Biochemical characterization of EGFR exon 20 insertion variants and their inhibitor sensitivities

HARVARD UNIVERSITY
Graduate School of Arts and Sciences

DISSERTATION ACCEPTANCE CERTIFICATE

The undersigned, appointed by the

Division of Medical Sciences

in the subject of Biological and Biomedical Sciences

have examined a dissertation entitled

Biochemical characterization of EGFR exon 20 insertion variants and their inhibitor sensitivities

presented by Hanchen Zhao

candidate for the degree of Doctor of Philosophy and hereby certify that it is worthy of acceptance.

Signature: Philip Cole
Philip Cole (Jan 20, 2023 17:09 EST)

Typed Name: Dr. Philip Cole

Signature: David Golan
David Golan (Jan 21, 2023 14:30 EST)

Typed Name: Dr. David Golan

Signature: Xin Zhou
Xin Zhou (Jan 22, 2023 09:45 GMT+8)

Typed Name: Dr. Xin Zhou

Signature: Margaret Stratton
Margaret Stratton (Jan 20, 2023 18:38 EST)

Typed Name: Dr. Margaret Stratton

Date: January 20, 2023

Biochemical characterization of EGFR exon 20 insertion variants and their inhibitor sensitivities

A dissertation presented

by

Hanchen Zhao

to

The Division of Medical Sciences

in partial fulfillment of the degree of

Doctor of Philosophy

in the subject of

Biological and Biomedical Sciences

Harvard University

Cambridge, Massachusetts

January 2023

Dissertation Advisor: Michael J. Eck Author: Hanchen Zhao

Biochemical characterization of EGFR exon 20 insertion variants and their inhibitor sensitivities

Abstract

Somatic mutations in the epidermal growth factor receptor (EGFR) are a major cause of non-small cell lung cancer (NSCLC). Among these structurally diverse alterations, exon 20 insertions represent a unique subset that rarely respond to EGFR tyrosine kinase inhibitors (TKIs), the leading class of targeted therapies for NSCLC in clinical use. Therefore, there is a significant need to develop inhibitors that are active against this class of activating mutations. Here, we conducted biochemical analysis of the two most frequent exon 20 insertion variants, V769_D770insASV (insASV) and D770_N771insSVD (insSVD), to better understand the molecular mechanisms underlying TKI sensitivity and resistance in these variants. From kinetic studies, we found that EGFR insASV and insSVD have lower $K_{m,\ ATP}$ values compared to the L858R variant, which contributes to their relative resistance to ATP competitive inhibitors. Biochemical, structural, and cellular studies of a diverse panel of EGFR inhibitors revealed that BAY-568, BAY-33, TAS6417, and TAK-788 inhibit EGFR insASV and insSVD in a mutant-selective manner, with BAY-568 being the most potent and mutant selective. Kinetic studies with

TAK788 showed that these covalent inhibitors have similar reversible binding affinities to wild-type EGFR and exon 20 insertion variants, but different covalent inactivation rates that may underlie their increased potency against insASV and insSVD. Similar experiments with BAY-33, a reversible inhibitor, suggest that mutant selectivity for this compound could be due to different rates of conformational change following reversible binding. Our systematic biochemical and structural analysis suggests that these mutant-selective compounds should be prioritized for clinical development, and in the long run our results should facilitate development of more mutant-selective compounds.

Table of Contents

List of Figures

List of Tables

Acknowledgements

There are so many people I would like to thank here during my scientific journey, especially over the last six years. Without every single one of you, it would be impossible for me to complete my thesis and pursue my PhD at Harvard. This journey has been filled with so many ups and downs, so I am grateful to all of you who helped me, guided me, or encouraged me along the way to reach the finish line.

First and foremost, I would like to sincerely thank my thesis advisor Professor Michael J. Eck. His patience and guidance have been enormous. I am really grateful to his advice and numerous discussions we had to push my thinking further, to help me make progress on my project. I am truly fortunate to be able to learn from him both as a student and a mentee. His great passion and commitment for science, as well as his persistent curiosity and careful thinking have taught me how to become not only a good scientist, but also a good person who could make contributions to our world in other science-related careers.

I have been fortunate to have many advisors throughout my graduate school. I would like to thank Professor Philip A. Cole, who served as chair of my thesis defense committee and a member of my dissertation advising committee (DAC). He is a very admirable professor with so much knowledge. His encouraging and insightful comments over the years on my thesis project not only helped me push my project further scientifically, but also supported me mentally by giving me confidence and inspiring my passion for science. I am also grateful to Professor Sara J. Buhrlage for serving as chair of my DAC. Her smile and encouragement made me believe in myself as a scientist throughout my PhD studies. Thank you to Professor Andrew Kruse for serving on my DAC. I appreciate his insightful ideas for my research as a brilliant scientist. I would

also like to thank Professor Stephen Buratowski for serving as chair of my qualifying exam committee. His advice for me to be a well-rounded scientist early on in my PhD has been really beneficial for my career, especially since I would be moving on to work on the interface of science and business after graduation. Thank you to all my rotation mentors Professor David M. Langenau, Professor Sichen Shao, and especially Professor Stephen C. Blacklow, for years of helpful comments in super group meetings and departmental talks. I also enjoyed his biochemistry class a lot, which reminded me of how much I like science.

Next, I would like to thank past and current members of the Eck lab. Kunhua Li, I have to thank you for your friendship and encouragement throughout my early PhD time, as those years are the hardest time for me due to constant failures in lab. Eunyoung Park, I would like to thank you for your mentorship during my rotation as well as later in my PhD. Javier Camallonga, thank you for being so sociable and positive. You are a sunshine to the lab. Bo Yang, thank you for being a good friend and always being supportive throughout my last year of PhD. Dongman Jang, thank you for being around as a great friend and also for your help with my structural figures. Your skills in PyMOL are impeccable and I have learned so much from you. Thank you to the rest of Eck lab members and alumni who make the lab a professional and friendly environment.

Not only do I need to thank Eck lab members, but I also have to express my gratitude towards some great neighbors. As a PhD student, I have done most of the experiments for my project by myself, but I have also been fortunate to connect and collaborate with Jie Jiang from the Jänne lab. She is a great collaborator who is always punctual and easy to communicate with. I am truly grateful to all the western blots and cell viability assays she did. I would also like to thank our old neighbor, Fischer lab members, for making the lab environment friendly and also for the many scientific discussions we had. Specifically, I want to thank Hong Yue for her

friendship and also Hojong Yoon, Jonathan Bushman, Cyrus Jin for being both great classmates and friends.

Outside of lab, I have to thank my BBS classmates, for making Harvard PhD life enjoyable. Francois Thelot, thank you for being such a close friend. I am glad that we have chosen similar career path, and I really enjoy our monthly catchup time. Shiyu Xia, thank you for being my friend for so long. We met during the Harvard interview weekend, and have since been so close to each other. Shiyu, Francois and I have a group chat, and we gather together from time to time to discuss science and career paths. Inwha Baek, thank you for being a nice and supportive friend. Thank you to Gaurav Luthria, Samantha Wong, and Heer Joisher for leading the student club with me. It was such an honor to be a co-president of the club and a student leader at Harvard, but what makes me proud is to expand the student club so much in such a short period of time, which would not have been possible without you. I would also like to thank all my Chinese friends in Boston, who have made me feel at home during my undergrad and PhD over the last decade.

Last but not least, I would like to acknowledge my family. Thank you to my parents for their unconditional love and support. Words cannot express how grateful I am to have such a positive and supportive family. It was very hard for them to let me leave the family at age of 17 for a foreign country, but they want me to have the best education and truly want me to succeed. I will forever be grateful to their love and support.

To my parents for their love and support

Chapter 1

Introduction

1.1: Significance of EGFR-mutated non-small cell lung cancer

EGFR and cancer

Overexpression of the epidermal growth factor receptor (EGFR) has been linked to the development and progression of several types of cancer, including non-small cell lung cancer (NSCLC), colorectal cancer, head and neck cancer, breast cancer, ovarian cancer, bladder cancer, and pancreatic cancer[1–4]. Among various types of cancer, lung cancer contributes most to the overall cancer-related deaths, with approximately 2.2 million new cases diagnosed and 1.8 million deaths each year worldwide[5]. As the most common type of lung cancer, NSCLC accounts for about 85% of all lung cancer cases, with overexpression of EGFR present in 40-80% of the cases[6–10]. Therefore, initial efforts were focused on developing inhibitors to leverage the oncogene addiction by targeting overexpression of the wild-type EGFR in NSCLC, but with little success[11–13].

Even though these tyrosine kinase inhibitors (TKI), such as gefitinib, did not show promising results as expected, a subgroup of patients experienced rapid and significant responses to them, which was later identified to harbor somatic mutations in EGFR[10,14,15]. It turned out that these inhibitors were fortuitously more potent against most EGFR mutants than against wild-type EGFR[16,17]. The discovery of the correlation between certain EGFR mutations and response to gefitinib treatment has emphasized the importance of targeting mutated kinases as a strategy for treating cancer and has also provided a way to identify those NSCLC patients that are likely to respond to EGFR-TKI therapy[4,10,15,18,19].

As one of the most frequent driver mutations causing NSCLC, somatic mutations in EGFR are present in up to 76% of NSCLC patients, with significantly more occurrences in Asian

populations than in other ethnicity groups[10,18,20–23]. These mutations cause over-activation of EGFR, thus enhancing downstream proliferation and anti-apoptotic signals through activation of MAPK/Erk and PI3K/Akt pathways, which contribute to cancer progression[4,10,24]. So far, a number of structurally diverse alterations have been identified, including small in-frame insertions and deletions, and point mutations. The L858R point mutation in exon 21 and deletions in exon 19 are the two most common oncogenic EGFR alterations and are often referred to as “classical” activating mutations, which collectively represent approximately 85% of all EGFR mutations[25,26]. The other EGFR mutations are much rarer, such as exon 20 insertions, point mutations of Gly 719, L861Q, and S768I[23,27,28]. Although not as prevalent as classical mutants, it is estimated that over 30,000 new diagnoses of NSCLC each year will involve patients with rare EGFR mutations, due to the high overall prevalence of lung cancer[27]. As the third most common mutation, exon 20 insertions represent 4-12% of all EGFR mutations and account for approximately 10,000 new cases of NSCLC annually around the world[29–32]. However, unlike the L858R and exon 19 deletions, exon 20 insertions have not been extensively studied biochemically and more effective treatment is urgently needed.

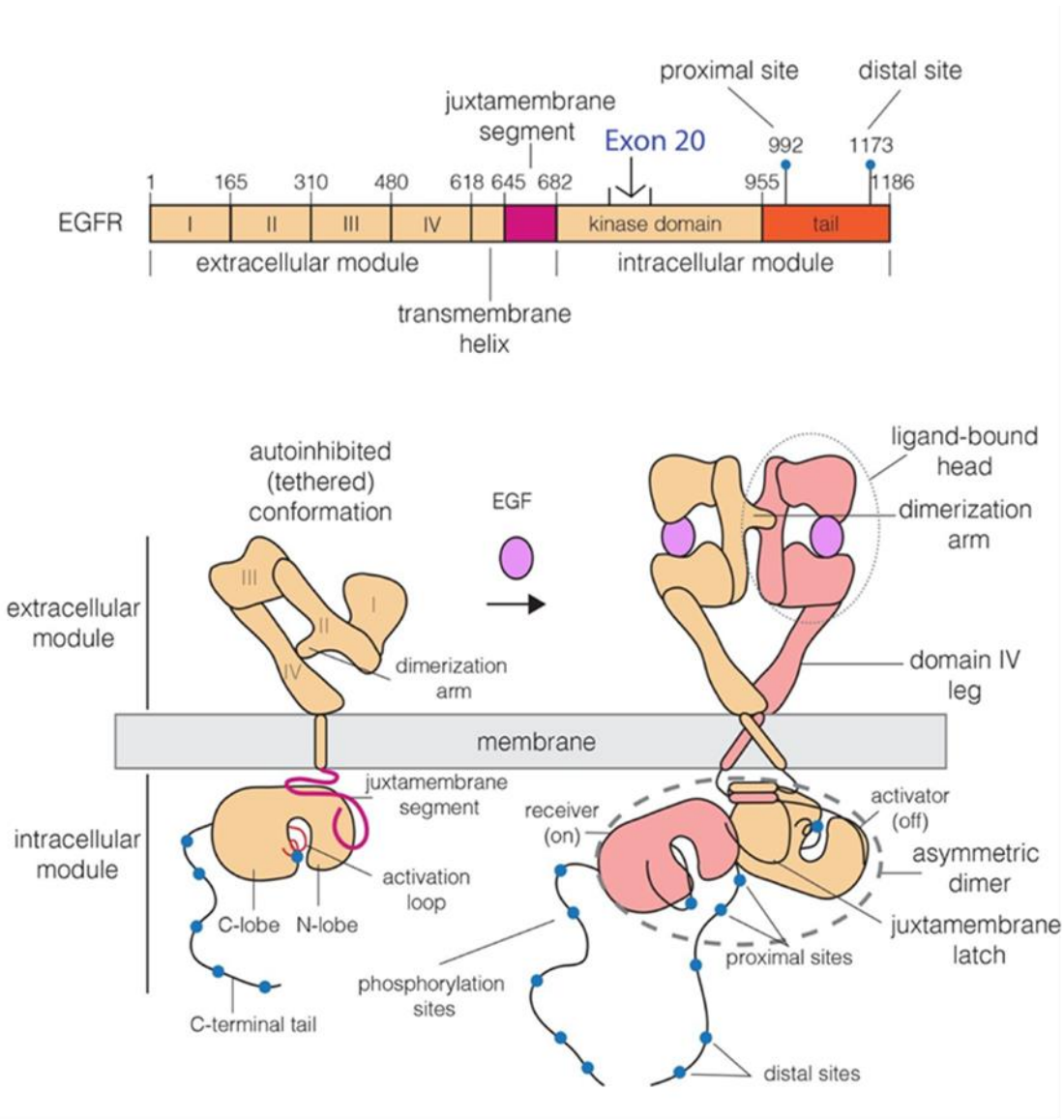


Figure 1.1. EGFR structure and regulation. A, Domain boundaries in human EGFR. The residue numbers do not include the 24-residue signal sequence. **B,** The canonical model for the monomer to dimer transition in EGFR that is triggered by ligand binding (Huang et al. 2016[33]).

EGFR signaling

Receptor tyrosine kinases (RTK) span the cell membrane and typically consist of an extracellular region that often functions as a ligand binding site, a single transmembrane helix, a juxtamembrane regulatory region, an intracellular tyrosine kinase domain, and a carboxy (C-) terminal tail[34–36]. EGFR, also known as HER1 or ErbB1, is the first identified member of the ErbB family of RTKs, which have received significant attention for their potential as targets for cancer therapeutics due to their abnormal expression in many epithelial tumors and their role in promoting the growth and survival of cancer cells[37–40]. The other ErbB family members are HER2 (ErbB2), HER3 (ErbB3), and HER4 (ErbB4)[2,4,41]. Although these RTKs are structurally similar to each other and share many functional characteristics within the ErbB family, there are some important differences to be noted[42–44]. Specifically, unlike EGFR or HER4, HER2 does not bind ligands and its extracellular domain is always in an extended orientation. HER3 is a pseudo-kinase with a PI3K binding motif in its C-terminal tail that is not present in other ErbB family members, which is important for the PI3K/Akt signaling pathway[45]. Although there are some reports of residual activity, it is generally thought that HER3 lacks tyrosine kinase activity and relies on other ErbB family members to activate downstream signaling pathways[43,46–51].

The full-length EGFR is a transmembrane protein, which consists of a 24-residue signal peptide, an extracellular module, a transmembrane helix, a juxtamembrane segment, a kinase domain, and a long C-terminal tail[33,52–54] (Figure 1.1, A). The signal peptide is cleaved after the nascent protein is transported to the endoplasmic reticulum (ER), resulting in the 1186-residue long mature transmembrane protein[33,55,56]. EGFR is activated by binding to specific ligands, such as epidermal growth factor (EGF)[2,33,57]. At least seven different activating ligands for EGFR have been identified in mammals: EGF, transforming growth factor α (TGF-α), epigen (EPG), amphiregulin (AR), heparin-binding (HB)-EGF, betacellulin (BTC), and epiregulin (EPR), with the

first four exclusive to EGFR and last three being bispecific ligands that regulate both EGFR and HER4[41,51,53,58–60]. All of these EGFR ligands are produced initially as transmembrane precursor proteins and then later are cleaved by cell-surface proteases, such as the ADAMs (a disintegrin and metalloproteinase), to make them soluble peptides ready to bind EGFR[53,59,60].

Under normal physiological conditions, a monomeric EGFR adopts an autoinhibited conformation, with its extracellular module in a tethered and closed state[33,39]. However, upon binding to ligands, the extracellular module of EGFR undergoes a conformational change, which allows it to form either a homodimer with another EGFR, or a heterodimer with another member of the ErbB family[40,58,61]. The extracellular module of HER2 is always in an extended and open state, which makes it ready to form such a heterodimer with ligand-bound EGFR, while HER3 and HER4 can also form heterodimers with ligand-bound EGFR when their activating ligands are present[40,46,53,59].

In contrast with most RTKs, where ligands bind at the dimer interface to bring together two receptors, dimerization of EGFR is entirely “receptor mediated”, which means that activating ligands do not directly contribute to the dimer interface[34,53]. Instead, ligand binding breaks the tether between domains II and IV of the extracellular module by simultaneously interacting with domains I and III, which exposes the dimerization arm of domain II to interact with another activated EGFR or another ErbB family member in the extended form[33,34,46,53] (Figure 1.1, B). As mentioned above, EGFR can be activated by different ligands, and each ligand generates a different signal, which could be partially explained by conformational differences of the ligand-binding head (domain I, II, and III) and the juxtamembrane segment when bound to ligands like EGF or TGF-α[62–66]. The difference was further revealed by full-length EGFR cryo-EM structures highlighting different abilities of EGF and TGF-α to stabilize the tips-separated conformation of EGFR dimers[67]. It has been shown that the tips-separated conformation of the extracellular

module is associated with the N-terminal dimer conformation of the transmembrane helices, which is correlated with increased intracellular kinase activity[67–69]. Since binding of EGF could better stabilize the tips-separated conformation, it sends stronger signals than binding of TGF-α[67] (Figure 1.2).

Similar to its extracellular module, the intracellular kinase domain of EGFR is also autoinhibited in the absence of activating ligand[34,51,70]. Unlike other RTKs, autophosphorylation of the activation loop is not required for the activation of EGFR kinase domain[46,53,71]. Instead, the dimerization of extracellular modules induces the transmembrane helices and juxtamembrane segments to dimerize, leading intracellular kinase domains to form an asymmetric dimer, in which the receiver kinase is activated by the activator kinase[33,70,72–76] (Figure 1.1, B). Activation of the receiver kinase results in the autophosphorylation of up to ten tyrosine residues in the C-terminal tail, such as Y1069, Y1092, Y1172, and Y1197, which subsequently recruits phosphotyrosine-binding (PTB) proteins containing Src homology 2 (SH2) or PTB domains, such as Grb2 and Shc, to initiate the downstream signaling cascades[39,77,78].

