ligands that target stabilization and conformational effects by the mutations, and/or through the rational design of agents that exploit the unique structural features, such as interactions, arising from the mutations.

8 References

Aertgeerts K, Skene R, Yano J, Sang BC, Zou H, Snell G, et al. Structural analysis of the mechanism of inhibition and allosteric activation of the kinase domain of HER2 protein. J Biol Chem. 2011;286(21):18756–65.

Agus DB, Akita RW, Fox WD, Lewis GD, Higgins B, Pisacane PI, et al. Targeting ligand-activated ErbB2 signaling inhibits breast and prostate tumor growth. Cancer Cell. 2002;2(2):127–37.

Arcaro A, Guerreiro A. The Phosphoinositide 3-Kinase Pathway in Human Cancer: Genetic Alterations and Therapeutic Implications. Curr Genomics. 2007;8(5):271–306.

Arcila ME, Chaft JE, Nafa K, Roy-Chowdhuri S, Lau C, Zaidinski M, et al. Prevalence, clinicopathologic associations, and molecular spectrum of ERBB2 (HER2) tyrosine kinase mutations in lung adenocarcinomas. Clin Cancer Res. 2012;18(18):4910–8.

Arcila ME, Nafa K, Chaft JE, Rekhtman N, Lau C, Reva BA, et al. EGFR exon 20 insertion mutations in lung adenocarcinomas: Prevalence, molecular heterogeneity, and clinicopathologic characteristics. Mol Cancer Ther. 2013;12(2):220–9.

Arteaga CL, Engelman JA. ERBB receptors: From oncogene discovery to basic science to mechanism-based cancer therapeutics. Cancer Cell. 2014;25(5):282–303.

Barretina J, Caponigro G, Stransky N, Venkatesan K, Margolin AA, Kim S, et al. The Cancer Cell Line Encyclopedia enables predictive modelling of anticancer drug sensitivity. Nature. 2012;483(7391):603–7.

Bateman A. UniProt: A worldwide hub of protein knowledge. Nucleic Acids Res. 2019;47(D1):D506–15.

Beenstock J, Mooshayef N, Engelberg D. How Do Protein Kinases Take a Selfie (Autophosphorylate)? Trends in Biochem Sci. 2016;41(11):938–53.

Berman HM, Westbrook J, Feng Z, Gilliland G, Bhat TN, Weissig H, et al. The Protein Data Bank. Nucleic Acids Res. 2000;28(1):235–42.

Bocharov E V., Lesovoy DM, Pavlov K V., Pustovalova YE, Bocharova O V., Arseniev AS. Alternative packing of EGFR transmembrane domain suggests that protein-lipid interactions underlie signal conduction across membrane. Biochim Biophys Acta - Biomembr. 2016;1858(6):1254–61.

Bose R, Kavuri SM, Searleman AC, Shen W, Shen D, Koboldt DC, et al. Activating HER2 mutations in HER2 gene amplification negative breast cancer. Cancer Discov. 2013;3(2):224–37.

Bousoik E, Montazeri Aliabadi H. "Do We Know Jack" About JAK? A Closer Look at JAK/STAT Signaling Pathway. Front Oncol. 2018;8:287.

Bouyain S, Longo PA, Li S, Ferguson KM, Leahy DJ. The extracellular region of ErbB4 adopts a tethered conformation in the absence of ligand. Proc Natl Acad Sci U S A. 2005;102(42):15024–9.

Bragin PE, Mineev KS, Bocharova O V., Volynsky PE, Bocharov E V., Arseniev AS. HER2 Transmembrane Domain Dimerization Coupled with Self-Association of Membrane-Embedded Cytoplasmic Juxtamembrane Regions. J Mol Biol. 2016;428(1):52–61.

Britsch S. The neuregulin-I/ErbB signaling system in development and disease. Adv Anat Embryol Cell Biol. 2007;190:1–65.

Burke CL, Stern DF. Activation of Neu (ErbB-2) Mediated by Disulfide Bond-Induced Dimerization Reveals a Receptor Tyrosine Kinase Dimer Interface. Mol Cell Biol. 1998;18(9):5371–9.

Carpenter G, King L, Cohen S. Epidermal growth factor stimulates phosphorylation in membrane preparations in vitro. Nature. 1978;276(5686):409–10.

Carpenter G, Lembach KJ, Morrison MM, Cohen S. Characterization of the binding of 125I labeled epidermal growth factor to human fibroblasts. J Biol Chem. 1975;250(11):4297–304.

Cho HS, Leahy DJ. Structure of the extracellular region of HER3 reveals an interdomain tether. Science. 2002;297(5585):1330–3.

Cho HS, Mason K, Ramyar KX, Stanley AM, Gabelli SB, Denney DW, et al. Structure of the extracellular region of HER2 alone and in complex with the Herceptin Fab. Nature. 2003;421(6924):756–60.

Citri A, Skaria KB, Yarden Y. The deaf and the dumb: The biology of ErbB-2 and ErbB-3. Exp Cell Res. 2003;284(1):54–65.

Citri A, Yarden Y. EGF-ERBB signalling: Towards the systems level. Nat Rev Mol Cell Biol. 2006;7:505–16.

Clark PA, Iida M, Treisman DM, Kalluri H, Ezhilan S, Zorniak M, et al. Activation of multiple ERBB family receptors mediates glioblastoma cancer stem-like cell resistance to EGFR-targeted inhibition. Neoplasia. 2012;14(5):420–8.

Cohen S. The stimulation of epidermal proliferation by a specific protein (EGF). Dev Biol. 1965;12(3):394–407.

Cohen S. Isolation of a mouse submaxillary gland protein accelerating incisor eruption and eyelid opening in the new-born animal. J Biol Chem. 1962;237(5):1555–62.

Collisson EA, Campbell JD, Brooks AN, Berger AH, Lee W, Chmielecki J, et al. Comprehensive molecular profiling of lung adenocarcinoma: The cancer genome atlas research network. Nature. 2014;511(7511):543–50.

Cross DAE, Ashton SE, Ghiorghiu S, Eberlein C, Nebhan CA, Spitzler PJ, et al. AZD9291, an irreversible EGFR TKI, overcomes T790M-mediated resistance to EGFR inhibitors in lung cancer. Cancer Discov. 2014;4(9):1046–61.

Van Cutsem E, Köhne C-H, Hitre E, Zaluski J, Chang Chien C-R, Makhson A, et al. Cetuximab and Chemotherapy as Initial Treatment for Metastatic Colorectal Cancer. N Engl J Med. 2009;360(14):1408–17.

Van Cutsem E, Peeters M, Siena S, Humblet Y, Hendlisz A, Neyns B, et al. Open-label phase III trial of panitumumab plus best supportive care compared with best supportive care alone in patients with chemotherapy- refractory metastatic colorectal cancer. J Clin Oncol. 2007;25(13):1658–64.

Cymer F, Schneider D. Transmembrane helix-helix interactions involved in ErbB receptor signaling. Cell Adh Migr. 2010;4(2):299–312.

Dawson JP, Bu Z, Lemmon MA. Ligand-Induced Structural Transitions in ErbB Receptor Extracellular Domains. Structure. 2007;15(8):942–54.

Degirmenci U, Wang M, Hu J. Targeting Aberrant RAS/RAF/MEK/ERK Signaling for Cancer Therapy. Cells. 2020;9(1):198.

Ding L, Getz G, Wheeler DA, Mardis ER, McLellan MD, Cibulskis K, et al. Somatic mutations affect key pathways in lung adenocarcinoma. Nature. 2008;455(7216):1069–75.

Downward J, Yarden Y, Mayes E, Scrace G, Totty N, Stockwell P, et al. Close similarity of epidermal growth factor receptor and v-erb-B

oncogene protein sequences. Nature. 1984;307(5951):521–7.

Eck MJ, Yun CH. Structural and mechanistic underpinnings of the differential drug sensitivity of EGFR mutations in non-small cell lung cancer. Biochim Biophys Acta. 2010;1804(3):559–66.

Ellis MJ, Ding L, Shen D, Luo J, Suman VJ, Wallis JW, et al. Whole-genome analysis informs breast cancer response to aromatase inhibition. Nature. 2012;486(7403):353–60.

Endres NF, Das R, Smith AW, Arkhipov A, Kovacs E, Huang Y, et al. Conformational coupling across the plasma membrane in activation of the EGF receptor. Cell. 2013;152(3):543–56.

Engelman JA, Zejnullahu K, Gale CM, Lifshits E, Gonzales AJ, Shimamura T, et al. PF00299804, an irreversible pan-ERBB inhibitor, is effective in lung cancer models with EGFR and ERBB2 mutations that are resistant to gefitinib. Cancer Res. 2007;67(24):11924–32.

Essmann U, Perera L, Berkowitz ML, Darden T, Lee H, Pedersen LG. A smooth particle mesh Ewald method. J Chem Phys. 1995;103(19):8577–93.

Ferguson KM, Berger MB, Mendrola JM, Cho HS, Leahy DJ, Lemmon MA. EGF activates its receptor by removing interactions that autoinhibit ectodomain dimerization. Mol Cell. 2003;11(2):507–17.

Ferreira PMP, Pessoa C. Molecular biology of human epidermal receptors, signaling pathways and targeted therapy against cancers: New evidences and old challenges. Brazilian J Pharm Sci. 2017;53(2):e16076.

Foster SA, Whalen DM, Özen A, Wongchenko MJ, Yin J, Yen I, et al. Activation Mechanism of Oncogenic Deletion Mutations in BRAF, EGFR, and HER2. Cancer Cell. 2016;29(4):477–93.

Franklin MC, Carey KD, Vajdos FF, Leahy DJ, De Vos AM, Sliwkowski MX. Insights into ErbB signaling from the structure of the ErbB2-pertuzumab complex. Cancer Cell. 2004;5(4):317–28.