Even though some phosphotyrosines could signal through negatively acting pathways, such as Cbl-dependent EGFR ubiquitination to trigger degradation, most of them signal for anti-apoptosis, proliferation, differentiation, and migration through MAPK/Erk, PI3K/Akt, PLCγ, STAT, and Src/FAK pathways, with the first two being the primary ones[2,41,79–81] (Figure 1.3). The MAPK/Erk pathway is initiated by binding of adaptor proteins, Shc and Grb2, to phosphorylated EGFR, recruiting effector proteins, such as Sos1 and Sos2, to activate Ras[41,82,83]. These effector proteins are called guanine nucleotide exchange factors (GEF), which promotes replacement of GDP by GTP on Ras, turning Ras into its activated form[82]. The GTP-bound Ras binds to Raf to relocate it to the plasma membrane and activates it from an autoinhibited state, which then phosphorylates and activates MEK[41,82,84]. After that, MEK phosphorylates and activates Erk,

which phosphorylates transcription factors or their interacting partners to ultimately stimulate cell proliferation and differentiation[2,41,84].

The PI3K/Akt pathway starts with phosphorylated EGFR recruiting the adaptor proteins, Grb2 and GAB1, which promotes the recruitment and activation of PI3K[41,85]. Then the activated PI3K phosphorylates a lipid second messenger, PIP_2, to generate PIP_3, which recruits and activates PDK1[86,87]. The activated PDK1 in turn phosphorylates and activates Akt (also known as PKB), phosphorylating transcription factors, such as the fork head box O (FOXO), to inhibit their binding to the target DNA and ultimately leading to suppression of apoptosis for cell survival[82,86,88]. As a result, overactivation of EGFR is associated with proliferation, and EGFR was also the first receptor directly linked to human cancer[39,51,89,90]. Since its first linkage to cancer, various therapeutic agents have been used to target EGFR, with TKIs and monoclonal antibodies being most promising and widely used[39,41].

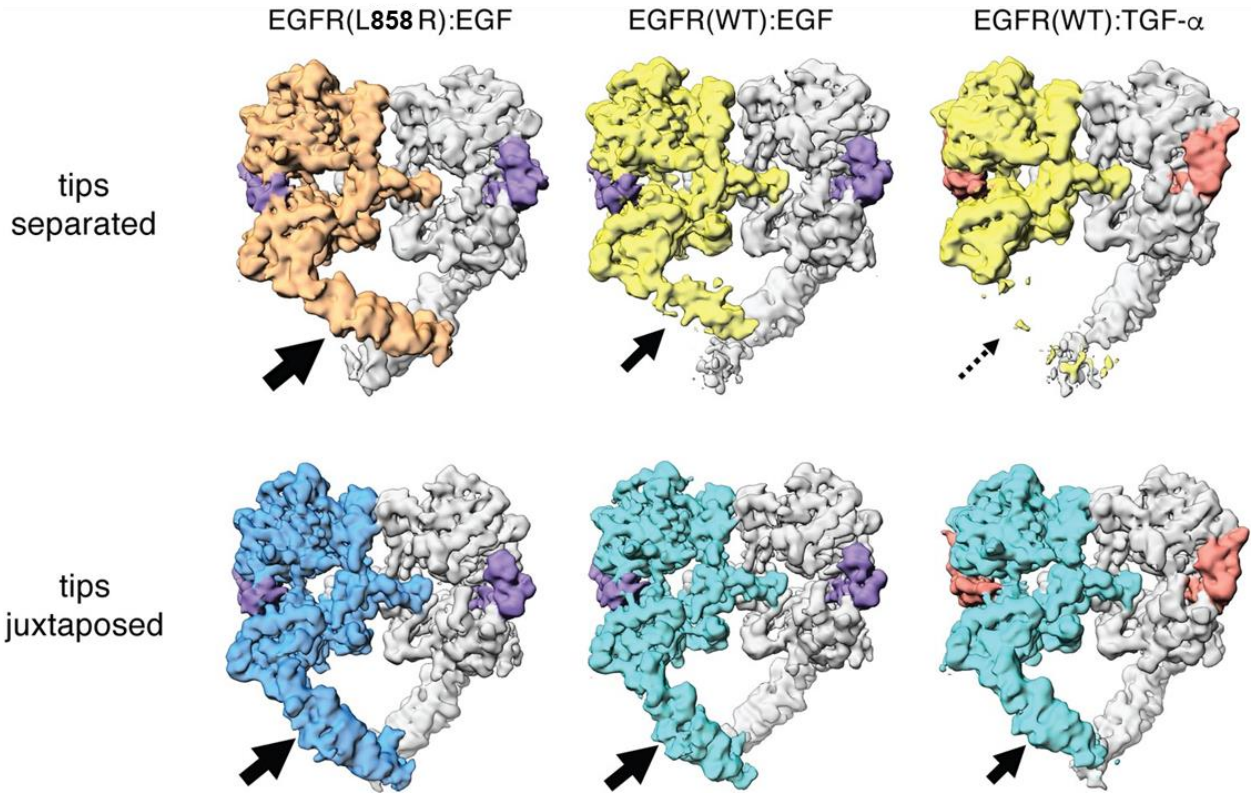


Figure 1.2. Tips-separated and tips-juxtaposed conformations of EGFR dimers. Comparison of the tips-separated and tips-juxtaposed conformations in different EGFR:ligand complexes. Cryo-EM density maps of EGFR(L858R):EGF, EGFR(WT):EGF, EGFR(WT):TGF-α in the tips-separated or the tips-juxtaposed conformations were shown (Huang et al. 2021[67]).

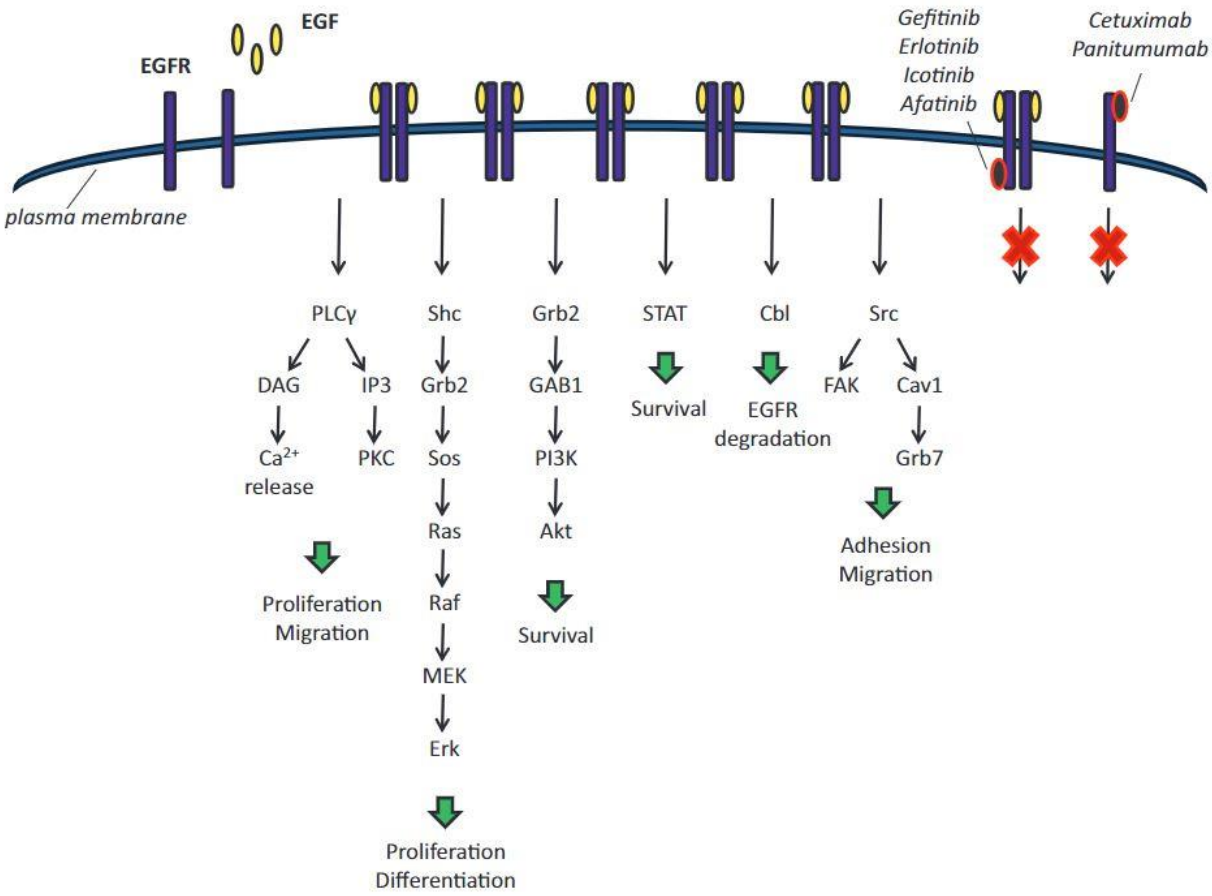


Figure 1.3. EGFR downstream pathways. Ligand binding induces EGFR dimerization and subsequent autophosphorylation of tyrosine residues in the C-terminal tails recruit different downstream effectors, leading to proliferation, differentiation, migration, survival, and in the case of Cbl, degradation. EGFR TKIs, such as gefitinib, were developed to target the kinase domain while anti-EGFR monoclonal antibodies, such as cetuximab, were developed to prevent dimerization of extracellular modules (Jones et al. 2014[41]).

EGFR kinase structure

Like other protein kinases, the kinase domain of EGFR is comprised of a smaller N-terminal lobe (N-lobe) with a five-stranded β-sheet and an α-helix called the C-helix or αC-helix, and a larger C-terminal lobe (C-lobe) with six α-helices and two short antiparallel β-sheets[28,91,92] (Figure 1.4, A). As a tyrosine kinase, EGFR catalyzes the phosphorylation reaction by transferring the γ-phosphate from ATP to a tyrosine residue of the substrate[91,93]. The ATP binding site is located at the interface between the N-lobe and the C-lobe, beneath a highly conserved glycine-rich loop called the phosphate binding loop, or P-loop, which coordinates the phosphates of ATP and positions the γ-phosphate for catalysis[94,95]. The peptide substrate utilizes the activation loop (A-loop), which contains a conserved DFG motif, as a platform to access ATP for reaction when EGFR is in the active state[93,94]. The catalytic loop, which contains a conserved HRD motif, is responsible for transferring the γ-phosphate of ATP to the peptide substrate[95,96].

The transition between the active and inactive states of EGFR kinase features transformation of activation loop and movement of the regulatory C-helix[28,97] (Figure 1.4, B). Under normal physiological conditions, the wild-type EGFR kinase adopts an auto-inhibited inactive conformation[70]. The activation loop stabilizes the C-helix in its inactive position and blocks its access to the ATP binding site[28,97–99]. Upon activation, the activation loop rearranges, allowing the C-helix to get close to the ATP binding site to make the EGFR kinase form the active conformation[28,97,99]. In the active state, the DFG motif is in a conformation that D855 is positioned towards bound ATP and coordinate with the magnesium ion to bind ATP, while in the inactive state of many kinases, the DFG motif is in a conformation that the conserved aspartate is oriented away from the ATP binding site and the phenylalanine blocks access of ATP[92,93,98,100,101]. However, the “Src-like inactive” conformation is observed in inactive EGFR structures[102]. The DFG motif is the same as the active conformation and the activation loop forms a helical turn that

interacts with the C-helix[98,99]. In addition, in the active conformation, E762 on the C-helix forms a salt bridge with K745 in the ATP binding site to coordinate the α and β phosphates of ATP, while the salt bridge is disrupted in the inactive conformation due to the displaced C-helix[28,94].

As mentioned before, the inactive EGFR is activated by dimerization of the extracellular modules upon ligand binding, which induces the formation of an asymmetric dimer of the kinase domains, without requiring phosphorylation of its activation loop[28,53,70,103]. Specifically, the extracellular dimerization is accompanied by dimerization of transmembrane and intracellular domains, including formation of an anti-parallel helix dimer of juxtamembrane segments and an asymmetric kinase dimer[74–76]. In the asymmetric dimer, the C-lobe of the activator kinase packs against the N-lobe of the receiver kinase, which "pushes" the C-helix of the receiver kinase into its active conformation[28,70] (Figure 1.5). This interaction between two EGFR kinase domains is mechanistically similar to the interaction between cyclin dependent kinases (CDK) and cyclins, where a cyclin packs against the C-helix of a CDK that "pushes" it inward for activation[70,104].

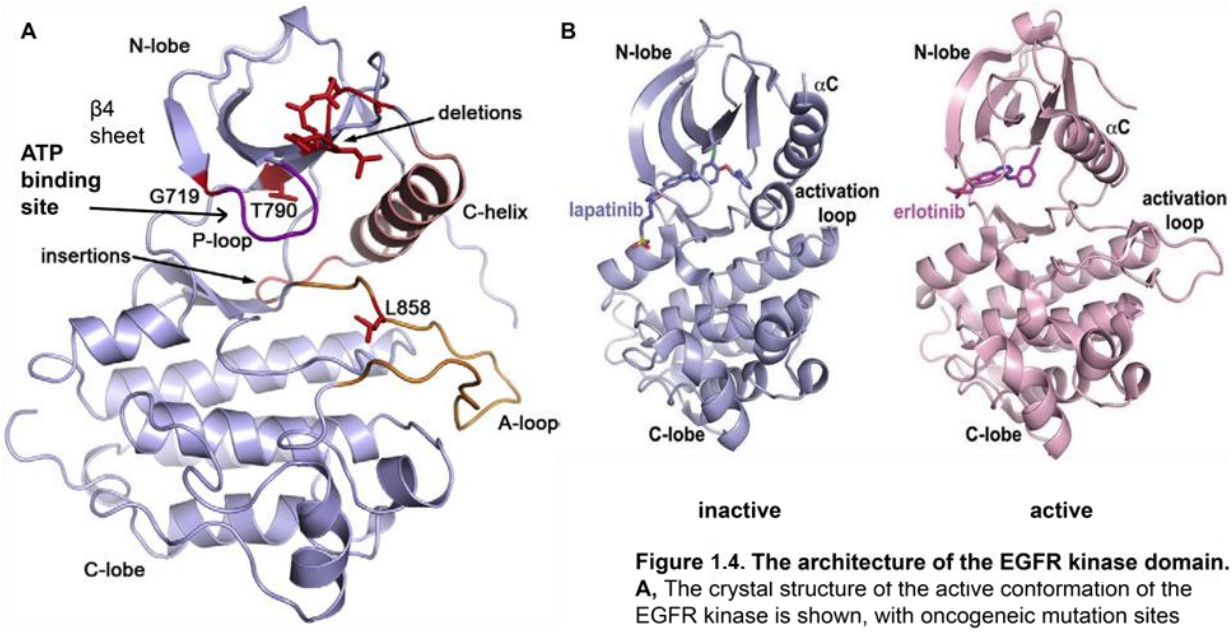


Figure 1.4. The architecture of the EGFR kinase domain. **A,** The crystal structure of the active conformation of the EGFR kinase is shown, with oncogeneic mutation sites highlighted (PDB ID: 1M14). **B,** The side-by-side comparison of the inactive conformation (left) (PDB ID: 1XKK) and the active conformation (right) (PDB ID: 1M17) of EGFR kinase is displayed. The regulatory αC helix swings in and out between the two conformations. (Eck et al. 2010[28]; Park et al. 2012[97]).

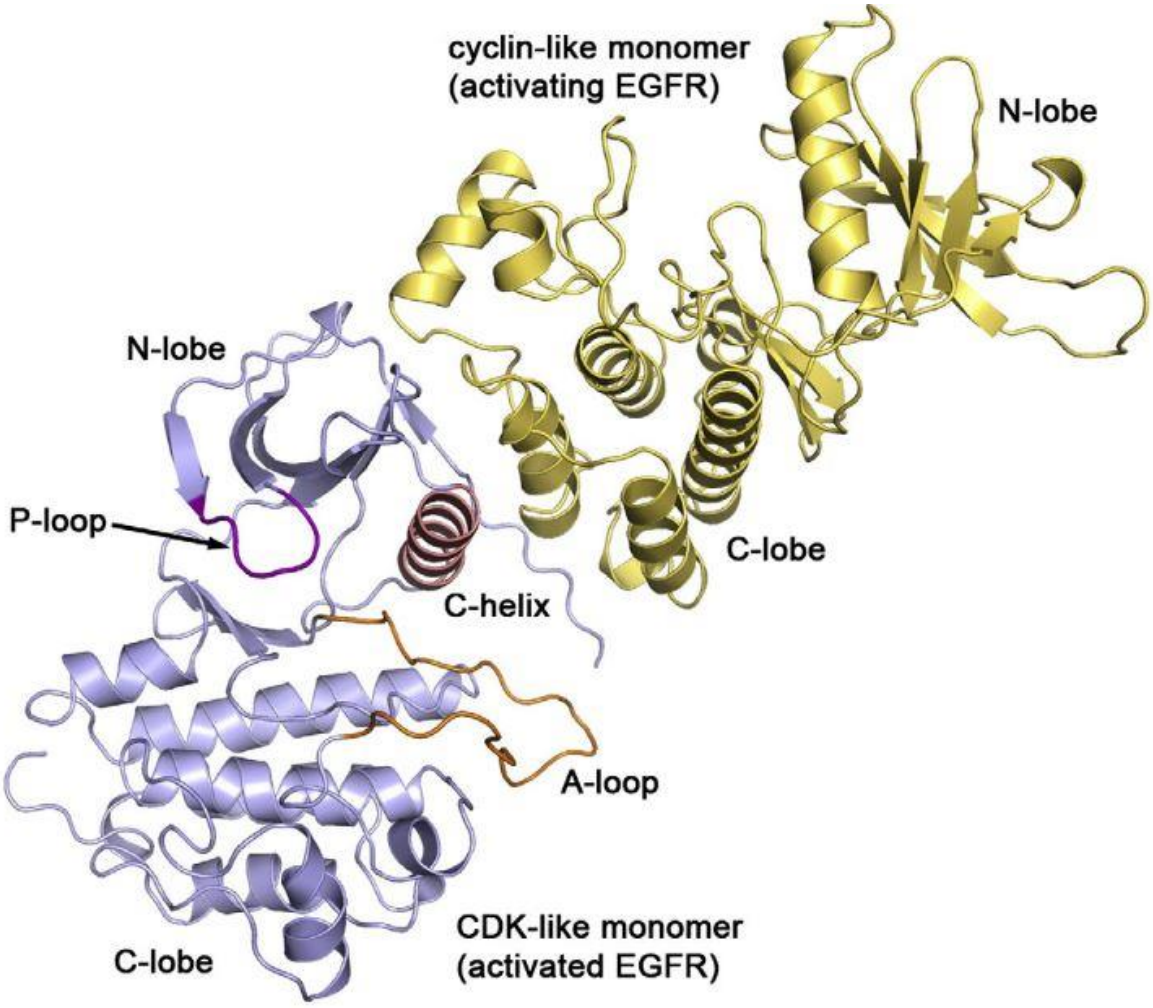


Figure 1.5. Asymmetric dimer of EGFR kinase domains. Crystal structure of an EGFR kinase dimer was shown. The C-lobe of the activating EGFR packs against the N-lobe of the activated EGFR to "push" its C-helix to the inward, active position, which is similar to the activation of cyclin dependent kinases (CDK) by cyclins (Eck et al. 2010[28]).