Fry DW, Kraker AJ, McMichael A, Ambroso LA, Nelson JM, Leopold WR, et al. A specific inhibitor of the epidermal growth factor receptor tyrosine kinase. Science. 1994;265(5175):1093–5.

Fuller SJ, Sivarajah K, Sugden PH. ErbB receptors, their ligands, and the consequences of their activation and inhibition in the myocardium. J Mol Cell Cardiol. 2008;44(5):831–54.

Furuyama K, Harada T, Iwama E, Shiraishi Y, Okamura K, Ijichi K, et al. Sensitivity and kinase activity of epidermal growth factor receptor (EGFR) exon 19 and others to EGFR-tyrosine kinase inhibitors. Cancer Sci. 2013;104(5):584–9.

Gajiwala KS. EGFR: Tale of the C-terminal tail. Protein Sci. 2013;22(7):995–9.

Gajiwala KS, Feng J, Ferre R, Ryan K, Brodsky O, Weinrich S, et al. Insights into the Aberrant Activity of Mutant EGFR Kinase Domain and Drug Recognition. Structure. 2013;21(2):209–19.

Garrett TPJ, McKern NM, Lou M, Elleman TC, Adams TE, Lovrecz GO, et al. The crystal structure of a truncated ErbB2 ectodomain reveals an active conformation, poised to interact with other ErbB receptors. Mol Cell. 2003;11(2):495–505.

Gassmann M, Casagranda F, Orloli D, Simon H, Lai C, Kleint R, et al. Aberrant neural and cardiac development in mice lacking the ErbB4 neuregulin receptor. Nature. 1995;378(6555):390–4.

Gazdar AF. Activating and resistance mutations of EGFR in non-small-cell lung cancer: Role in clinical response to EGFR tyrosine kinase

inhibitors. Oncogene. 2009;28:S24–31.

Ghosh R, Narasanna A, Wang SE, Liu S, Chakrabarty A, Balko JM, et al. Trastuzumab has preferential activity against breast cancers driven by HER2 homodimers. Cancer Res. 2011;71(5):1871–82.

Gilmore JL, Gallo RM, Riese DJ. The epidermal growth factor receptor (EGFR)-S442F mutant displays increased affinity for neuregulin-2β and agonist-independent coupling with downstream signalling events. Biochem J. 2006;396(1):79–88.

Giri DK, Ali-Seyed M, Li L-Y, Lee D-F, Ling P, Bartholomeusz G, et al. Endosomal Transport of ErbB-2: Mechanism for Nuclear Entry of the Cell Surface Receptor. Mol Cell Biol. 2005;25(24):11005–18.

Goldstein NI, Prewett M, Zuklys K, Rockwell P, Mendelsohn J. Biological efficacy of a chimeric antibody to the epidermal growth factor receptor in a human tumor xenograft model. Clin Cancer Res. 1995;1(11):1311-8

Goss G, Tsai CM, Shepherd FA, Bazhenova L, Lee JS, Chang GC, et al. Osimertinib for pretreated EGFR Thr790Met-positive advanced non-small-cell lung cancer (AURA2): a multicentre, open-label, single-arm, phase 2 study. Lancet Oncol. 2016;17(12):1643–52.

Gridelli C, De Marinis F, Di Maio M, Cortinovis D, Cappuzzo F, Mok T. Gefitinib as first-line treatment for patients with advanced non-small-cell lung cancer with activating epidermal growth factor receptor mutation: Review of the evidence. Lung Cancer. 2010;71:249–57.

Guha U, Chaerkady R, Marimuthu A, Patterson AS, Kashyap MK, Harsha HC, et al. Comparisons of tyrosine phosphorylated proteins in cells expressing lung cancer-specific alleles of EGFR and KRAS. Proc Natl Acad Sci U S A. 2008;105(37):14112–7.

de Gunst MM, Gallegos-Ruiz MI, Giaccone G, Rodriguez JA. Functional analysis of cancer-associated EGFR mutants using a cellular assay with YFP-tagged EGFR intracellular domain. Mol Cancer. 2007;6:56.

Guo Y, Pan W, Liu S, Shen Z, Xu Y, Hu L. ERK/MAPK signalling pathway and tumorigenesis. Exp Ther Med. 2020;19(3):1997-2007

Hanker AB, Brewer MR, Sheehan JH, Koch JP, Sliwoski GR, Nagy R, et al. An acquired HER2T798I gatekeeper mutation induces resistance to neratinib in a patient with HER2 mutant-driven breast cancer. Cancer Discov. 2017;7(6):575–85.

Harari D, Yarden Y. Molecular mechanisms underlying ErbB2/HER2 action in breast cancer. Oncogene. 2000;19(53):6102–14.

Herter-Sprie GS, Greulich H, Wong KK. Activating mutations in ERBB2 and their impact on diagnostics and treatment. Front Oncol. 2013;3:86.

Ho J, Moyes DL, Tavassoli M, Naglik JR. The Role of ErbB Receptors in Infection. Trends in Microbiol. 2017;25(11):942–52.

Humphrey W, Dalke A, Schulten K. VMD: Visual molecular dynamics. J Mol Graph. 1996;14(1):33–8.

Huse M, Kuriyan J. The Conformational Plasticity of Protein Kinases. Cell. 2002;109(3):275–82.

Hynes NE, Watson CJ. Mammary gland growth factors: roles in normal development and in cancer. Cold Spring Harb Perspect Biol. 2010;2(8):a003186.

Jaiswal BS, Kljavin NM, Stawiski EW, Chan E, Parikh C, Durinck S, et al. Oncogenic ERBB3 Mutations in Human Cancers. Cancer Cell. 2013;23(5):603–17.

Johnston SRD, Leary A. Lapatinib: A novel EGFR/HER2 tyrosine kinase inhibitor for cancer. Drugs Today (Barc). 2006;42(7):441–53.

Jorgensen WL, Chandrasekhar J, Madura JD, Impey RW, Klein ML. Comparison of simple potential functions for simulating liquid water. J Chem Phys. 1983;79(2):926–35.

Jura N, Endres NF, Engel K, Deindl S, Das R, Lamers MH, et al. Mechanism for Activation of the EGF Receptor Catalytic Domain by the Juxtamembrane Segment. Cell. 2009a;137(7):1293–307.

Jura N, Shan Y, Cao X, Shaw DE, Kuriyan J. Structural analysis of the catalytically inactive kinase domain of the human EGF receptor 3. Proc Natl Acad Sci U S A. 2009b;106(51):21608–13.

Jura N, Zhang X, Endres NF, Seeliger MA, Schindler T, Kuriyan J. Catalytic Control in the EGF Receptor and Its Connection to General Kinase Regulatory Mechanisms. Mol Cell. 2011;42(1):9–22.

Kabsch W, Sander C. Dictionary of protein secondary structure: Pattern recognition of hydrogen-bonded and geometrical features. Biopolymers. 1983;22(12):2577–637.

Kannan S, Pradhan MR, Tiwari G, Tan W-C, Chowbay B, Tan EH, et al. Hydration effects on the efficacy of the Epidermal growth factor receptor kinase inhibitor afatinib. Sci Rep. 2017;7(1):1540.

Kornev AP, Haste NM, Taylor SS, Ten Eyck LF. Surface comparison of active and inactive protein kinases identifies a conserved activation mechanism. Proc Natl Acad Sci U S A. 2006;103(47):17783–8.

Kovacs E, Das R, Wang Q, Collier TS, Cantor A, Huang Y, et al. Analysis of the Role of the C-Terminal Tail in the Regulation of the Epidermal Growth Factor Receptor. Mol Cell Biol. 2015;35(17):3083–102.

Kraus MH, Issing W, Miki T, Popescu NP, Aaronson SA. Isolation and characterization of ERBB3, a third member of the ERBB/epidermal growth factor receptor family: Evidence for overexpression in a subset of human mammary tumors. Proc Natl Acad Sci U S A. 1989;86(23):9193–7.

Kurppa KJ, Denessiouk K, Johnson MS, Elenius K. Activating ERBB4 mutations in non-small cell lung cancer. Oncogene. 2016;35(10):1283–91.

Leach A. Molecular Modelling: Principles and Applications, 2nd Edition. 2001.

Lau C, Killian KJ, Samuels Y, Rudloff U. ERBB4 mutation analysis: Emerging molecular target for melanoma treatment. Methods Mol Biol. 2014;1102:461–80.

Leahy DJ. Structure and Function of the Epidermal Growth Factor (EGF/ErbB) Family of Receptors. Adv Protein Chem. 2004;68:1–27.

Lee JW, Soung YH, Seo SH, Kim SY, Park CH, Wang YP, et al. Somatic mutations of ERBB2 kinase domain in gastric, colorectal, and breast carcinomas. Clin Cancer Res. 2006;12(1):57–61.

Lee KF, Simon H, Chen H, Bates B, Hung MC, Hauser C. Requirement for neuregulin receptor erbB2 in neural and cardiac development. Nature. 1995;378(6555):394–8.

Li D, Ambrogio L, Shimamura T, Kubo S, Takahashi M, Chirieac LR, et al. BIBW2992, an irreversible EGFR/HER2 inhibitor highly effective in preclinical lung cancer models. Oncogene. 2008;27(34):4702–11.

Linggi B, Carpenter G. ErbB receptors: new insights on mechanisms and biology. Trends Cell Biol. 2006;16:649–56.

Lipsick J. A history of cancer research: Tyrosine kinases. Cold Spring Harb Perspect Biol. 2019;11(2):a035592.

Liu P, Cleveland TE, Bouyain S, Byrne PO, Longo PA, Leahy DJ. A single ligand is sufficient to activate EGFR

dimers. Proc Natl Acad Sci U S A. 2012;109(27):10861–6.