1.2. Development of 1st-4th generation of EGFR tyrosine kinase inhibitors (TKIs)

Significance of tyrosine kinase inhibitors against EGFR mutants

Somatic mutations in EGFR have been identified in various cancer patient tissue samples, such as gliomas, lung, ovarian, colorectal, head and neck, prostate, and breast cancers[40,58,105]. Among them, NSCLC patients, representing 85% of lung cancer patients, are affected by EGFR mutations the most, with an occurrence rate of 10-20% in Caucasian patients and greater than 50% in Asian patients[4,7,27,28,106]. The L858R point mutation and exon 19 deletions are the two most common EGFR mutations in NSCLC, which collectively represent 85% of all EGFR mutations[27,106,107]. The other mutations include point mutations in exon 18 (G719X) or exon 21 (L861Q), and the exon 20 insertion mutations, which account for 4-12% of EGFR mutations in NSCLC as the third most common mutation[32,107] (Figure 1.4, A). These mutations cause over-activation of EGFR and thus enhance downstream anti-apoptotic signals, which make cells proliferate without any control, and thus contribute to cancer progression[10,39].

Remarkable success has been achieved over the past decade by treating advanced cancer patients with genotype-directed targeted therapy, especially for those patients harboring somatic mutations in EGFR[108–111]. Although monoclonal antibodies, such as cetuximab, were developed to target the EGFR ectodomain to prevent dimerization and achieved some clinical success[112–114], the activation of some EGFR mutants, such as the exon 20 insertion mutants, is independent of dimerization[115]. Therefore, small molecule inhibitors targeting mutant EGFR kinase domains are of particular importance for selectively killing cells harboring EGFR mutations while sparing normal cells[116–118].

1st generation EGFR TKIs

The 1st generation EGFR/HER2 TKIs were designed as reversible ATP-competitive inhibitors almost two decades ago with two of them approved by the FDA, including gefitinib and erlotinib[14,119,120]. However, the initial development of these TKIs was focused on tumors overexpressing wild-type EGFR, which showed little clinical activity[10–12]. Fortunately, it turned out that patients with somatic mutations in EGFR are more sensitive to TKI treatment than those without mutations[10,15,18,121]. Patients with the L858R mutation or exon 19 deletions generally show radiographic response to treatment of these TKIs[10,15,18,19]. As a result, gefitinib received accelerated approval by the Federal Drug Administration (FDA) in 2003 for treatment of advanced NSCLC after failure of chemotherapies, as the first FDA-approved EGFR TKI[14]. One year later, erlotinib was also approved by FDA as the second EGFR TKI on the market to treat advanced NSCLC[120,122].

In addition to these two EGFR TKIs, lapatinib, which is a reversible HER2 TKI, was also approved by the FDA in 2007 to treat HER2-positive metastatic breast cancer[123–125]. Notably, lapatinib is a dual inhibitor of EGFR and HER2 that binds to a Src-like inactive conformation, where the DFG motif is flipped "in" and the C-helix is pushed "out", which is classified as a type-1.5 inhibitor[102,126]. Unlike type-2 inhibitors that bind to the "DFG-out C-helix-out" inactive conformation, type-1.5 inhibitors like lapatinib bind to kinases with a "DFG-in" conformation, most commonly seen in an active kinase[95]. The other two clinically approved 1st generation TKIs, gefitinib and erlotinib, are type-1 inhibitors that bind to active kinases with a "DFG-in C-helix-in" conformation, similar to most other ATP-competitive inhibitors[95,126].

Biochemical studies revealed that classical EGFR mutants (L858R and exon 19 deletions) have much weaker affinities for ATP than wild-type EGFR, which gives advantages to ATP-

competitive inhibitors like 1st generation EGFR TKIs for targeting mutants[16,17,28,127,128]. The studies also showed that 1st generation TKIs bind more tightly to classical EGFR mutants than wild-type EGFR, which further explains the difference in treating NSCLC patients with or without EGFR mutations[16,128].

2nd generation EGFR TKIs

Unfortunately, even though 1st generation EGFR TKIs have achieved remarkable initial success, all responding patients eventually develop acquired resistance to TKI therapy, with over 50% of cases due to T790M, a secondary mutation in exon 20 of the EGFR kinase domain[129–132]. The point mutation at the T790 gatekeeper residue, which is a conserved residue at the opening of the ATP-binding pocket, is analogous to the acquired T315I mutation in ABL following imatinib (Gleevec) treatment[133]. Small-cell transformation, HER2 amplification, and MET amplification to bypass EGFR signaling were also identified in those patients that developed resistance[134–136]. Although T790M most often arises in those patients that received 1st generation TKI treatment, 2nd generation EGFR TKIs, designed with a Michael acceptor site to covalently bind at C797, were reported to be effective against the gatekeeper mutation in preclinical studies[121,127,137–139] (Figure 1.6).

Biochemical studies have shown that T790M restores affinity for ATP in EGFR mutants to a level comparable to wild-type EGFR, and that the mutation also induces compromised inhibitor binding affinities, so that reversible ATP-competitive inhibitors like 1st generation EGFR TKIs would not be able to selectively target EGFR mutants[127]. The 2nd generation TKIs were developed to irreversibly inhibit EGFR kinase activity, thus are much more potent than 1st generation TKIs[119,140]. In 2013, as a novel EGFR inhibitor at that time, afatinib received FDA approval for first-

line treatment of NSCLC patients with EGFR mutations[137]. Later in 2018, another 2nd generation EGFR TKI, dacomitinib, also received approval by the FDA as a first-line therapy for metastatic NSCLC patients with EGFR mutations, after showing promising results from clinical trials[138,141]. Notably, neratinib and pyrotinib, which are covalent structural analogs of lapatinib and dual inhibitors of EGFR and HER2, also received approval to treat HER2-positive breast cancer in the U.S. and China, respectively[142–145].

3rd generation EGFR TKIs

These covalent 2nd generation EGFR TKIs, however, are not clinically effective against TKI acquired resistance and often cause severe side effects to patients, resulting in frequent dose interruptions[140,146]. The most common 2nd generation TKI treatment-related adverse events (TRAE) are skin rashes, diarrhea, and stomatitis, which are observed in vast majority of cases and often lead to drug interruptions or discontinuation[137,147]. Although 2nd generation EGFR TKIs showed activity against T790M in preclinical studies, they are also very potent inhibitors of wild-type EGFR and other ErbB family members[148–150]. They do not have a wide therapeutic window to treat patients because normal cells are also killed alongside cancer cells[151]. To widen the therapeutic window, 3rd generation EGFR TKIs were rationally designed to target the mutant proteins with selectivity, first of which was WZ4002[152] (Figure 1.6). As the first published 3rd generation EGFR TKI that selectively inhibits T790M, WZ4002 was designed with an anilinopyrimidine scaffold that targets Methionine 790 and a Michael acceptor to bind at Cysteine 797[152,153]. Although WZ4002 did not progress to clinical trials, it provided proof-of-concept for mutant-selective irreversible inhibition of T790M[153–155].

Five years later, osimertinib was reported to potently inhibit T790M while sparing wild-type in preclinical experiments, and preliminary results from clinical trials also indicated that it is effective with minimal side effects in treating NSCLC patients who had disease progression following TKI treatment, which ultimately led to its first FDA approval in 2015 to treat metastatic T790M-positive NSCLC patients[153,155,156]. Due to its clinical success, osimertinib received additional FDA approval as the first-line therapy for metastatic NSCLC patients with EGFR mutations in 2018, and two years later, it was approved by the FDA for adjuvant therapy after surgical resection in NSCLC patients with EGFR mutations[157,158]. Since osimertinib can cross the blood-brain carrier to overcome brain metastasis and also has a much wider therapeutic window than the first two generations of TKIs, it has gradually become the first choice of TKI therapy against classical EGFR mutants with or without the acquired resistance mutation T790M[158–161]. Some other examples of 3rd generation TKIs are rociletinib, olmutinib, naquotinib, mavelertinib and nazartinib, but none of them has been approved by the FDA, with rociletinib abandoned after the FDA refused to grant accelerated approval and olmutinib withdrawn as a result of serious toxicities[162,163] (Figure 1.6).

4th generation EGFR TKIs and allosteric inhibitors

Even though osimertinib has shown remarkable success in treating NSCLC patients with EGFR mutations, C797S often appears at an occurrence rate of 40-66% after extensive 3rd generation TKI treatment, which relies on formation of covalent bond with C797[164–168]. As a result, patients with the acquired resistance mutation C797S are resistant to 3rd generation TKIs, so new inhibitors are being studied and developed to overcome the additional acquired mutation C797S[163,169]. Notably, when osimertinib is used as the first-line therapy for patients with classical

EGFR mutations, C797S develops as a secondary on-target resistance mutation in the absence of T790M, so 1st generation TKIs could retain their efficacy and be used as a second-line or combination therapy to treat these patients[157,164–166]. However, for those patients with C797S as a tertiary mutation in the presence of a classical EGFR mutation and T790M, none of the first three generations of TKIs are effective[163,165,166]. Therefore, allosteric inhibitors and reversible 4th generation TKIs emerged recently as a strategy to overcome the C797S resistance mutation following 3rd generation TKI treatment[170–183] (Figure 1.6).

As the first EGFR allosteric inhibitor to target the T790M and C797S mutations, EAI045 was identified from a large screen of around 2.5 million compounds against the purified L858R/T790M kinase, with further optimization through medicinal chemistry[174]. Since top hits from the screen were assayed at different ATP concentrations and also were counter-screened against the wild-type kinase, non-ATP competitive mutant-selective compounds could be found for further studies[174,184]. Although EAI045 was not effective as a single agent against the double mutant L858R/T790M, it exhibited synergistic effect and induced significant tumor shrinkage in the double mutant as well as the triple mutant L858R/T790M/C797S murine models when used in combination with cetuximab, a monoclonal antibody that blocks EGFR dimerization[169,174]. This suggests that EAI045 target monomeric kinases and would be more effective in a combination therapy[174,184].

Following the discovery of EAI045, more allosteric inhibitors have been developed, with some of them, such as JBJ-09-063, effective as a single agent and not requiring combination therapies with other agents like cetuximab[175–177]. Notably, these allosteric inhibitors exhibit mechanistic synergy with EGFR TKIs like osimertinib by co-binding the kinase, leading to enhanced efficacies under coadministration[176,177,185]. However, allosteric inhibitors are not effective against exon 19 deletions since their mutant selectivity comes from an allosteric pocket

created by the L858R mutation[174,178,179,182]. By contrast, 4th generation TKIs, such as BI-4020 and BLU-945, have been developed to target the T790M and C797S resistance mutations in the presence of either classical mutation[178–182,186,187]. These reversible ATP-competitive inhibitors selectively target the classical mutants, the double mutants (del19/T790M, L858R/T790M), and the triple mutants (del19/T790M/C797S, L858R/T790M/C797S) in preclinical studies, with BLU-945, BBT-176, and BPI-361175 currently in phase I/II clinical trials to overcome these EGFR resistance mutations[177–182,186–190].

EGFR exon 20 insertion inhibitors

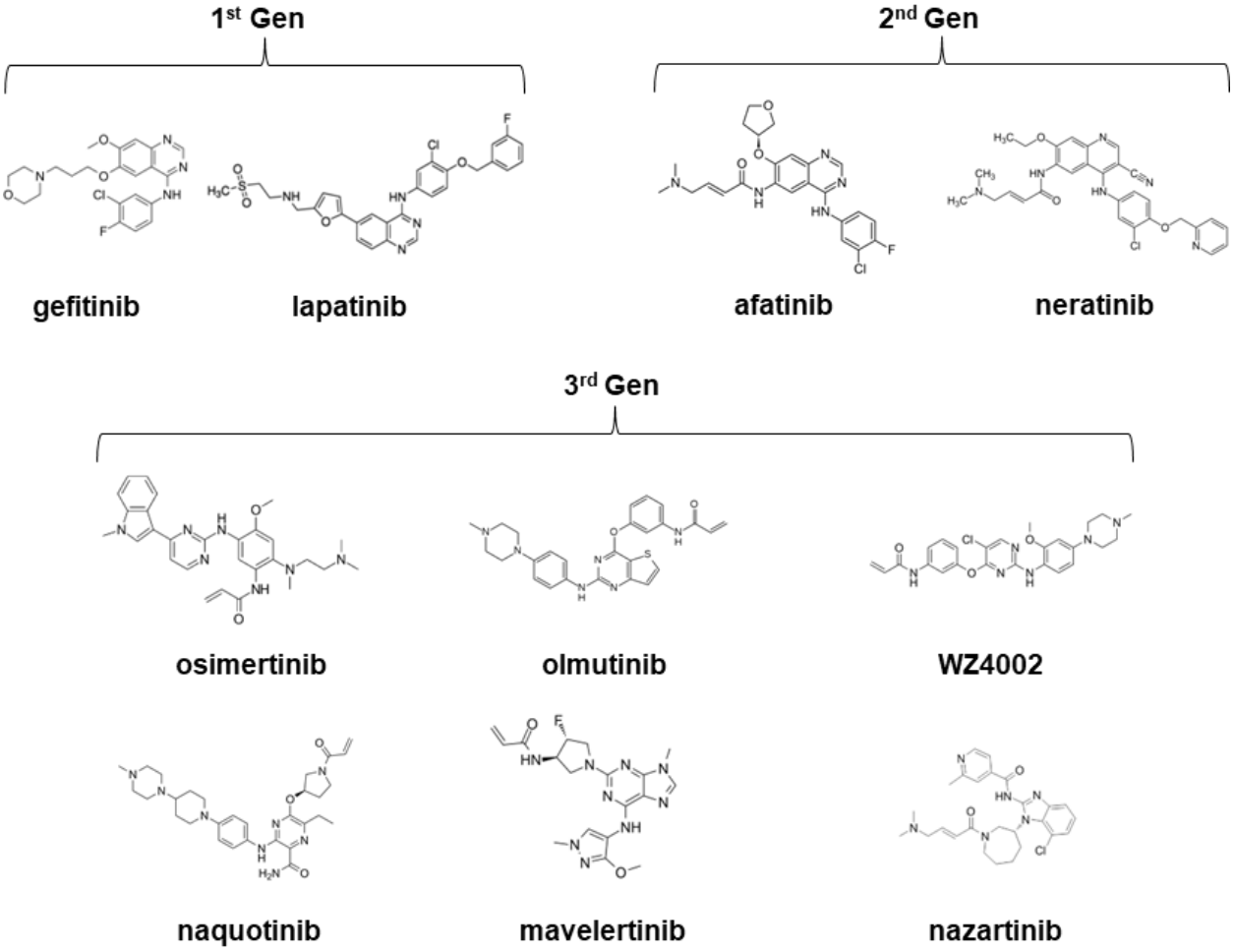
Even though the first four generations of TKIs have been able to selectively target many EGFR mutations, from classical mutations with or without T790M to L858R/T790M/C797S and del19/T790M/C797S triple mutations, patients harboring exon 20 insertions rarely show any response to these TKIs[29,191–196]. Treatment for these patients using clinically available TKIs is generally associated with a low progression-free survival (PFS) and objective response rate (ORR)[32,197]. Therefore, there has been an urgent unmet need for a potent and safe treatment against EGFR exon 20 insertions, given that they are the third most common EGFR mutations in NSCLC. As a result, a variety of drugs, including TKIs and monoclonal antibodies, are under preclinical investigations or have entered clinical trials to overcome these insertions, with TAK-788 (mobocertinib, or formerly AP32788) and amivantamab (Rybrevant, or formerly JNJ-61186372) recently granted accelerated approval by the FDA to treat metastatic NSCLC with EGFR exon 20 insertions[27,191,198–202].

Poziotinib, a covalent EGFR and HER2 inhibitor, is one of the first inhibitors reported to have activity against exon 20 insertions and is currently in a phase II clinical trial to evaluate its

efficacy and safety to treat NSCLC patients with EGFR or HER2 exon 20 insertions[198,203,204]. It is a potent inhibitor that shows strong inhibition of exon 20 insertions in preclinical studies, but it also strongly inhibits the wild-type EGFR, which leads to serious toxicities seen in vast majority of patients in clinical trials[193,198]. As a result, the FDA has recently rejected its approval for treating patients harboring HER2 exon 20 insertions, citing its high level of toxicities with a Complete Response Letter (CRL) to request a different randomized controlled study. Although the phase II clinical trial against EGFR exon 20 insertions is still ongoing, Spectrum Pharmaceuticals, the licensee of poziotinib in the U.S., announced that the poziotinib program would be urgently de-prioritized following the FDA's CRL[205,206]. Notably, in 2018, the FDA already refused to grant breakthrough therapy designation to poziotinib for treating patients with EGFR exon 20 insertions due to safety concerns[192,193,207]. Similar to poziotinib, BDTX-189 (tuxobertinib), a structural analog of neratinib previously in a clinical trial for exon 20 insertions, is a potent irreversible EGFR/HER2 inhibitor that can target both EGFR and HER2 exon 20 insertions, but has serious toxicity issues, which led to discontinuation of the trial and further development[208,209].

First reported in 2016, TAS6417 (CLN-081, TPC-064, or zipalertinib) is also an ATP-competitive covalent inhibitor that binds to C797[200,210]. Its preclinical and initial clinical results have demonstrated that it is a pan-mutation-selective EGFR TKI effective against most clinically relevant EGFR mutations with a wide therapeutic window, which leads to the FDA's issuance of breakthrough therapy designation for NSCLC earlier this year[200,211,212]. Similar to TAS6417, Sunvozertinib (DZD9008) is another recently developed covalent EGFR TKI that received breakthrough therapy designation from the FDA earlier this year[213–215]. It is currently still in a phase I/II clinical trial for EGFR or HER2 mutations (exon 20 insertions included), but its preliminary result in patients with EGFR exon 20 insertions was promising, with ORR around 41% and frequency of TRAE around 33%[213].