Lo HW, Ali-Seyed M, Wu Y, Bartholomeusz G, Hsu SC, Hung MC. Nuclear-cytoplasmic transport of EGFR involves receptor endocytosis, importin β1 and CRM1. J Cell Biochem. 2006;98(6):1570–83.

Lu C, Mi LZ, Schürpf T, Walz T, Springer TA. Mechanisms for kinase-mediated dimerization of the epidermal growth factor receptor. J Biol Chem. 2012;287(45):38244–53.

Lynch TJ, Bell DW, Sordella R, Gurubhagavatula S, Okimoto RA, Brannigan BW, et al. Activating Mutations in the Epidermal Growth Factor Receptor Underlying Responsiveness of Non–Small-Cell Lung Cancer to Gefitinib. N Engl J Med. 2004;350(21):2129–39.

MacKenzie KR. Folding and stability of α-helical integral membrane proteins. Chem Rev. 2006;106(5):1931–77.

Maier JA, Martinez C, Kasavajhala K, Wickstrom L, Hauser KE, Simmerling C. ff14SB: Improving the Accuracy of Protein Side Chain and Backbone Parameters from ff99SB. J Chem Theory Comput. 2015;11(8):3696–713.

Mazières J, Peters S, Lepage B, Cortot AB, Barlesi F, Beau-Faller M, et al. Lung cancer that harbors an HER2 Mutation: Epidemiologic characteristics and therapeutic perspectives. J Clin Oncol. 2013;31(16):1997–2003.

Meagher KL, Redman LT, Carlson HA. Development of polyphosphate parameters for use with the AMBER force field. J Comput Chem. 2003;24(9):1016–25.

Mei L, Nave KA. Neuregulin-ERBB signaling in the nervous system and neuropsychiatric diseases. Neuron. 2014;83(1):27–49.

Miettinen PJ, Berger JE, Meneses J, Phung Y, Pedersen RA, Werb Z, et al. Epithelial immaturity and multiorgan failure in mice lacking epidermal growth factor receptor. Nature. 1995;376(6538):337–41.

Miller BR, McGee TD, Swails JM, Homeyer N, Gohlke H, Roitberg AE. MMPBSA.py : An Efficient Program for End-State Free Energy Calculations. J Chem Theory Comput. 2012;8(9):3314–21.

Mishra R, Hanker AB, Garrett JT. Genomic alterations of ERBB receptors in cancer: Clinical implications. Oncotarget. 2017;8(69):114371–92.

Molina MA, Codony-Servat J, Albanell J, Rojo F, Arribas J, Baselga J. Trastuzumab (Herceptin), a humanized anti-HER2 receptor monoclonal antibody, inhibits basal and activated HER2 ectodomain cleavage in breast cancer cells. Cancer Res. 2001;61(12):4744–9.

Murray S, Dahabreh IJ, Linardou H, Manoloukos M, Bafaloukos D, Kosmidis P. Somatic mutations of the tyrosine kinase domain of epidermal growth factor receptor and tyrosine kinase inhibitor response to TKIs in non-small cell lung cancer: an analytical database. J Thorac Oncol. 2008;3(8):832–9.

Niederst MJ, Hu H, Mulvey HE, Lockerman EL, Garcia AR, Piotrowska Z, et al. The allelic context of the C797S mutation acquired upon treatment with third-generation EGFR inhibitors impacts sensitivity to subsequent treatment strategies. Clin Cancer Res. 2015;21(17):3924–33.

O'Connor CM, Adams JU. Essentials of Cell Biology. 2010.

Oxnard GR, Lo PC, Nishino M, Dahlberg SE, Lindeman NI, Butaney M, et al. Natural history and molecular characteristics of lung cancers harboring egfr exon 20 insertions. J Thorac Oncol. 2013;8(2):179–84.

Paez JG, Jänne PA, Lee JC, Tracy S, Greulich H, Gabriel S, et al. EGFR Mutations in Lung Cancer: Correlation with Clinical Response to Gefitinib Therapy. Science. 2004;304(5676):1497–500.

Pao W, Miller V, Zakowski M, Doherty J, Politi K, Sarkaria I, et al. EGF receptor gene mutations are common in lung cancers from "never smokers" and are associated with sensitivity of tumors to gefitinib and erlotinib. Proc Natl Acad Sci U S A. 2004;101(36):13306–11.

Pao W, Miller VA, Politi KA, Riely GJ, Somwar R, Zakowski MF, et al. Acquired Resistance of Lung Adenocarcinomas to Gefitinib or Erlotinib Is Associated with a Second Mutation in the EGFR Kinase Domain. Liu ET, editor. PLoS Med. 2005;2(3):e73.

Park JH, Liu Y, Lemmon MA, Radhakrishnan R. Erlotinib binds both inactive and active conformations of the EGFR tyrosine kinase domain. Biochem J. 2012;448(3):417–23.

Park K, Tan EH, O'Byrne K, Zhang L, Boyer M, Mok T, et al. Afatinib versus gefitinib as first-line treatment of patients with EGFR mutation-positive non-small-cell lung cancer (LUX-Lung 7): A phase 2B, open-label, randomised controlled trial. Lancet Oncol. 2016;17(5):577–89.

Pedersen MW, Meltorn M, Damstrup L, Poulsen HS. The type III epidermal growth factor receptor mutation. Biological significance and potential target for anti-cancer therapy. Ann Oncol. 2001;12(6):745–60.

Penington DJ, Bryant I, Riese DJ. Constitutively active ErbB4 and ErbB2 mutants exhibit distinct biological activities. Cell Growth Differ. 2002;13(6):247–56.

Pettersen EF, Goddard TD, Huang CC, Couch GS, Greenblatt DM, Meng EC, et al. UCSF Chimera--a visualization system for exploratory research and analysis. J Comput Chem. 2004;25(13):1605–12.

Pickering CR, Zhou JH, Lee JJ, Drummond JA, Peng SA, Saade RE, et al. Mutational landscape of aggressive cutaneous squamous cell carcinoma. Clin Cancer Res. 2014;20(24):6582–92.

Pines G, Huang PH, Zwang Y, White FM, Yarden Y. EGFRvIV: A previously uncharacterized oncogenic mutant reveals a kinase autoinhibitory mechanism. Oncogene. 2010;29(43):5850–60.

Plowman GD, Culouscou JM, Whitney GS, Green JM, Carlton GW, Foy L, et al. Ligand-specific activation of HER4/p180erbB4, a fourth member of the epidermal growth factor receptor family. Proc Natl Acad Sci U S A. 1993;90(5):1746–50.

Pon JR, Marra MA. Driver and passenger mutations in cancer. Annu Rev Pathol Mech Dis. 2015;10:25–50.

Prickett TD, Agrawal NS, Wei X, Yates KE, Lin JC, Wunderlich JR, et al. Analysis of the tyrosine kinome in melanoma reveals recurrent mutations in ERBB4. Nat Genet. 2009;41(10):1127–32.

Qiu C, Tarrant MK, Choi SH, Sathyamurthy A, Bose R, Banjade S, et al. Mechanism of Activation and Inhibition of the HER4/ErbB4 Kinase. Structure. 2008;16(3):460–7.

Del Re M, Cucchiara F, Petrini I, Fogli S, Passaro A, Crucitta S, et al. ErbB in NSCLC as a molecular target: Current evidences and future directions. ESMO Open. 2020;5(4):e000724.

Reckamp KL, Krysan K, Morrow JD, Milne GL, Newman RA, Tucker C, et al. A phase I trial to determine the optimal biological dose of celecoxib when combined with erlotinib in advanced non-small cell lung cancer. Clin Cancer Res. 2006;12(11 I):3381–8.

Reguart N, Cardona AF, Rosell R. Role of erlotinib in first-line and maintenance treatment of advanced non-small-cell lung cancer. Cancer Manag Res. 2010;2:143–56.

Riese DJ, Gallo RM, Settleman J. Mutational activation of ErbB family receptor tyrosine kinases: Insights into mechanisms of signal transduction and tumorigenesis. BioEssays. 2007;29(6):558–65.

Riethmacher D, Sonnenberg-Riethmacher E, Brinkmann V, Yamaai T, Lewin GR, Birchmeier C. Severe neuropathies in mice with targeted mutations in the ErbB3 receptor. Nature. 1997;389(6652):725–30.

Robichaux JP, Elamin YY, Tan Z, Carter BW, Zhang S, Liu S, et al. Mechanisms and clinical activity of an EGFR and HER2 exon 20-selective kinase inhibitor in non-small cell lung cancer. Nat Med. 2018;24(5):638–46.

Robinson DR, Wu YM, Lin SF. The protein tyrosine kinase family of the human genome. Oncogene. 2000;19(49):5548–57.

Roe DR, Cheatham TE. PTRAJ and CPPTRAJ: Software for processing and analysis of molecular dynamics trajectory data. J Chem Theory Comput. 2013;9(7):3084–95.

Roskoski R. ErbB/HER protein-tyrosine kinases: Structures and small molecule inhibitors. Pharmacol Res. 2014a;87:42–59.

Roskoski R. The ErbB/HER family of protein-tyrosine kinases and cancer. Pharmacol Res. 2014b;79:34–74.

Ruan Z, Kannan N. Mechanistic Insights into R776H Mediated Activation of Epidermal Growth Factor Receptor Kinase. Biochemistry. 2015;54(27):4216–25.

Šali A, Blundell TL. Comparative Protein Modelling by Satisfaction of Spatial Restraints. J Mol Biol. 1993;234(3):779–815.

Sanchez-Soria P, Camenisch TD. ErbB signaling in cardiac development and disease. Semin Cell Dev Biol. 2010;21(9):929–35.