Since all of the EGFR exon 20 insertion inhibitors mentioned above are covalent TKIs, almost inevitably resistance mutations like T790M or C797S would eventually appear after extensive treatment of these inhibitors[198,202,216]. As a result, a reversible TKI that could selectively target exon 20 insertions is highly desirable, even better if its potency remains in the presence of T790M. To this end, Bayer has developed reversible compounds that selectively inhibit EGFR exon 20 insertions, and BAY 2927088 is currently in phase I clinical trial to evaluate its safety and preliminary anti-tumor activity[217,218]. The chemical structure of the Bayer clinical candidate, BAY 2927088, has not been disclosed. However, after reviewing the patent they filed, the Example 33 (BAY-33) stood out, so we synthesized it in-house and included in my studies[219]. Later, the chemical structure of one of their preclinical candidate, BAY 2476568 (BAY-568), also became publicly available[218] (Figure 1.6).


1st Gen
2nd Gen
gefitinib
lapatinib
afatinib
neratinib
3rd Gen
osimertinib
olmutinib
WZ4002
naquotinib
mavelertinib
nazartinib

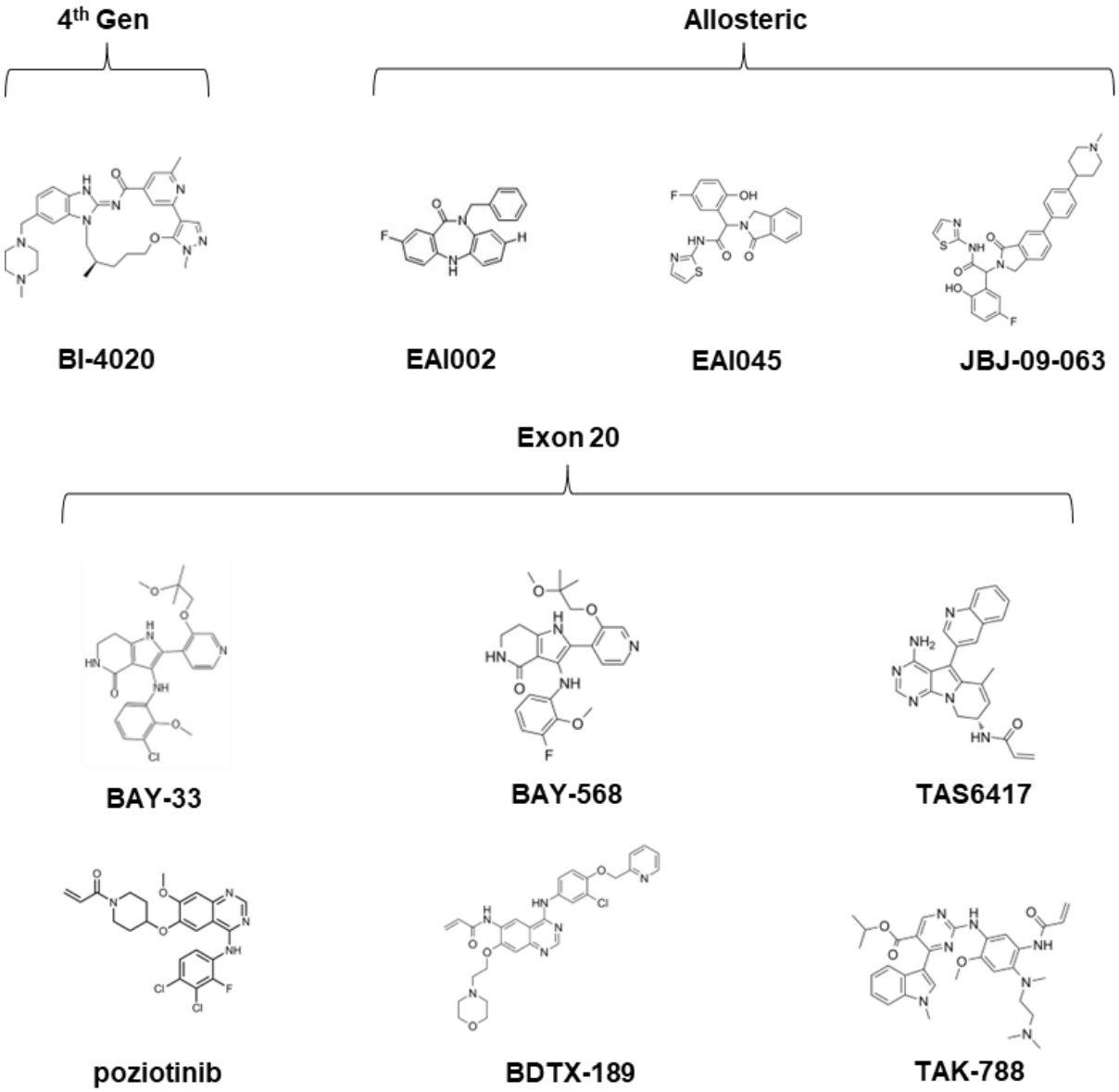


Figure 1.6. Chemical structures of selected EGFR inhibitors. Inhibitors grouped by generation, EGFR allosteric inhibitors, and inhibitors currently in clinical trials or approved for treatment of NSCLC patients with EGFR exon 20 insertions.

1.3. Significance and challenges of targeting EGFR exon 20 insertion mutants

Heterogeneity of EGFR exon 20 insertions

EGFR exon 20 insertions are heterogeneous, with insertions of 1-7 residues found in the region from D761 to C775[192,220]. At least 53 distinct insertion variants have been detected so far, and most incidences (>90%) have insertions after the C-helix in the αC-β4 loop[29,220] (Figure 1.7). As the two most frequent exon 20 insertions, V769_D770insASV (ASV inserted between V769 and D770) and D770_N771insSVD (SVD inserted between D770 and N771) account for about 40% of patients harboring insertions in exon 20[202,221]. Clinical data indicate that they are both resistant to the 1st–3rd generation TKIs[146,153,195,197]. Previously, the kinase domain of EGFR D770_N771insNPG, a rare exon 20 insertion mutant, was characterized biochemically and structurally in our lab. The insNPG mutant shares similar TKI resistance with the insASV and insSVD mutants *in vitro*[128]. In addition, since their insertion sites are either identical or only one amino acid away, it suggests that these mutants might have similar biochemical and structural properties. Notably, the A763_Y764insFQEA mutant, unlike most other insertion mutants, was found sensitive to the 1st generation TKIs, including gefitinib and erlotinib, implying that the location of an insertion is important for its TKI sensitivity[128,192,222]. This hypothesis was later supported by other studies showing location-dependent sensitivities to poziotinib[194,198,223].

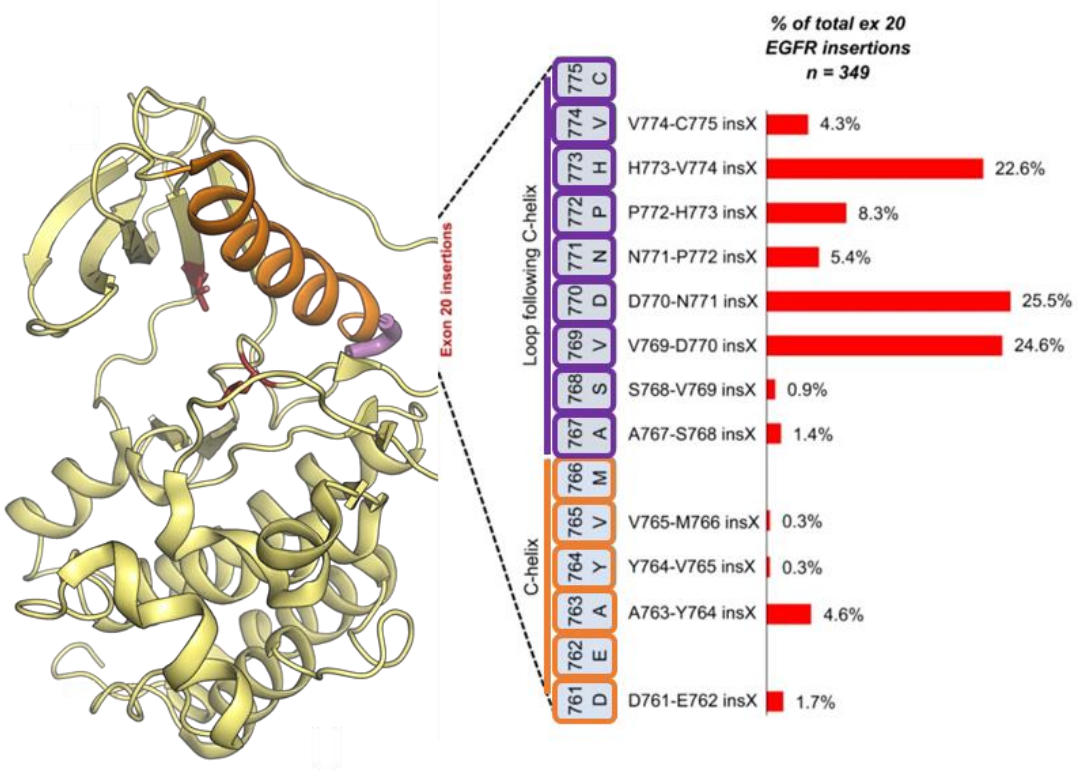


Figure 1.7. Heterogeneous EGFR exon 20 insertion variants. The crystal structure of insNPG is shown with L858R and T790M highlighted in red to indicate their relative positions to the insertion site (PDB: 4LRM). The C-helix is highlighted in orange while the loop following C-helix is highlighted in magenta. Insertion sites are widely distributed in the C-helix and the loop following C-helix. Most of the insertion sites are located in the loop following C-helix, with V769_D770, D770_N771, and H773_V774 being the three most common insertion sites (Vyse et al. 2019[192]).

Structure of EGFR exon 20 insertion kinase domain

The two classical EGFR mutations, L858R and exon 19 deletions, activate the kinase by destabilizing the inactive conformation, which shifts the conformational dynamics towards the active conformation[16,128,224]. Specifically, in the case of L858R, it is highly unfavorable for the inactive conformation to adopt the large and basic arginine in its hydrophobic core, where the N-terminal portion of the activation loop forms a short helical turn that packs with the C-helix (Figure 1.8, A). In the active conformation, the activation loop is extended, which makes the basic R858 favorably exposed to solvent (Figure 1.8, B). Therefore, oncogenic activation of the L858R mutant relies on heavily shifted dynamics towards the active state[16]. Similarly, the exon 19 deletions truncate the αC-β3 loop, which makes the "C-helix-out" inactive conformation highly unfavorable, thus promoting the active conformation[16,224]. Intuitively, one would expect the exon 20 insertions to activate the EGFR kinase in a similar fashion by inserting residues from the other end of the regulatory C-helix. However, it would be difficult to confirm the mechanism of activation without a clinically relevant mutant structure.

Fortunately, the crystal structure of EGFR insNPG kinase domain in complex with a TKI, PD168393, was determined at a resolution of 3.5Å (PDB ID: 4LRM) (Figure 1.9, A). In the crystal structure, the C-helix of the mutant EGFR kinase is observed in the inward active conformation[128]. Compared to the active conformation of wild-type EGFR kinase domain (PDB ID: 1M17), the active conformation of insNPG has an extra-long αC-β4 loop due to the inserted residues (Figure 1.9, B). Other than the inserted residues, the insNPG structure aligns almost perfectly with the wild-type EGFR, raising the question of whether a selective inhibitor of the mutant is possible.

However, the insertion mutation may alter the dynamics of the kinase in a manner that affects inhibitor binding. The additional loop structure formed by the inserted residues could hinder

the repositioning of C-helix, which makes it difficult for the mutant kinase to adopt the inactive conformation (Figure 1.9, C-D). In addition, Y827 hydrogen bonds with the carbonyl of V769 and the amino of the inserted glycine at the end of C-helix, which is not seen in the active wild-type structure, suggesting a more stabilized active conformation in the mutant kinase (Figure 1.9, B). The steric hindrance and additional hydrogen bonding interactions at the end of C-helix could limit its reorientation, which may in turn control the access of TKI to the adjacent active site, leading to TKI resistance in the mutant kinase.

In 2021, Gonzalvez et al. published a crystal structure of EGFR insNPG/V948R kinase in complex with osimertinib (PDB ID: 7LGS)[201]. As a dimer-disrupting mutation, V948R would force the EGFR kinase to be in the monomeric form, contrary to its active dimeric state[70]. Unlike all other EGFR V948R mutant kinases, the insNPG/V948R kinase is in the active conformation, which further suggests an altered equilibrium between active and inactive states for insNPG. In the monomeric state, EGFR insNPG kinase adopts the active conformation, rather than the inactive conformation adopted by wild-type EGFR.

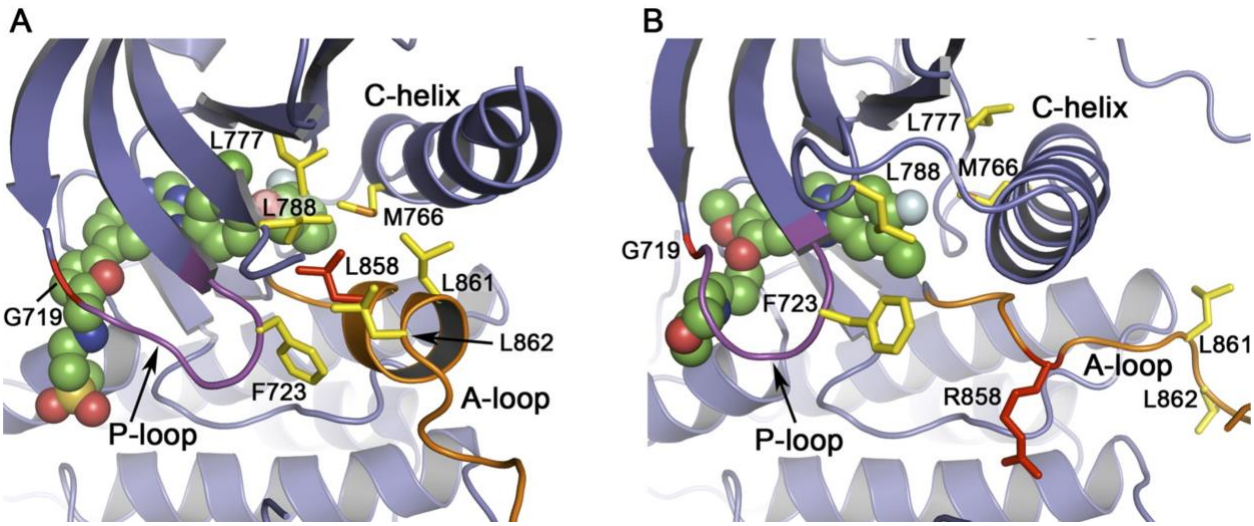


Figure 1.8. Mechanism of activation for L858R. A, The crystal structure of the inactive wild-type EGFR kinase in complex with lapatinib. L858 is highlighted in red, its surrounding hydrophobic residues are highlighted in yellow, and the N-terminal portion of the activation loop is highlighted in orange. **B,** The crystal structure of the active L858R kinase in complex with gefitinib (Yun et al. 2007[16]).

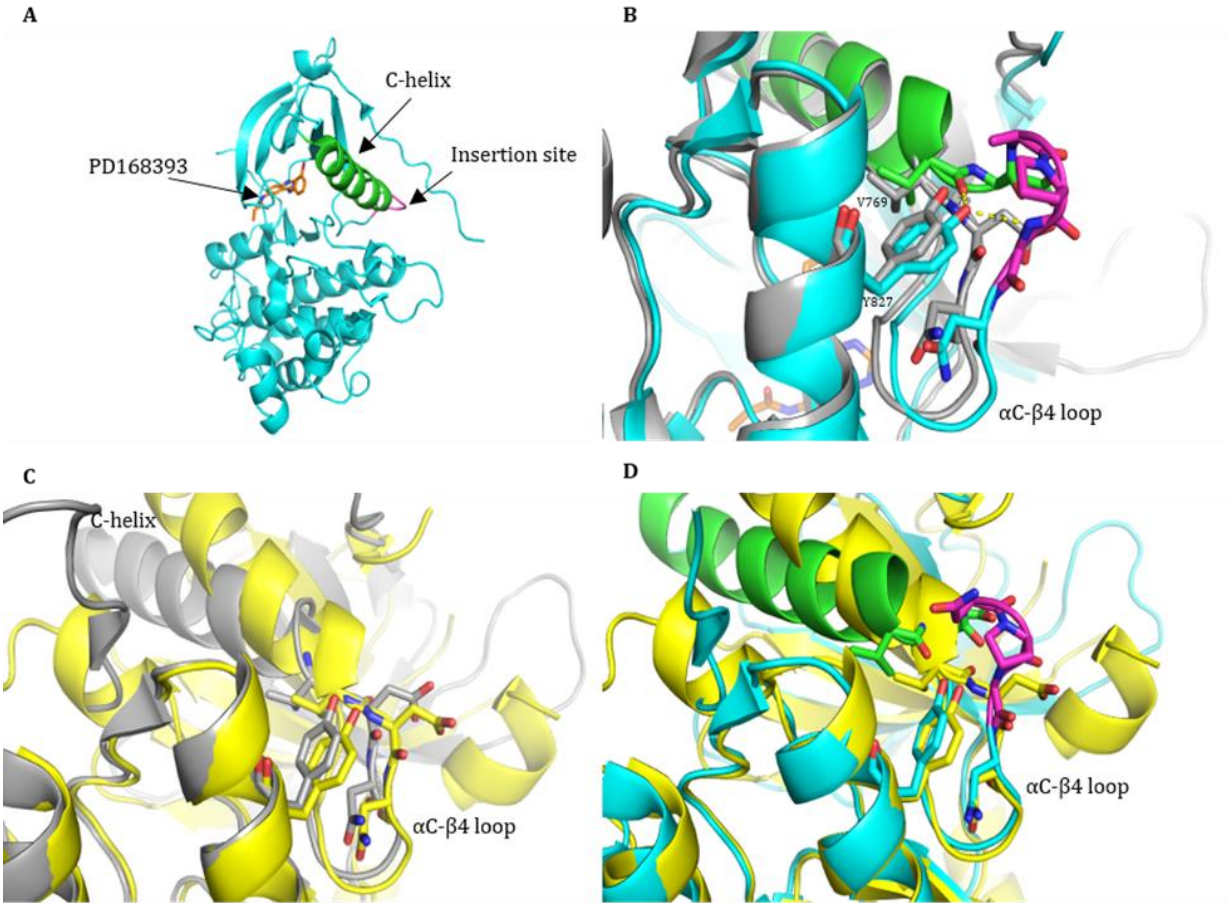


Figure 1.9. Steric hindrance and additional hydrogen bonding interactions at the end of C-helix might alter conformational dynamics of exon 20 insertion mutant kinases. A, Crystal structure of insNPG kinase domain in complex with PD168393, which occupies the ATP site. **B,** Structures of insNPG superimposed with wild-type (colored in gray), both in active conformation. **C and D,** Structures of active conformation of wild-type and insNPG, respectively, superimposed with inactive conformation of wild-type (colored in yellow).

Current treatment options for EGFR exon 20 insertions

EGFR exon 20 insertion mutations represent a unique set of EGFR-activating mutations, which previously have been described as insensitive to clinically available EGFR TKIs[128,197,225]. Even though osimertinib has been used as the preferred first-line therapy for most EGFR mutations, it shows limited success against exon 20 insertions[193]. As a result, standard chemotherapy had still been considered the frontline treatment for this subgroup of EGFR mutations until recently[221].

Amivantamab, a bispecific antibody targeting EGFR and mesenchymal-epithelial transition factor (MET, or hepatocyte growth factor receptor), received accelerated approval from the FDA for treatment of NSCLC patients harboring exon 20 insertions in May 2021[226]. However, it is unlikely to be mutant selective due to its mechanism of action. Amivantamab is produced in cell lines with low fucose Fc production, which enhances the binding of its Fc region to Fc gamma receptors on immune cells[227,228]. The enhanced binding leads to stronger target cell elimination through antibody-dependent cell-mediated cytotoxicity[228,229] (ADCC). In addition to preventing dimerization by binding to the receptor, Amivantamab also eliminates cancer cells by inducing ADCC. This mechanism does not differentiate cancer cells from normal, because wild-type EGFR is abundant in our normal cells and has the same extracellular domain as the mutant EGFR[227]. Likewise, the activities of EGFR exon 20 insertion mutants do not rely on dimerization, which explains why they are not sensitive to other monoclonal antibodies like cetuximab[115]. As a result, amivantamab has a narrow therapeutic window, which limits its efficacy[230]. This is consistent with the clinical data so far; amivantamab treatment against exon 20 insertions has an ORR of only 40% and treatment-related severe (grade ≥3) side effects led to dose interruptions in 78% of patients[226]. Although it is better than most of the known TKIs against EGFR exon 20 insertions, it

is still not good enough if compared to the ORR of 87% and frequency of dose interruptions of 37% when using osimertinib as first-line treatment against classical mutations[157].

As a newly developed EGFR TKI, TAK-788 also received accelerated approval from the FDA last year[231], but 40% of patients in the clinical trials experienced treatment-related severe side effects, possibly due to its strong potency against wild-type EGFR[232]. Many patients required dose reductions (25%) or even discontinuation (17%), which contributed to the relatively low ORR of around 28%[232]. Around 10% of patients had TRAEs leading to death, further questioning its potential safety issues[233]. In addition, like most other EGFR exon 20 insertion inhibitors, TAK-788 is a covalent ATP-competitive TKI, which makes it vulnerable to the acquired resistance mutation C797S (Figure 1.6). Therefore, it is important to find an inhibitor, a reversible one preferred, with a greater mutant-selectivity to further widen the therapeutic window.

Chapter 2

Biochemical characterization of EGFR insASV and insSVD

2.1: Introduction

EGFR mutations have been well characterized over the last two decades[16,127,174]. However, as the third most common EGFR mutations, exon 20 insertions are still poorly understood, with limited biochemical characterization available. Although EGFR exon 20 insertions are heterogeneous, >90% of insertions occur in the αC-β4 loop following the C-helix[29]. Among these insertions, EGFR insASV and insSVD are the two most common ones (Figure 2.1, A), together representing >40% of all these insertions[234]. However, despite their high prevalence, there has not been any biochemical characterization of these mutants yet.

EGFR exon 20 insertions are known to be oncogenic, but there are many questions remain to be answered. How much more active are exon 20 insertion mutants compared to wild-type EGFR? What causes oncogenic activation of these mutants? To answer these questions, it is important to study these insertion variants on a molecular level. Previously, EGFR insNPG, a rare exon 20 insertion variant, was biochemically and structurally characterized (Figure 2.1, B), along with some cellular potency measurement of 1st generation EGFR tyrosine kinase inhibitors (TKI) against representative EGFR exon 20 insertion variants[128]. However, EGFR exon 20 insertions are heterogeneous, and no biochemical characterization or systematic inhibitor potency measurement has been performed on the two most common insertion variants, EGFR insASV and insSVD. Therefore, I started my project by making efforts to optimize expression and purification of the kinase domain of these two most prevalent EGFR exon 20 insertion mutants, hoping to examine their catalytic activities to compare them with wild-type EGFR as well as other clinically relevant EGFR mutants.

In addition to examining catalytic activities of these insertion variants, it is also very important to measure $K_{m,\ ATP}$ values to investigate their affinities to ATP. As outlined in the previous chapter, ATP affinity (or more precisely $K_{m,\ ATP}$) can have a major effect on drug sensitivity of EGFR mutants. Classical EGFR mutants are inherently sensitive to the 1st generation EGFR TKIs due to their weaker affinities to ATP compared to wild-type EGFR, and the gatekeeper mutation T790M confers resistance to 1st generation TKIs by restoring ATP affinities to similar levels with the wild-type EGFR[16,127]. Notably, even though most exon 19 deletions, such as ΔE746-A750, are sensitive to 1st generation TKIs, some less prevalent exon 19 deletions, such as ΔL747-E749, are resistant to 1st and 3rd generation TKIs, with 2nd generation TKIs potent but not mutant selective. The difference in sensitivities of these exon 19 deletions to 1st generation TKIs turned out to result from marked differences of their ATP affinities[224]. Therefore, it is significant to carefully measure the $K_{m,\ ATP}$ values of these insertion variants and compare them to that of the wild-type EGFR as well as other EGFR mutants to understand to what extent this property may affect drug sensitivity.

In this chapter, I will cover three topics. First, I will describe the design of a series of Baculogold virus/insect cell expression constructs for insASV and insSVD. Second, I will present the results of my purification optimization of buffer conditions to obtain active, soluble EGFR insASV and insSVD proteins. Finally, I will cover the enzyme kinetic characterization of EGFR insASV and insSVD kinases, and compare them to L858R, L858R/T790M, as well as the wild-type EGFR.

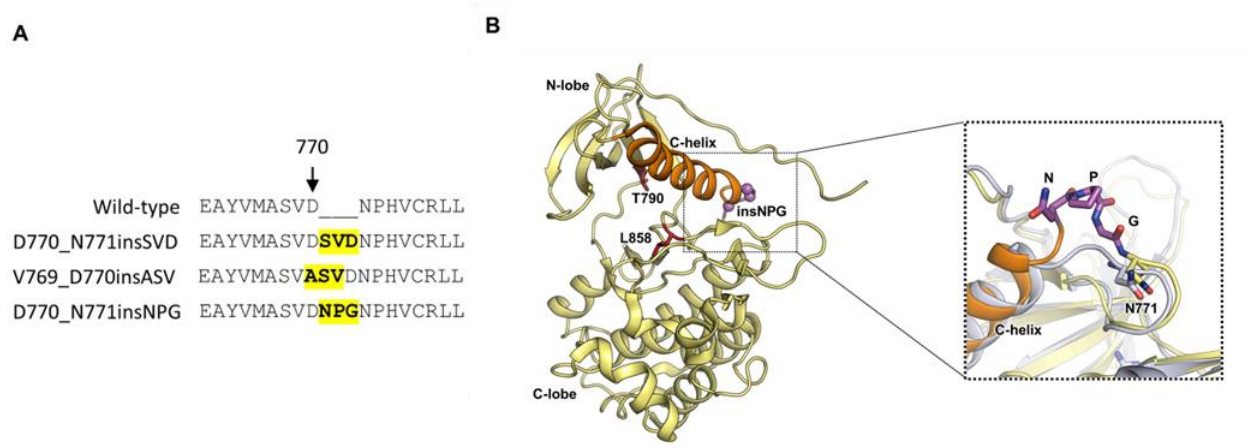


Figure 2.1. Location of EGFR exon 20 insertion. A, Primary sequences of EGFR wild-type, D770_N771insSVD, V769_D770insASV, and D770_N771insNPG are aligned with inserted residues highlighted in yellow. **B,** Crystal structure of EGFR insNPG in the active conformation (PDB: 4lrm), highlighting the location of exon 20 insertion mutations studied here and a detailed view of the NPG insertion. The insNPG structure is in yellow (C-helix highlighted in orange) and inserted residues in magenta in stick form. The structure is superimposed on the wild-type EGFR structure in gray (PDB: 2itx).

2.2: Materials and Methods

Cloning, expression, and purification of EGFR kinase domain

The EGFR kinase domain constructs (residues 696-1022) bearing the insertions V769_D770insASV or D770_N771insSVD were expressed and purified in Sf9 insect cells using the Baculogold system (Pharmingen). Insertions were first introduced into the human wild-type EGFR sequence in the context of the pAC8NKRed vector using the QuickChange XL site-Directed Mutagenesis Kit (Stratagene) or using gene blocks ordered from IDT DNA. The resulting constructs were confirmed by Sanger sequencing and were expressed as N-6X His, N-Strep, C-Strep, N-10X His C-TwinStrep, or N-TwinStrep C-10X His fusion proteins (Figure 2.2, A). EGFR L858R, L858R/T790M, T790M/V948R, D770_N771insNPG and wild-type EGFR with construct residues 696-1022 were expressed as N-6X His-GST fusion proteins using the pcDNA3.1 expression vector (Invitrogen)[121,128,235].

The pAC8NKRed vector used in the cases of insASV or insSVD and the pcDNA3.1 vector in the cases of L858R, L858R/T790M, T790M/V948R, insNPG or wild-type EGFR are transfer plasmids, with the polyhedrin promoter (PH prom in the example figure) flanked by Autographa californica multiple nucleopolyhedrovirus (AcMNPV) DNA on both sides. To utilize insect cells for protein production, the transfer plasmid containing the EGFR construct was co-transfected with linearized baculovirus AcMNPV DNA, which is lack of the polyhedrin gene that is essential for infection[236]. Then homologous recombination between the transfer plasmid and linearized viral DNA made the virus contain the EGFR construct and the polyhedrin gene, which enables the virus for infection[237]. Subsequently, the virus was amplified by passaging through four rounds of infection cycles with *Spodoptera frugiperda* (Sf9) insect cells for a higher viral titer stock. For

protein expression, the viral stock was added to a suspension of Sf9 insect cells in a 1:100 volume ratio when the insect cells grew to a density of 2 million cells per mL. After 2-3 days, the cells were harvested using centrifuge when they looked infected and cell viability dropped to 90%.

Cell pellets of insASV or insSVD were lysed by sonication in 50 mM Tris-HCl (pH 8.0), 500 mM NaCl, and 20% glycerol supplemented with 1 mM TCEP [tris(2-carboxyethyl) phosphine hydrochloride], 20 mM imidazole, and complete protease inhibitor mixture (Roche). After centrifugation (40,000 rpm, 2 hr, 4 °C), the supernatant was loaded to a Ni-NTA column with 1mL of beads and flowed through it twice. The beads were washed with 100 mL wash buffer (50 mM Tris-HCl, 500 mM NaCl, 20% Glycerol, 50 mM imidazole, 1 mM TCEP [pH 8.0]) and then eluted with two fractions of 5 mL elution buffer (50 mM Tris-HCl, 500 mM NaCl, 20% Glycerol, 200 mM imidazole, 1 mM TCEP [pH 8.0]). The eluted EGFR kinase proteins (not concentrated) were further purified by size-exclusion chromatography (Superdex 200) with 120 mL of bed volume in SEC buffer (50 mM Tris-HCl, 500 mM NaCl, 20% Glycerol, 1 mM TCEP [pH 8.0]) with a flow rate of 0.5 mL/min and were concentrated by centrifuging to 30 μg/mL (insASV) or 80 μg/mL (insSVD), and aliquots were flash frozen in liquid nitrogen and stored at −80 °C. EGFR WT, L858R, L858R/T790M, T790M/V948R and insNPG were purified in a similar way except that the glycerol concentration in the purification buffers was 5% instead of 20%[127,128].

Streptavidin purification buffer conditions:

1. Lysis/Wash buffer: 100mM Tris-HCl (pH 8.0), 150 mM NaCl, 5% glycerol, 1mM TCEP
2. Strep Elution buffer: 100mM Tris-HCl (pH 8.0), 150 mM NaCl, 5% glycerol, 5mM desthiobiotin, 1mM TCEP

Ni-NTA purification buffer conditions:

1. Lysis/Ni-NTA Binding buffer: 1X TBS with 1 mM TCEP, 10% Glycerol
2. Wash buffer: 50mM Tris-HCl (pH 8.0), 500 mM NaCl, 1 mM TCEP, 5% Glycerol
3. Imidazol wash buffer: 50mM Tris-HCl (pH 8.0), 500 mM NaCl, 1 mM TCEP, 5% Glycerol, 20 mM Imidazole
4. Ni-NTA Elution buffer: 50mM Tris-.-HCl (pH 8.0), 500 mM NaCl, 1 mM TCEP, 5% Glycerol, 500 mM Imidazole

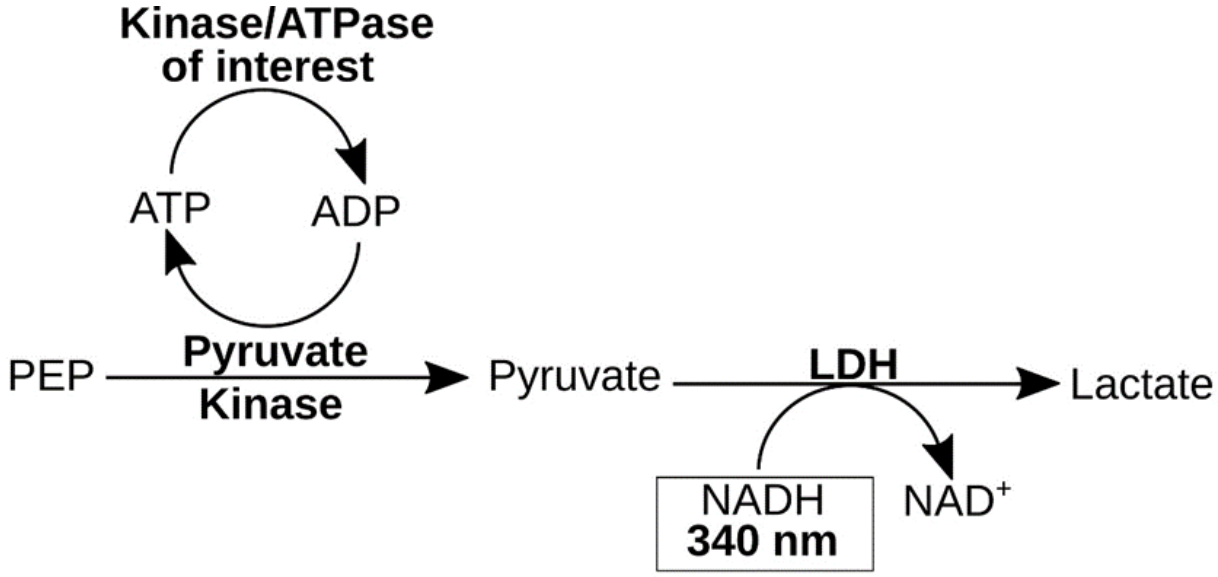
EGFR enzyme kinetic assays

Kinetic parameters were determined using a continuous coupled enzyme assay in which pyruvate kinase (PK) and lactate dehydrogenase (LDH) are used to couple production of ADP to oxidation of NADH to NAD^+. In this ATP/NADH coupled enzyme assay, ATP consumption is coupled with dephosphorylation of phosphoenolpyruvate (PEP) by PK, which generates pyruvate and converts ADP back to ATP. Pyruvate is then converted to lactate by LDH, which is accompanied by oxidation of NADH to NAD^+. Since NADH shows strong absorbance at wavelength of 340 nm, the ATP consumption by EGFR could be correlated to the signal reduction of NADH, which indirectly monitors the reaction progress of EGFR[238,239] (Scheme 2.1).

In my experiments, a 96-well plate was used for the ATP/NADH coupled assay system. The reaction mixture contained 0.5 mg/mL BSA, 2 mM $MnCl_2$, 1 mM phosphoenolpyruvate (PEP), 1 mM TCEP, 1 mM ATP, 0.1 M HEPES 7.5, 1/50 volume of PK/LDH, 0.5 mM NADH, 1 mM tyrosine kinase substrate poly-[Glu_4Tyr], in addition to the indicated EGFR variant[16]. WT or mutant EGFR kinases were used at the following concentrations: 30 nM (L858R), 60 nM (L858R/T790M),

500 nM (insSVD), 100 nM (insASV), and 1 µM (WT). ATP at the indicated concentrations were added last to start the reaction. Steady state initial velocity data were drawn from the slopes of the A_{340} curves and fit to the Michaelis-Menten Equation in GraphPad Prism to determine v_{max} and K_m values. The reduction of absorbance at 340 nm was converted to µM using the initial concentration of NADH in the reaction mixture. Then k_{cat} values were calculated using v_{max} and the active enzyme concentration. The active enzyme concentration of each EGFR variant was determined by titration with the irreversible inhibitor poziotinib and then expected enzyme concentrations (determined by A_{280} using NanoDrop) were adjusted based on those values[16] (Figure 2.5).

Scheme 2.1. Illustration of the ATP/NADH coupled enzyme assay system (McFarlane et al. 2020[239])



2.3: Results

EGFR inASV and insSVD construct design

A key obstacle in biochemical characterization of EGFR insASV and insSVD has been difficulty in obtaining sufficient quantities of the mutant kinase domains in soluble form. To overcome this challenge, I created a total of 30 constructs spanning different residue ranges in the cytoplasmic domain of EGFR (Table 2.1) and tested multiple purification tags, insect cell infection conditions, and purification buffer compositions. I chose six different amino acid ranges: 669-985, 669-1022, 669-1026, 696-985, 696-1022, and 696-1026. Starting at the position 669 includes the juxtamembrane segment, which is known for promoting and stabilizing the formation of the asymmetric dimer, while starting at the position 696 is commonly used in previous EGFR kinase constructs in the literature[74,240]. Ending at the position 985 minimizes the C-terminal tail included in the construct, while ending at the position 1026 includes four more residues than the traditionally used position 1022 to extend the C-terminal tail in the construct.

There are also five different purification tag systems examined in my project: N-6X His, N-Strep, C-Strep, N-10X His C-TwinStrep, and N-TwinStrep C-10X His. N-6X His is one of the most widely used tag system in recombinant protein purification, and N-Strep or C-Strep is generally associated with high purity of functional proteins since purification takes place under moderate conditions[241–243]. N-10X His C-TwinStrep and N-TwinStrep C-10X His are both very high affinity double-tag systems, with 10X His and TwinStrep being the enhanced versions of 6X His and Strep, respectively[242,244,245]. These two tag systems were designed to significantly increase the purity of protein yield.