Sanchez-Vega F, Mina M, Armenia J, Ciriello G, Sander C, Schultz N, et al. Oncogenic Signaling Pathways in The Cancer Genome Atlas Article Oncogenic Signaling Pathways in The Cancer Genome Atlas. Cell. 2018;173(2):321-337.e10.

Schechter AL, Stern DF, Vaidyanathan L, Decker SJ, Drebin JA, Greene MI, et al. The neu oncogene: An erb-B-related gene encoding a 185,000-Mr tumour antigen. Nature. 1984;312(5994):513–6.

Schlessinger J. Cell signaling by receptor tyrosine kinases. Cell. 2000;103(2):211–25.

Schneider MR, Werner S, Paus R, Wolf E. Beyond wavy hairs: The epidermal growth factor receptor and its ligands in skin biology and pathology. Am J Pathol. 2008;173(1):14–24.

Ségaliny AI, Tellez-Gabriel M, Heymann MF, Heymann D. Receptor tyrosine kinases: Characterisation, mechanism of action and therapeutic interests for bone cancers. J Bone Oncol. 2015;4(1):1–12.

Shah SP, Morin RD, Khattra J, Prentice L, Pugh T, Burleigh A, et al. Mutational evolution in a lobular breast tumour profiled at single nucleotide resolution. Nature. 2009;461(7265):809–13.

Shan Y, Arkhipov A, Kim ET, Pan AC, Shaw DE. Transitions to catalytically inactive conformations in EGFR kinase. Proc Natl Acad Sci U S A. 2013;110(18):7270–5.

Shi F, Telesco SE, Liu Y, Radhakrishnan R, Lemmona MA. ErbB3/HER3 intracellular domain is competent to bind ATP and catalyze autophosphorylation. Proc Natl Acad Sci U S A. 2010;107(17):7692–7.

Shigematsu H, Gazdar AF. Somatic mutations of epidermal growth factor receptor signaling pathway in lung cancers. Int J Cancer; 2006;118(2):257–62.

Singh V, Feldman R, Sukari A, Kim C, Mamdani H, Spira AI, et al. Characterization of ERBB2 alterations in non-small cell lung cancer. J Clin Oncol. 2020;38(15_suppl):e21553–e21553.

Slamon DJ, Clark GM, Wong SG, Levin WJ, Ullrich A, McGuire WL. Human breast cancer: Correlation of relapse and survival with amplification of the HER-2/neu oncogene. Science. 1987;235(4785):182–91.

Sogabe S, Kawakita Y, Igaki S, Iwata H, Miki H, Cary DR, et al. Structure-based approach for the discovery of pyrrolo[3,2- d]pyrimidine-based EGFR T790M/L858R mutant inhibitors. ACS Med Chem Lett. 2013;4(2):201–5.

Solca F, Dahl G, Zoephel A, Bader G, Sanderson M, Klein C, et al. Target binding properties and cellular activity of afatinib (BIBW 2992), an irreversible ErbB family blocker. J Pharmacol Exp Ther. 2012;343(2):342–50.

Sordella R, Bell DW, Haber DA, Settleman J. Gefitinib-Sensitizing EGFR Mutations in Lung Cancer Activate Anti-Apoptotic Pathways. Science. 2004;305(5687):1163–7.

Soung YH, Lee JW, Kim SY, Wang YP, Jo KH, Moon SW, et al. Somatic mutations of the ERBB4 kinase domain in human cancers. Int J Cancer. 2006;118(6):1426–9.

Stamos J, Sliwkowski MX, Eigenbrot C. Structure of the epidermal growth factor receptor kinase domain alone and in complex with a 4-anilinoquinazoline inhibitor. J Biol Chem. 2002;277(48):46265–72.

Su J, Zhong W, Zhang X, Huang Y, Yan H, Yang J, et al. Molecular characteristics and clinical outcomes of EGFR exon 19 indel subtypes to EGFR TKIs in NSCLC patients. Oncotarget. 2017;8(67):111246–57.

Sweeney SM, Cerami E, Baras A, Pugh TJ, Schultz N, Stricker T, et al. AACR project genie: Powering precision medicine through an international consortium. Cancer Discov. 2017;7(8):818–31.

Tan CS, Kumarakulasinghe NB, Huang YQ, Ang YLE, Choo JRE, Goh BC, et al. Third generation EGFR TKIs: Current data and future directions. Mol Cancer. 2018;17:1–14.

Tanaka K, Nosaki K, Otsubo K, Azuma K, Sakata S, Ouchi H, et al. Acquisition of the T790M resistance mutation during afatinib treatment in EGFR tyrosine kinase inhibitor-naïve patients with non-small cell lung cancer harboring EGFR mutations. Oncotarget. 2017;8(40):68123–30.

Taylor SS, Kornev AP. Protein kinases: Evolution of dynamic regulatory proteins. Trends Biochem Sci; 2011;36(2):65–77.

Tvorogov D, Sundvall M, Kurppa K, Hollmén M, Repo S, Johnson MS, et al. Somatic mutations of ErbB4: Selective loss-of-function phenotype affecting signal transduction pathways in cancer. J Biol Chem. 2009;284(9):5582–91.

Uings IJ, Farrow SN. Cell receptors and cell signalling. Mol Pathol. 2000;53(6):295–9.

Ullrich A, Coussens L, Hayflick JS, Dull TJ, Gray A, Tam AW, et al. Human epidermal growth factor receptor cDNA sequence and aberrant expression of the amplified gene in A431 epidermoid carcinoma cells. Nature. 1984;309(5967):418–25.

Ushiro H, Cohen S. Identification of phosphotyrosine as a product of epidermal growth factor-activated protein kinase in A-431 cell membranes. J Biol Chem. 1980;255(18):8363–5.

Vogelstein B, Papadopoulos N, Velculescu VE, Zhou S, Diaz LA, Kinzler KW. Cancer genome landscapes. Science. 2013;339(6127):1546–58.

Wang S, Li J. Second-generation EGFR and ErbB tyrosine kinase inhibitors as first-line treatments for non-small cell lung cancer. Onco Targets Ther. 2019;12:6535–48.

Wang S, Zhang Z, Peng H, Zeng K. Recent advances on the roles of epidermal growth factor receptor in psoriasis. Am J Transl Res. 2019;11(2):520–8.

Wang SC, Hung MC. Nuclear translocation of the epidermal growth factor receptor family membrane tyrosine kinase receptors. Clin Cancer Res. 2009;15(21):6484–9.

Ward MD, Leahy DJ. Kinase activator-receiver preference in erbb heterodimers is determined by intracellular regions and is not coupled to extracellular asymmetry. J Biol Chem. 2015;290(3):1570–9.

Ward WHJ, Cook PN, Slater AM, Davies DH, Holdgate GA, Green LR. Epidermal growth factor receptor tyrosine kinase. Investigation of catalytic mechanism, structure-based searching and discovery of a potent inhibitor. Biochem Pharmacol. 1994;48(4):659–66.

Wee P, Wang Z. Epidermal growth factor receptor cell proliferation signaling pathways. Cancers(Basel). 2017:9(5):52.

Wen W, Chen WS, Xiao N, Bender R, Ghazalpour A, Tan Z, et al. Mutations in the kinase domain of the HER2/ERBB2 gene identified in a wide variety of human cancers. J Mol Diagnostics. 2015;17(5):487–95.

Wieduwilt MJ, Moasser MM. The epidermal growth factor receptor family: Biology driving targeted therapeutics. Cell Mol Life Sci. 2008;65(10):1566–84.

Wood ER, Truesdale AT, McDonald OB, Yuan D, Hassell A, Dickerson SH, et al. 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. 2004;64(18):6652–9.

Wu YL, Cheng Y, Zhou X, Lee KH, Nakagawa K, Niho S, et al. Dacomitinib versus gefitinib as first-line treatment for patients with EGFR-mutation-positive non-small-cell lung cancer (ARCHER 1050): a randomised, open-label, phase 3 trial. Lancet Oncol. 2017;18(11):1454–66.

Yamamoto T, Hihara H, Nishida T, Kawai S, Toyoshima K. A new avian erythroblastosis virus, AEV-H, carries erbB gene responsible for the induction of both erythroblastosis and sarcomas. Cell. 1983;34(1):225–32.

Yang XD, Jia XC, Corvalan JRF, Wang P, Davis CG. Development of ABX-EGF, a fully human anti-EGF receptor monoclonal antibody, for cancer therapy. Crit Rev Oncol Hematol. 2001;38(1):17–23.

Yarden Y, Sliwkowski MX. Untangling the ErbB signalling network. Nat Rev Mol Cell Biol. 2001;2:127–37.

Yasuda H, Kobayashi S, Costa DB. EGFR exon 20 insertion mutations in non-small-cell lung cancer: Preclinical data and clinical implications. Lancet Oncol; 2012;13(1):e23-31.

Yasuda H, Park E, Yun CH, Sng NJ, Lucena-Araujo AR, Yeo WL, et al. Structural, biochemical, and clinical characterization of epidermal growth factor receptor (EGFR) exon 20 insertion mutations in lung cancer. Sci Transl Med. 2013;5(216):216ra177.

Yu D, Hung MC. Overexpression of ErbB2 in cancer and ErbB2-targeting strategies. Oncogene. 2000;19:6115–21.

Yun C-H, Boggon TJ, Li Y, Woo MS, Greulich H, Meyerson M, et al. Structures of lung cancer-derived EGFR mutants and inhibitor complexes: mechanism of activation and insights into differential inhibitor sensitivity. Cancer Cell. 2007;11(3):217–27.

Zhang X, Gureasko J, Shen K, Cole PA, Kuriyan J. An Allosteric Mechanism for Activation of the Kinase Domain of Epidermal Growth Factor Receptor. Cell. 2006;125(6):1137–49.

ISBN 978-952-12-4105-5