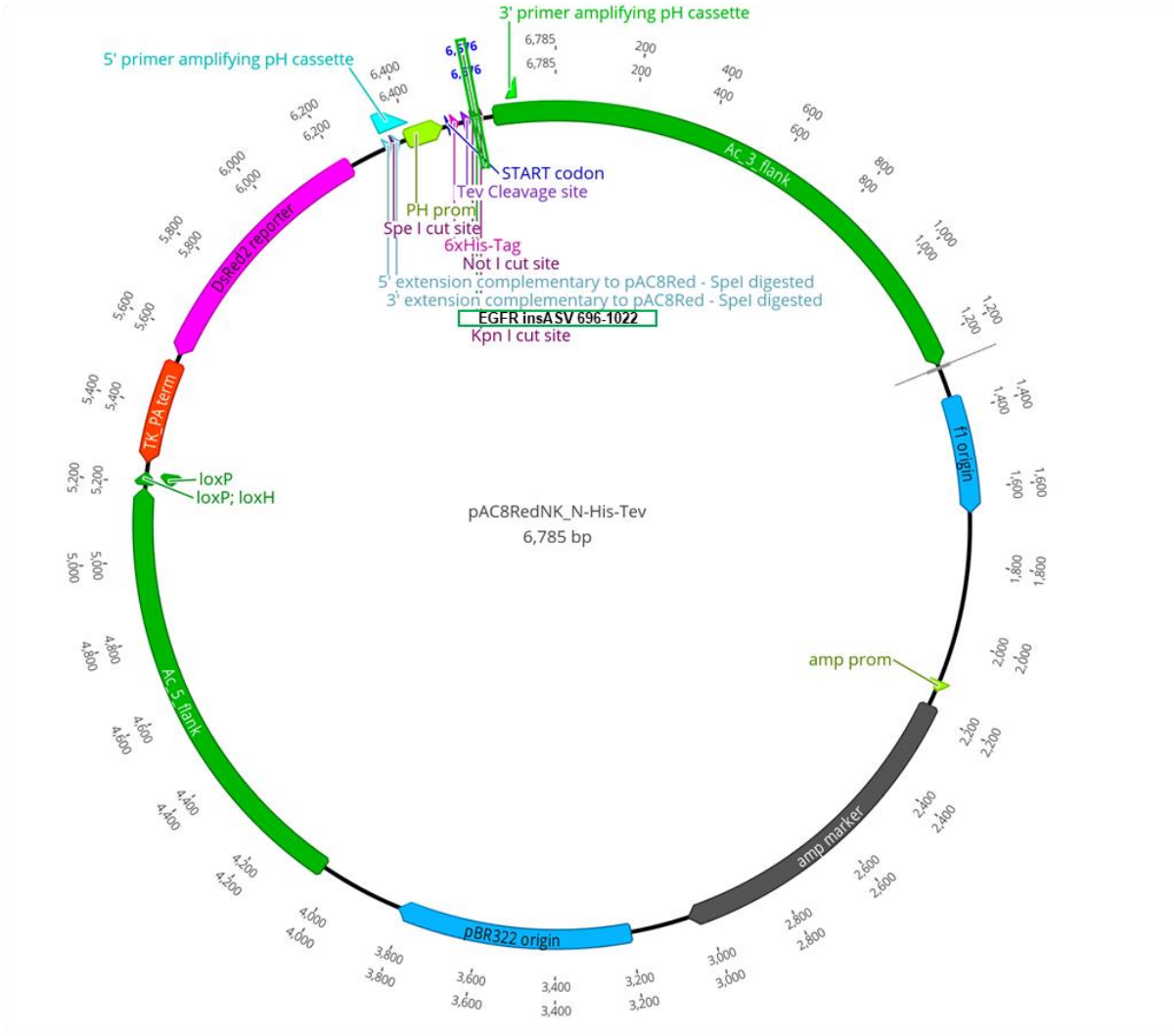
Table 2.1

Constructs for EGFR insASV

	Vector	Construct	Expression Level
1	pAC8NKRed N-His	EGFRinsASV 669-1026	Poor
2	pAC8NKRed N-Strep	EGFRinsASV 669-1026	Moderate
3	pAC8NKRed N-Strep	EGFRinsASV 669-985	Moderate
4	pAC8NKRed N-His	EGFRinsASV 669-985	Moderate
5	pAC8NKRed N-Strep	EGFRinsASV 669-1022	Good
6	pAC8NKRed N-His	EGFRinsASV 669-1022	Good
7	pAC8NKRed N-Strep	EGFRinsASV 696-985	Moderate
8	pAC8NKRed N-His	EGFRinsASV 696-985	Moderate
9	pAC8NKRed N-Strep	EGFRinsASV 696-1022	Good
10	pAC8NKRed N-His	EGFRinsASV 696-1022	Excellent
11	pAC8NKRed N-Strep	EGFRinsASV 696-1026	Good
12	pAC8NKRed N-His	EGFRinsASV 696-1026	Good
13	pAC8NKRed C-Strep	EGFRinsASV 669-985	Good
14	pAC8NKRed C-Strep	EGFRinsASV 669-1022	Good
15	pAC8NKRed C-Strep	EGFRinsASV 669-1026	Moderate
16	pAC8NKRed C-Strep	EGFRinsASV 696-985	Moderate
17	pAC8NKRed C-Strep	EGFRinsASV 696-1022	Excellent
18	pAC8NKRed C-Strep	EGFRinsASV 696-1026	Good
19	pAC8NKRed N-TwinStrep C-10X His	EGFRinsASV 669-985	Moderate
20	pAC8NKRed N-TwinStrep C-10X His	EGFRinsASV 669-1022	Poor
21	pAC8NKRed N-TwinStrep C-10X His	EGFRinsASV 669-1026	Moderate
22	pAC8NKRed N-TwinStrep C-10X His	EGFRinsASV 696-985	Moderate
23	pAC8NKRed N-TwinStrep C-10X His	EGFRinsASV 696-1022	Excellent
24	pAC8NKRed N-TwinStrep C-10X His	EGFRinsASV 696-1026	Good
25	pAC8NKRed N-10X His C-TwinStrep	EGFRinsASV 669-985	Moderate
26	pAC8NKRed N-10X His C-TwinStrep	EGFRinsASV 669-1022	Poor
27	pAC8NKRed N-10X His C-TwinStrep	EGFRinsASV 669-1026	Moderate
28	pAC8NKRed N-10X His C-TwinStrep	EGFRinsASV 696-985	Moderate
29	pAC8NKRed N-10X His C-TwinStrep	EGFRinsASV 696-1022	Excellent
30	pAC8NKRed N-10X His C-TwinStrep	EGFRinsASV 696-1026	Good

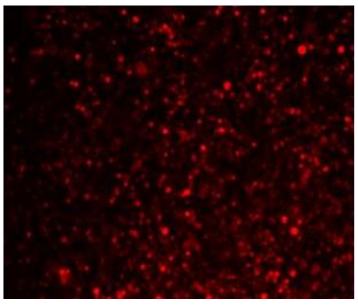
Figure 2.2

A



B

C



D

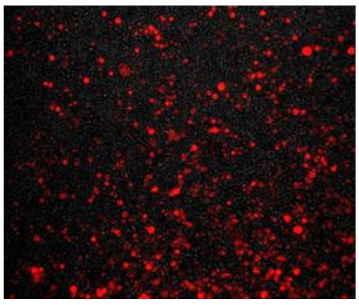


Figure 2.2. Cloning and transfection. A, Example of a pAC8NKRed vector with N-His EGFR insASV 696-1022 is shown. The pAC8NKRed vector contains a fluorescence reporter gene that expresses DsRed as a transfection marker. The vector is a transfer plasmid, with the polyhedrin promoter (PH prom in the example figure) flanked by AcMNPV DNA on both sides. **B-D,** The transfection of pAC8NKRed N-His EGFR insASV 696-1022 and linear baculoviral DNA into sf9 cells was done, and the fluorescence images were taken after 48 hours. **B,** the total cells were shown. **C,** The transfected cells were shown. **D,** The superimposion of the transfected cell image over the total cell image was displayed, which is an indication of the percentage of successfully transfected cells.

Expression and purification of EGFR insASV and insSVD

The EGFR constructs were cloned into the pAC8NKRed vector (Figure 2.2, A) and successfully co-transfected into Sf9 insect cells (Figure 2.2, B-D). Our test purifications indicated that the expression levels of EGFR insASV or insSVD proteins generally vary with different residue ranges. While tags were not as important in changing expression levels, I found that those with strep tags would aggregate on the strep-tactin beads, which significantly decreased protein yield. Among the five different residue ranges I tested, 696-1022 generated most production of EGFR insASV or insSVD proteins (Figure 2.3, A-C). Of all the constructs tested, pAC8NKRed N-His10 C-TwinStrep 696-1022 produced the best purity and yield with only Ni-NTA purification (Figure 2.3, B, D). Thus I selected this construct for large scale expression and purification of insASV and insSVD with Ni-NTA purification and further size exclusion chromatography. Initial efforts were focused on a two-step affinity purification to improve the purity of EGFR kinase, but since most of the proteins would aggregate on the strep-tactin beads, only Ni-NTA purification was performed before proceeding to the final stage of purification.

Although this construct expressed well, size exclusion chromatography showed that most of the insASV or insSVD protein was aggregated and eluted in the void volume. With extensive trials of different buffer conditions, I found that inclusion of 20% glycerol in the purification buffer was important to prevent aggregation of EGFR insASV or insSVD variants during purification (Figure 2.4). With 5% glycerol present in the purification buffer, a condition that works well for most other EGFR variants, EGFR insASV or insSVD eluted in the void volume, and showed minimal catalytic activity, indicating little presence of active EGFR insASV or insSVD kinase. Interestingly, double bands usually appeared on SDS-PAGE, which was later confirmed as phosphorylated insASV or insSVD using λ-phosphatase. Follow-up mass spectrometry showed

heterogeneous phosphorylation status of insASV or insSVD, including Y869 on the activation loop, which was shown previously not to impact EGFR kinase activity (data not shown here).

Measurement of active enzyme concentration

EGFR L858R, L858R/T790M, insNPG, as well as wild-type EGFR kinases were expressed and purified as described in section 2.2[127,128]. The active concentration of each EGFR variant was measured by titrating poziotinib, a potent covalent EGFR inhibitor[198]. For the titration, each variant with an expected concentration of 1 nM (measured by A_{280}) was incubated with increasing concentrations of poziotinib for 2 hours before initiating kinase reactions. Product formation after 20 minutes were measured and linearly fitted to derive the active enzyme concentration using the X-intercept (Figure 2.5, A-F). To confirm the active enzyme concentrations, 1 µg of each variant was run by SDS-PAGE that showed same amount (Figure 2.5, G).

Kinetic analysis of EGFR insASV and insSVD

We used a continuous coupled enzyme assay to determine kinetic parameters for these EGFR variants, as well as for WT, L858R, and L858R/T790M EGFR for comparison. This well-established assay employs PK and LDH to couple ADP production to oxidation of NADH to NAD+, which is measured as a decrease in absorbance at 340 nm[238]. We used poly(Glu_4,Tyr_1) as a peptide substrate at a concentration of 5 mM. Reaction velocities at increasing ATP concentrations were plotted to determine Michaelis-Menten kinetic parameters (Figure 2.6, A-F).

Both insASV and insSVD were highly active as compared with WT EGFR (38- and 27-fold higher than WT, respectively), and were comparable to the L858R variant (41-fold higher than WT). We observed that while insASV and insSVD showed higher $K_{m, ATP}$ values compared to wild-type EGFR, their $K_{m, ATP}$ values were lower than that of the inhibitor-sensitive L858R mutant, indicating stronger affinity to ATP (Table 2.2; Figure 2.6, B-F). The $K_{m, ATP}$ values of insASV and insSVD were similar to the previously reported value of the insNPG variant, although their catalytic rates were faster[128].

Although $K_{m, ATP}$ values for the insertion variants were increased relative to WT EGFR, their catalytic efficiencies ($k_{cat}/K_{m, ATP}$) were nevertheless 6-fold higher due to their much faster catalytic rates. Their $k_{cat}/K_{m, ATP}$ values were also ~2 times higher than the L858R mutant, due to their tighter affinities for ATP (Table 2.2; Figure 2.6, B-F).

I also measured the Michaelis constant for peptide ($K_{m, peptide}$) of each EGFR variant. Interestingly, with poly(Glu_4,Tyr_1) as the peptide substrate, wild-type EGFR and L858R mutant showed very weak affinities for the peptide ($K_{m, peptide}$[WT] >5000 μM, $K_{m, peptide}$[L858R] =3542 μM) while insASV, insSVD and L858R/T790M showed much stronger affinities to poly(Glu_4,Tyr_1) ($K_{m, peptide}$ values range from 200-500 μM). Taken together, insASV and insSVD variants display increased activity and increased ATP affinity compared to L858R variants, suggesting exon 20 insertion variants may be more difficult to inhibit using ATP-competitive inhibitors.

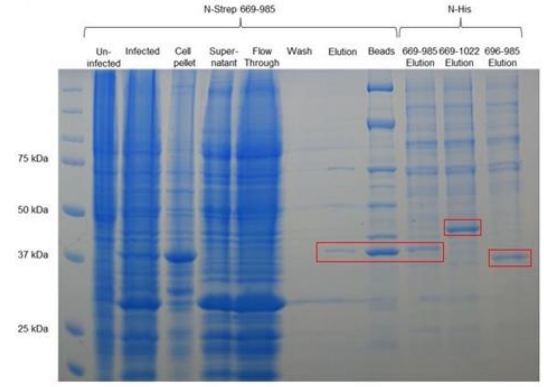
Table 2.2

Enzyme kinetic parameters of wild-type (WT) and mutant EGFR kinases

	k_{cat} (s^{-1})	Fold activity versus WT	$K_{m, ATP}$ (μM)	$k_{cat}/K_{m, ATP}$ (s^{-1} μM^{-1})	$K_{m, peptide}$ (μM)
WT	0.026	1	6±4	4.33E-3	>5000
insASV	0.994	38	39±12	2.55E-2	489±148
insSVD	0.710	27	27±7	2.63E-2	212±43
L858R	1.063	41	79±14	1.35E-2	3542±1129
L858R/T790M	0.880	34	15±7	5.87E-2	481±238

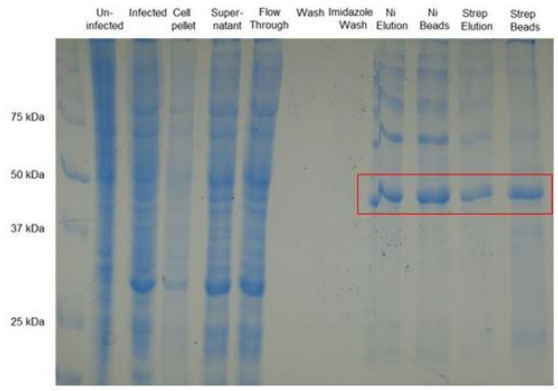
Figure 2.3

A



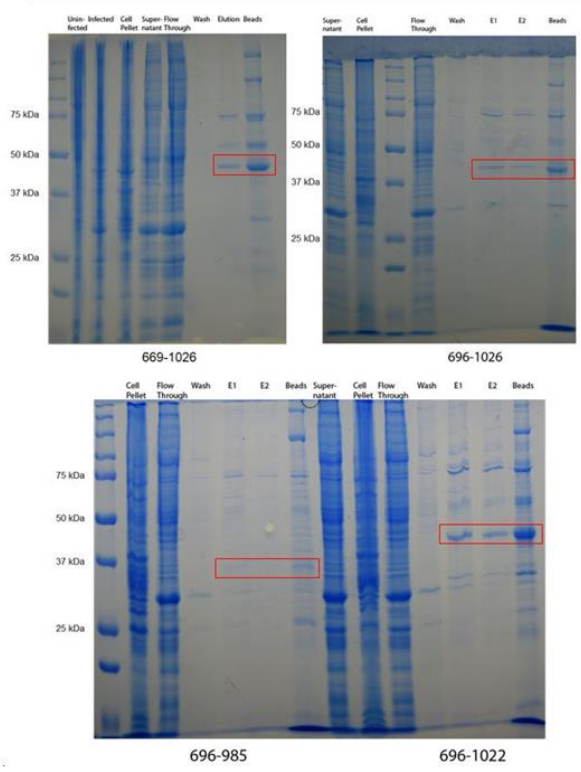
	Vector	Construct	Molecular Weight (kDa)
3	pAC8NKRed N-Strep	EGFRinsASV 669-985	39
4	pAC8NKRed N-His	EGFRinsASV 669-985	39
6	pAC8NKRed N-His	EGFRinsASV 669-1022	43
8	pAC8NKRed N-His	EGFRinsASV 696-985	36

B



	Vector	Construct	Molecular Weight (kDa)
29	pAC8NKRed N-His10 C-TwinStrep	EGFRinsASV 696-1022	44

C



	Vector	Construct	Molecular Weight (kDa)
15	pAC8NKRed C-Strep	EGFRinsASV 669-1026	43
16	pAC8NKRed C-Strep	EGFRinsASV 696-985	42
17	pAC8NKRed C-Strep	EGFRinsASV 696-1022	39
18	pAC8NKRed C-Strep	EGFRinsASV 696-1026	40

D

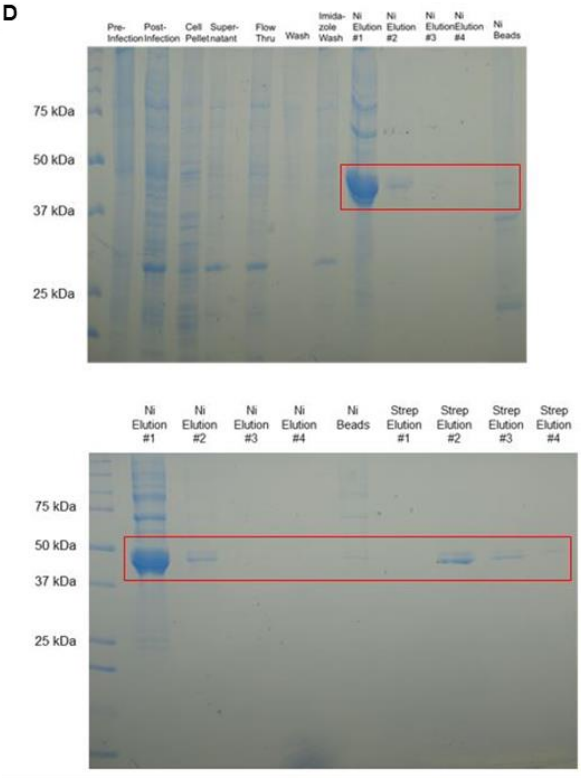


Figure 2.3. Test expression and purification of EGFR insASV kinase. Some examples of test purification results were shown. The insASV protein bands were highlighted by red boxes. **A,** SDS-PAGE of strep purification process of pAC8NKRed N-Strep EGFR insASV 669-985, in comparison of pAC8NKRed N-His EGFR insASV 669-985, 669-1022, and 696-985. **B,** SDS-PAGE of two-step purification of pAC8NKRed N-His10 C-TwinStrep EGFR insASV 696-1022. **C,** SDS-PAGE of strep purification process of pAC8NKRed C-Strep EGFR insASV 669-1026, 696-1026, 696-985, and 696-1022, respectively. E1 and E2 represent elution fraction #1, and elution fraction #2, respectively. **D,** SDS-PAGE of large scale (5L Sf9 cells) two-step purification of pAC8NKRed N-His10 C-TwinStrep EGFR insASV 696-1022.

Figure 2.4

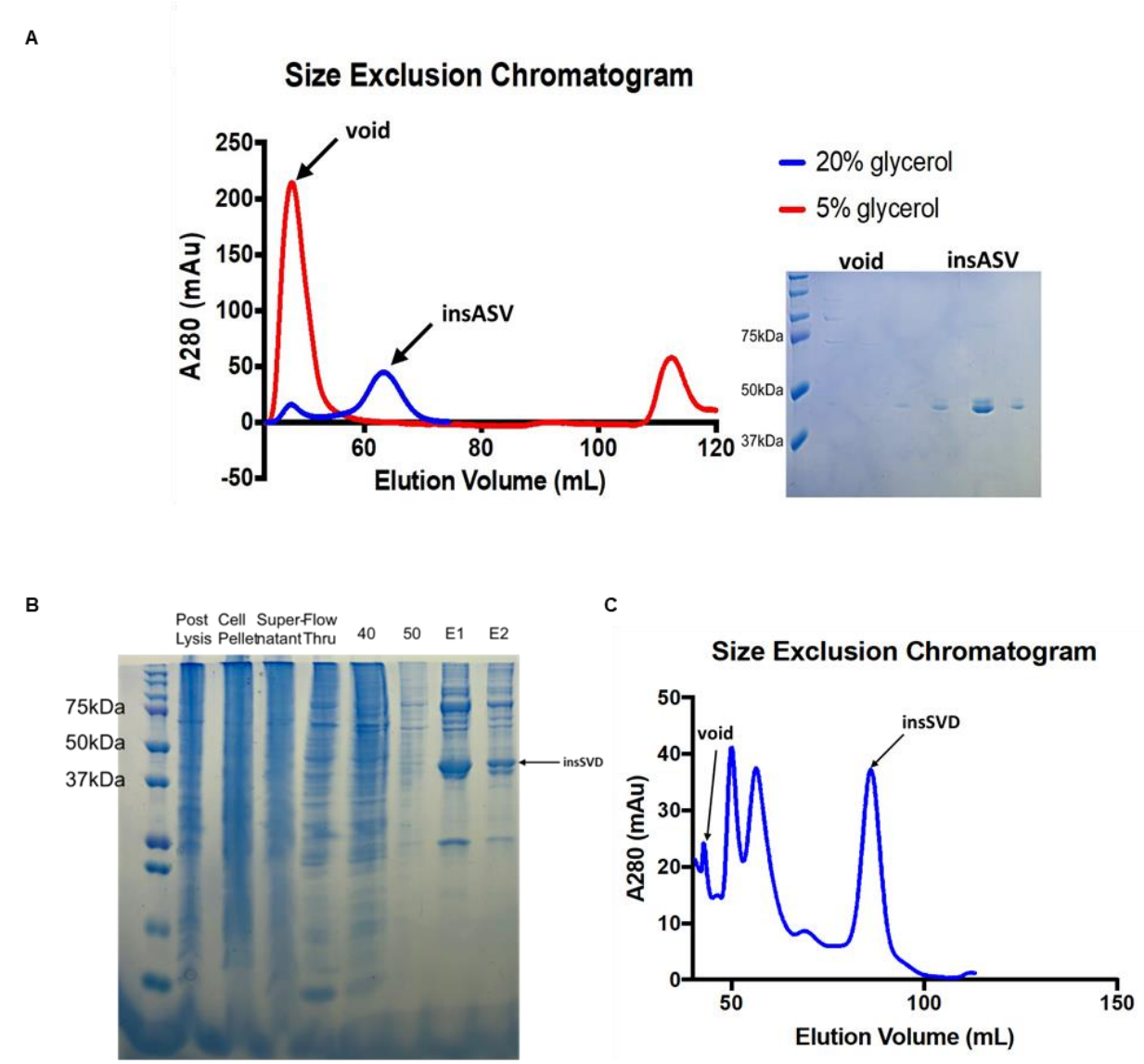

A
Size Exclusion Chromatogram
void
insASV
A280 (mAu)
Elution Volume (mL)
20% glycerol
5% glycerol
void
insASV
75kDa
50kDa
37kDa
B
Post Lysis
Cell Pellet
Super-natant
Flow Thru
40
50
E1
E2
75kDa
50kDa
37kDa
insSVD
C
Size Exclusion Chromatogram
void
insSVD
A280 (mAu)
Elution Volume (mL)

Figure 2.4. Size exclusion chromatography of purified EGFR insASV and insSVD. A, Chromatogram from size exclusion chromatography (SEC) showing insASV eluting as a monomer in the presence of 20% glycerol and aggregate in the void peak in the presence of 5% glycerol (Superdex 75 column). SDS-PAGE of SEC fractions from the 20% glycerol condition confirmed the purity of insASV. **B,** SDS-PAGE showing Ni-NTA purification of EGFR insSVD. The column was washed two times with 40 mM and 50 mM imidazole, respectively. The sample was eluted in two fractions: E1 and E2. **C,** SEC showing insSVD appeared as a monomer peak, which could be well separated from other peaks.

Figure 2.5

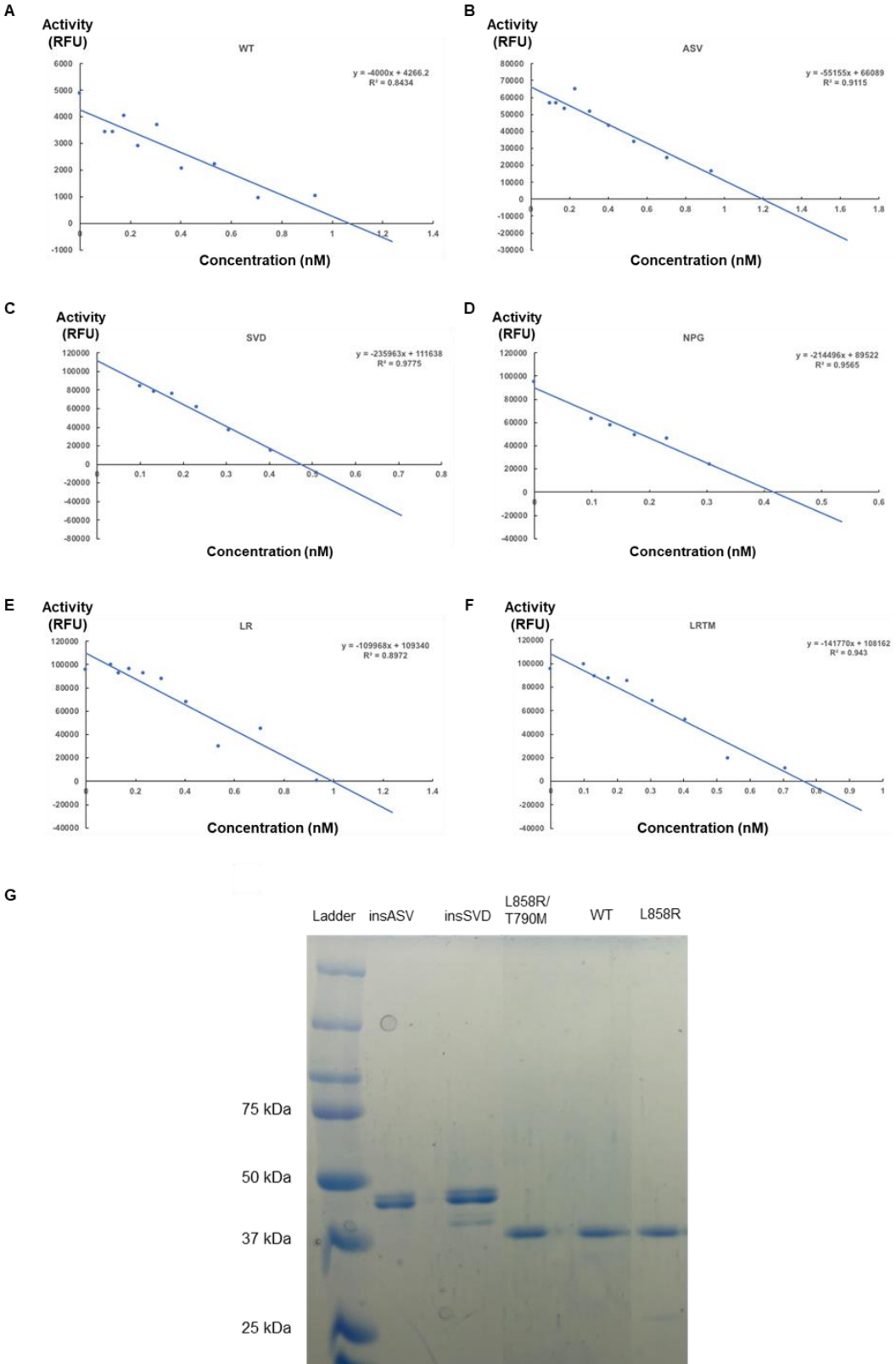

A
Activity (RFU)
WT
y = -4000x + 4266.2
R² = 0.8434
Concentration (nM)
B
Activity (RFU)
ASV
y = -55155x + 66089
R² = 0.9115
Concentration (nM)
C
Activity (RFU)
SVD
y = -235963x + 111638
R² = 0.9775
Concentration (nM)
D
Activity (RFU)
NPG
y = -214496x + 89522
R² = 0.9565
Concentration (nM)
E
Activity (RFU)
LR
y = -109968x + 109340
R² = 0.8972
Concentration (nM)
F
Activity (RFU)
LRTM
y = -141770x + 108162
R² = 0.943
Concentration (nM)
G
Ladder
insASV
insSVD
L858R/ T790M
WT
L858R
75 kDa
50 kDa
37 kDa
25 kDa

Figure 2.5. Determination of active enzyme concentrations. A-F, Plots of covalent titration of different EGFR variants using a potent irreversible inhibitor, poziotinib. X-axis shows poziotinib concentration used for each data point, while Y-axis shows the fluorescence signal generated by formation of the product. The amount of product formation when different concentrations of poziotinib are used could be linearly fitted to derive the active enzyme concentration in the solution using the X-intercept. For all the enzymes used, the expected concentration is 1 nM based on A_{280}. **A,** Wild-type EGFR is titrated by incubating with different concentrations of poziotinib for 2 hours. The X-intercept of the linear fit is 1.07 nM. **B,** EGFR insASV is titrated by incubating with different concentrations of poziotinib for 2 hours. The X-intercept of the linear fit is 1.20 nM. **C,** EGFR insSVD is titrated by incubating with different concentrations of poziotinib for 2 hours. The X-intercept of the linear fit is 0.47 nM. **D,** EGFR insNPG is titrated by incubating with different concentrations of poziotinib for 2 hours. The X-intercept of the linear fit is 0.42 nM. **E,** EGFR L858R is titrated by incubating with different concentrations of poziotinib for 2 hours. The X-intercept of the linear fit is 0.99 nM. **F,** EGFR L858R/T790M is titrated by incubating with different concentrations of poziotinib for 2 hours. The X-intercept of the linear fit is 0.76 nM. **G,** SDS-PAGE showing same amount (by measured mass of 1 µg) of EGFR insASV, insSVD, L858R, L858R/T790M, and wild-type EGFR.

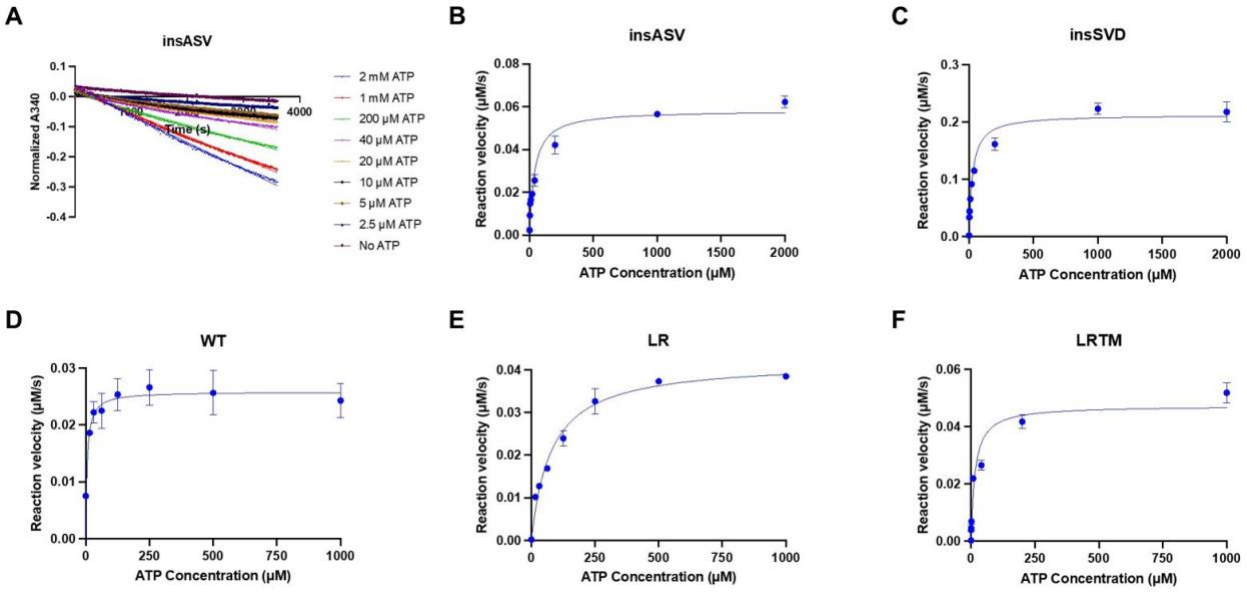


Figure 2.6. Kinetic characterization of EGFR insASV and insSVD mutant kinases. A, Representative reaction progress curves of coupled enzyme assay and purified EGFR insASV, in which consumption of ATP is coupled to oxidation of NADH. Absorbance at 340 nm was monitored at different ATP concentrations. **B-F,** Michaelis-Menten curves of EGFR wild-type, insASV, insSVD, L858R, or L858R/T790M. Reaction velocities were derived from the slopes of the initial linear regions of the reaction progress curves. Data are shown as mean ± SD (n=3 shown).

2.4: Discussion

After over one year of expression and purification trials, I found that the insASV or insSVD construct with residues ranging from 696-1022 produced best expression level and that the presence of 20% glycerol was critical to obtaining functional, monomeric EGFR insASV or insSVD kinases. The kinetic characterization showed that insASV and insSVD were much more active than WT EGFR, as expected for oncogenic mutants. The $K_{m,\ ATP}$ values of insASV and insSVD were lower than that of the drug-sensitive L858R mutant and comparable to the L858R/T790M mutant, which is resistant to 1st and 2nd generation EGFR TKIs[127]. The $K_{m,\ peptide}$ values of insASV and insSVD were much lower than WT or L858R, similar to L858R/T790M.

Why does insASV or insSVD need such a high concentration of glycerol during purification? One hypothesis is that EGFR insASV and insSVD may be prone to aggregation via formation of "chains" of asymmetric dimers or oligomerization, perhaps more so than other EGFR variants, because they are effectively "locked" in the active conformation. As I mentioned in Chapter 1, EGFR kinase molecules would form asymmetric dimers upon activation, where the receiver kinase needs to be in the active conformation to allow formation of the dimer. If EGFR kinase molecules heavily favor active conformation, they would be more likely to form chains of asymmetric dimers or oligomers as most of the molecules would be in the active conformation. It is possible that insASV and insSVD tend to oligomerize, which eventually would lead to aggregation especially if the protein concentration is relatively high.

In some of my trials with size exclusion chromatography in the presence of buffers with 5% glycerol, I could see that insASV or insSVD would appear as a very broad peak that overlaps with the void peak (Figure 2.7, A), and that insASV or insSVD could be seen almost anywhere

throughout the broad peak based on SDS-PAGE analysis (Figure 2.7, B). This supports my hypothesis that EGFR insASV and insSVD tend to oligomerize or form chains of asymmetric dimers, and they exist in different oligomerization states.

As most EGFR TKIs are ATP competitive, the inhibitor sensitivity of EGFR L858R and exon 19 deletions to EGFR TKIs results from their markedly decreased affinity for ATP compared to the wild-type EGFR. Our studies revealed that, unlike classical oncogenic EGFR mutants, the most prevalent EGFR exon 20 insertion mutants bind ATP with only a modestly lower affinity than wild-type EGFR, which suggests that developing ATP competitive inhibitors for these variants would be more challenging than for L858R variants. This could partially explain why exon 20 insertion variants are not as sensitive to traditional EGFR TKIs as L858R or exon 19 deletion variants, and emphasizes the need for inhibitors that differentiate wild type and mutant EGFR, rather than relying on different ATP affinities for mutant-selective inhibition.

My studies also showed that insASV and insSVD have stronger affinity for the peptide substrate poly(Glu_4,Tyr_1) than WT or L858R, similar to L858R/T790M. One association that crossed my mind is that the insertion mutants and L858R/T790M do not rely on the formation of extracellular dimer to be active, while WT and L858R are dimerization dependent[115]. Could differences in substrate affinity contribute to this? Is it possible that *cis*-autophosphorylation (e.g., phosphorylation of its own C-terminal tail) only occurs in those dimerization-independent variants? How different are affinities for endogenous peptide substrates among EGFR variants and how would kinetic parameters change with these endogenous peptides? To answer these questions, I would first purify the C-terminal tail of EGFR to substitute the use of poly(Glu_4,Tyr_1) in the coupled enzyme assay to compare the k_{cat} and K_m values to confirm the difference in substrate affinity is not an artifact.

One complication with my experimental setup is that in my kinetic characterization, I used 2 mM $MnCl_2$ instead of $MgCl_2$, where Mg^{2+} is the more physiologically relevant cation. The replacement of Mg^{2+} by Mn^{2+} was due to widely reported better facilitation of phosphoryl transfer reactions in tyrosine kinases[246]. However, Mn^{2+} significantly increases affinity for ATP and thus lowers $K_{m, ATP}$ values compared to Mg^{2+} ions, which would affect my measured kinetic parameters. Specifically, under physiological conditions, the $K_{m, ATP}$ values are expected to be higher due to the mostly Mg^{2+} environment, and the catalytic activities are also expected to change. However, this is not likely to change my conclusions when comparing across EGFR variants as the assay conditions were the same for different variants.

Another important question still remains to be answered. The gatekeeper mutation T790M is also known to restore ATP affinity to a level similar to that of wild-type EGFR. Then why are 3rd or 4th generation of EGFR TKIs not effective against these exon 20 insertion variants even though they are effective against variants with T790M? One possible reason is that T790M is located in the active site, which could affect the binding affinities of EGFR TKIs. However, exon 20 insertions are located in the loop following the C-helix, which is far from the ATP binding site and could not have as much impact on inhibitor binding affinities. The binding site is structurally the same as that of the wild-type EGFR, so it is very difficult to find an inhibitor that would have stronger binding affinities to exon 20 insertion variants. Therefore, it is significant to use my purified insASV and insSVD proteins to carefully study their inhibitor sensitivities to identify promising directions for further drug development and screen for more mutant-selective inhibitors.

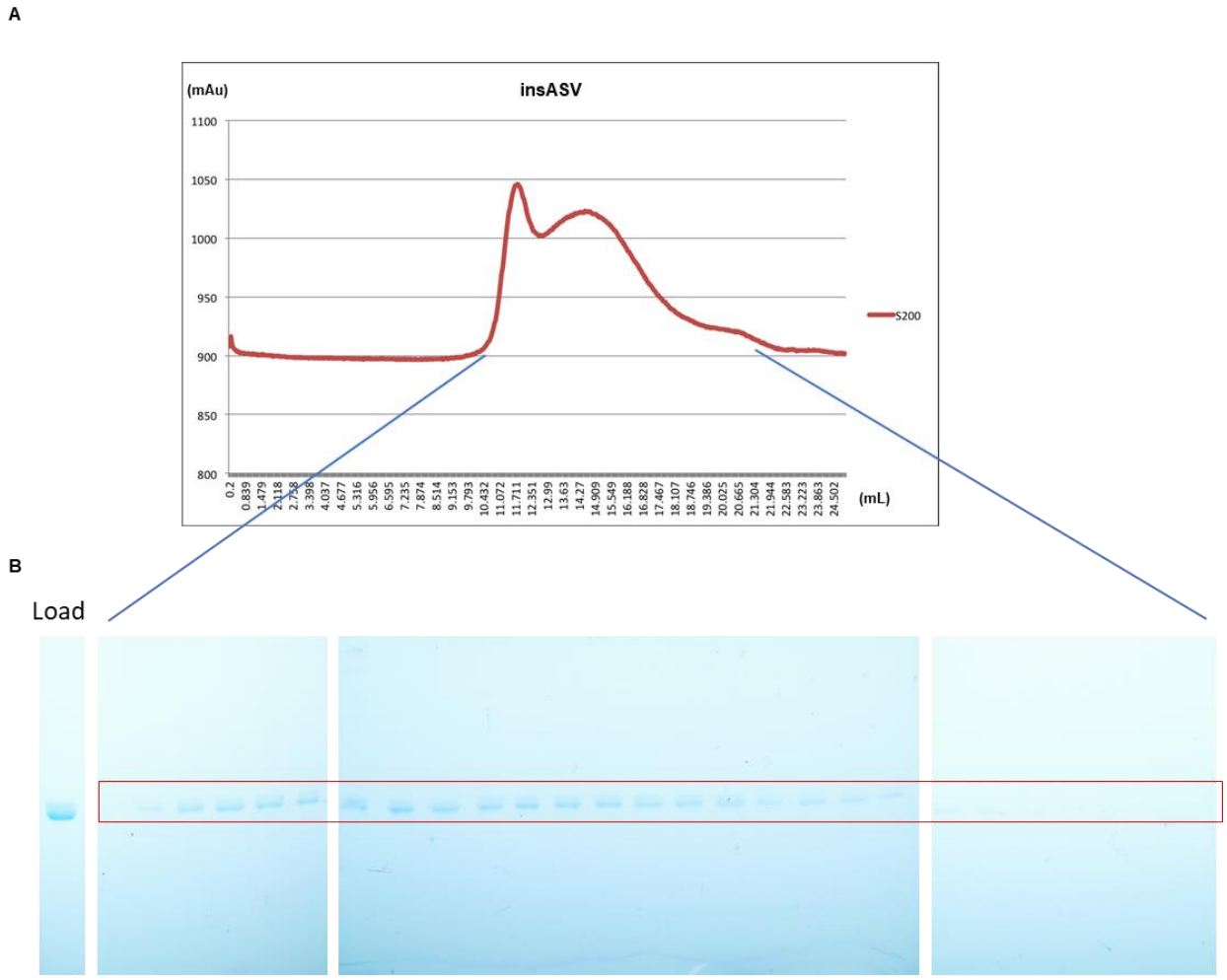


Figure 2.7. Representative unsuccessful size exclusion chromatography of EGFR insASV. Purified EGFR insASV loaded onto the size exclusion column. **A,** Size exclusion chromatogram of pAC8NKRed N-His10 C-TwinStrep EGFR insASV 696-1022. Note that the Superdex 200 column instead of the Superdex 75 column was used here. **B,** SDS-PAGE of fractions coming out of size exclusion column. These fractions correspond to the size exclusion profile. The sample before loading onto size exclusion column was also included for reference.

Chapter 3

Biochemical and cellular potencies of EGFR inhibitors

3.1: Introduction

Patients with the L858R mutation or exon 19 deletions show clinical response to treatment by the 1st generation TKIs[15,19,25,247–249]. The increased $K_{m,\ ATP}$ values of the mutants make ATP-competitive inhibitors more effective when compared to wild-type EGFR, since these reversible TKIs need to compete with ATP for binding[16]. However, the gatekeeper resistance mutation T790M often arises in those patients after TKI treatment, restoring ATP affinity to a level similar to wild-type EGFR[127]. Therefore, 2nd generation EGFR TKIs, which are irreversible inhibitors, were examined for their activities against the resistance mutation. Although they showed potent inhibition against T790M in preclinical studies, they were not effective in the clinical setting due to a narrow therapeutic window[152]. As a result, 3rd generation inhibitors, such as osimertinib, were developed to selectively target these variants[152,156,159]. Recently, allosteric inhibitors[174,176,177] and 4th generation TKIs[178,179] emerged as strategies to overcome the additional C797S resistance mutation following extensive 3rd generation TKI treatment[164,166]. However, tumors harboring insertions in exon 20 show limited clinical response to these TKIs[30,133,250], resulting in low PFS and ORR[197].

Although therapeutic antibodies that prevent dimerization and activation of EGFR, such as cetuximab[251], have proven successful for some variants[112], the activation of EGFR exon 20 insertion mutants is independent of dimerization[115]. Therefore, small molecule kinase inhibitors targeting exon 20 insertion variants represent an unmet clinical need[194,200,209,218,232]. Recently, several inhibitors targeting exon 20 insertions have entered pre-clinical or clinical studies, including BDTX-189[209], poziotinib[194], TAS6417 (CLN-081)[200], BAY 2927088[217], and TAK-788 (mobocertinib)[232], which recently got accelerated approval by FDA for exon 20 insertions

treatment[231]. Even though potent inhibitors including poziotinib demonstrated clinical activity, a limited therapeutic window due to potent inhibition of wild-type EGFR caused severe side effects to patients, such as skin rash, diarrhea, and pneumonitis, leading to frequent dose interruption, reduction, and discontinuation[203,252].

In 2021, the Heymach group divided EGFR mutations, including exon 20 insertions, into four distinct subgroups based on their structural impact on ATP and drug binding, hoping to help patients and physicians choose the right treatment and clinical trials. They showed that almost none of the inhibitors in their portfolio were selective against exon 20 insertions in Ba/F3 cells, which raised the question of whether a mutant-selective inhibitor against these insertions is possible[223]. Therefore, further research is needed to develop exon 20 insertion-selective inhibitors, widen the therapeutic window, and improve clinical outcomes[29]. To facilitate development of such mutant-selective inhibitors, it is important to improve our understanding of the molecular mechanisms underlying oncogenic activation by EGFR exon 20 insertions and why they are insensitive to current TKIs.

In this chapter, I will cover three topics. First, I will present biochemical potencies of a structurally diverse panel of EGFR inhibitors against different EGFR variants to compare their inhibitor sensitivities. I will show that only BAY-33, BAY-568, TAS6417, and TAK-788 are selective against EGFR insASV and insSVD. Then I will describe my findings from the selectivity plots of inhibitor biochemical potencies to show that exon 20 insertions are not inherently resistant to inhibitors but require different scaffolds. Finally, I will present cellular studies of EGFR exon 20 insertion inhibitors against insASV, insSVD or wild-type, and illustrate that BAY-568, TAS6417 and TAK-788 effectively blocked downstream signaling.

3.2: Materials and Methods

Inhibitors

Unless otherwise noted, inhibitors used in this study were purchased from commercial vendors and are >95% pure per the vendors' analyses. EAI002, EAI045, and JBJ-09-063 were synthesized in-house for prior studies[174,175,177]. BAY-33 corresponds to example #33 in patent WO2019081486[219] and was synthesized by PepTech Corp. (>95% pure). Compound stocks of 10 mM were prepared in DMSO and stored at -20 ºC.

EGFR inhibition assay

Inhibition assays were performed using the homogenous time-resolved fluorescence (HTRF) KinEASE tyrosine kinase assay kit (Cisbio) according to the manufacturer's protocol. Inhibitors (10 mM DMSO stocks) were dispensed into black 384-well plates (Corning) using an HP D300e dispenser (Hewlett-Packard) and were normalized to a 1% final DMSO concentration. For inhibitor experiments, the reaction mixture containing purified EGFR at a final concentration of 5 nM (wild-type EGFR) or 0.5 nM (L858R, L858R/T790M, insNPG, insSVD and insASV) was dispensed using a Multidrop Combi dispenser (ThermoFisher) and incubated with compounds at room temperature for 30 min. Reactions were initiated with 1 mM ATP and allowed to proceed for 30 min (Wild-type) or 10 min (mutants) at room temperature before being quenched using the detection reagent with EDTA from the HTRF KinEASE assay kit.

In a typical experiment, the final reaction mixture containing ATP had a volume of 10 µL in each well, and another 10 µL detection reagent was later added to quench the reaction. The

FRET signal ratio was measured at 665 and 620 nm using a PHERAstar microplate reader (BMG LABTECH). Data were processed using GraphPad Prism and fit to a three-parameter dose-response model with a Hill slope constrained to -1.

Ba/F3 cell culture and anti-proliferation assay

EGFR exon 20 insASV Ba/F3 and EGFR exon 20 insSVD Ba/F3 cells were cultured in RPMI media (Life technologies, Cat# 11875119) containing 10% fetal bovine serum (Life technologies, Cat# 10437028) and 1% Penicillin/Streptomycin (Life technologies, Cat# 10378016). WT EGFR Ba/F3 cells were cultured in RPMI media containing 10 ng/mL EGF (Life Technologies, Cat# PHG0311L),10% fetal bovine serum and 1% Penicillin/Streptomycin. All the cell lines were cultured at 37°C in 5% CO2 humidified air and tested for mycoplasma negative.

In the anti-proliferation assay, EGFR exon 20 insASV Ba/F3, EGFR exon 20 insSVD Ba/F3 or WT EGFR Ba/F3 (+50 ng/mL EGF) cells were seeded at the density of 1000 cells/well in 384-well plates. The compounds were added to the cells and incubated for 72 hr. Cell viability was determined by using CellTiter-Glo (Promega #G7571) according to the manufacturer's instructions, measuring luminescence using an Envision plate-reader (PerkinElmer Inc.). Dose-response curves were generated using non-linear regression curve fit in GraphPad Prism 9 (GraphPad Software).

Immunoblotting and Antibodies

Cells were lysed in RIPA buffer (150 mM NaCl, 1.0% IGEPAL® CA-630, 0.5% sodium deoxycholate, 0.1% SDS, 50 mM Tris, pH 8.0) (Sigma, Cat# R0278) with protease inhibitor and

phosphatase inhibitor (Roche). The protein concentrations were measured by BCA analysis (Thermo Fisher Scientific, Cat# PI23225). Equal amounts of protein were resolved by 4-12% Tris-Base gels (Life Technologies), and then transferred to the Immuno-Blot PVDF membrane (BioRad, Cat# 1620177). Proteins were probed with appropriate primary antibodies at 4 ℃ overnight and then with IRDye®800-labeled goat anti-rabbit IgG (LICOR Biosciences, Cat# 926-32211), IRDye®800-labeled goat anti-mouse IgG (LICOR Biosciences, Cat# 926-32210) or IRDye 680RD goat anti-Mouse IgG (LICOR Biosciences, Cat# 926-68070) secondary antibodies at room temperature for 1 hour. The membranes were detected on Odyssey CLx system.

Antibodies used in this study include anti-following proteins: EGFR (Cell signaling Technology, #4267S, 1:1000), p-EGFR (Tyr1068) (Cell signaling Technology, #3777S, 1:1000), ERK1/2 (Cell signaling Technology, #4696S, 1:0000), p-ERK1/2 (Cell Signaling Technology, #4370S, 1:1000), Akt (Cell signaling Technology, #9272L, 1:1000), p-Akt (Cell Signaling Technology, #4060S, 1:1000) and β-Actin (Cell Signaling Technology, #3700, 1:1000).

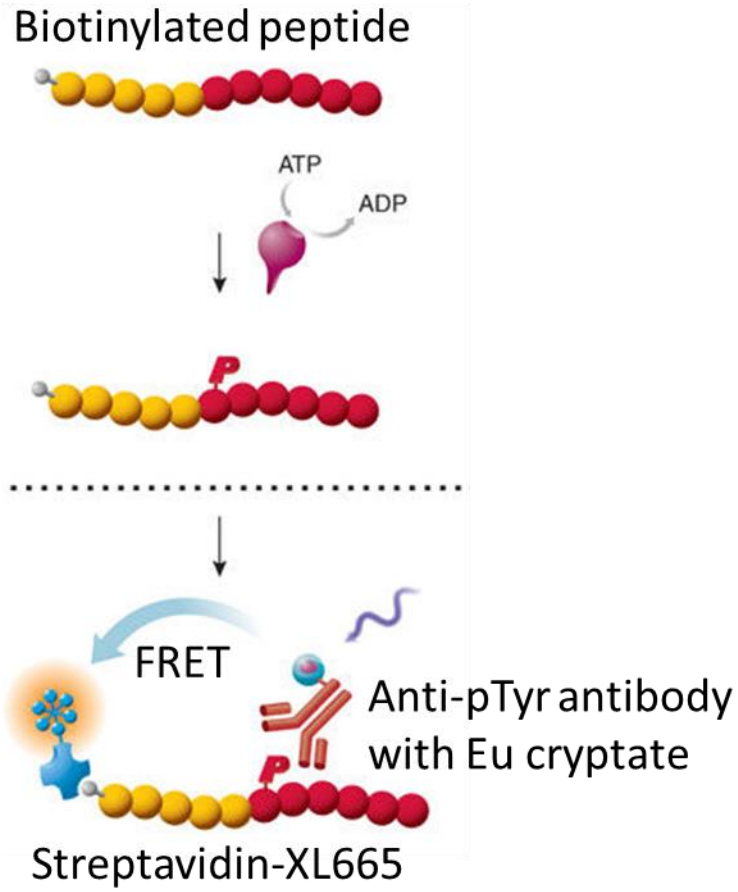


Figure 3.1. Illustrative diagram of homogenous time-resolved fluorescence (HTRF) assay. When the biotinylated peptide is phosphorylated by EGFR kinase, anti-pTyr antibody with Eu cryptate is recruited to the phosphorylated Tyr. Streptavidin-XL665, recruited by the biotin on the peptide substrate, generates fluorescence signals at 665 nm when the bound Eu crypate is excited with fluorescence emitted at 620 nm (Degorce et al. 2009[253]).

3.3: Results

Biochemical sensitivity of EGFR exon 20 insertion mutations to EGFR inhibitors

To better understand the inhibitor sensitivity of EGFR exon 20 insertion variants, we measured biochemical potencies of a diverse panel of EGFR inhibitors (chemical structures shown in Figure 1.6) against three representative inhibitor-resistant EGFR exon 20 variants with insertions at the C-terminus of the αC-helix, insASV, insSVD, and insNPG. We compared these to the EGFR L858R and L858R/T790M variants, which are more effectively targeted by TKIs, and wild-type EGFR. We used recombinant EGFR kinase domain (prepared as described in Chapter 2) together with a physiological 1 mM ATP concentration and a homogenous time-resolved fluorescence (HTRF) assay to obtain dose-response curves for this diverse panel of EGFR inhibitors including compounds from different generations of TKIs in addition to those inhibitors specifically developed to target EGFR exon 20 insertions.

In the HTRF assay, when the biotinylated peptide is phosphorylated by EGFR kinase, the anti-phosphotyrosine antibody with Eu cryptate is recruited to the phosphorylated tyrosine. Streptavidin-XL665, recruited by the biotin on the peptide substrate, generates fluorescence signals at 665 nm when the bound Eu crypate is excited with fluorescence emitted at 620 nm[253] (Figure 3.1). The fluorescence resonance energy transfer (FRET) efficiency is calculated using the fluorescence signal at 665 nm divided by the background fluorescence signal at 620 nm, multiplied by 10,000. The HTRF assay was used for my inhibitor studies because it is highly sensitive, which allows me to use small quantities of EGFR proteins for accurate measurement of inhibitor sensitivity.

Dose-response curves for these 20 inhibitors profiled against the six EGFR variants are presented in Figure 3.2, and the corresponding IC_{50} values are collected in Table 3.1. BAY-33 was included in the panel because it was one of the most mutant-selective inhibitors against EGFR exon 20 insertions from Bayer's patent, while the preclinical activity of BAY-568 was reported by Bayer in 2021[218,219]. My inhibitor studies showed that gefitinib, a 1st generation EGFR TKI, was selective against L858R, but was not potent against L858R/T790M or exon 20 insertions, which is consistent with prior studies[128,152]. Lapatinib, a reversible dual inhibitor of EGFR and HER2, was more potent against WT EGFR than the mutants. As expected, afatinib, a 2nd generation EGFR TKI, was potent but not selective against L858R or L858R/T790M. Neratinib, an irreversible analog of lapatinib, was potent but not selective against L858R. Similar to lapatinib, it was much more potent against WT EGFR than L858R/T790M or exon 20 insertions, although some potency was restored against L858R/T790M due to the covalent modification. Most of the 3rd generation EGFR TKIs displayed mutant selectivity against L858R or L858R/T790M, with strong potency against the double mutant, except that mavelertinib was not selective against any mutant.

BI-4020, a 4th generation EGFR TKI designed to overcome C797S, was selective against L858R or L858R/T790M, despite a compromised potency against the classical mutant without T790M. Allosteric inhibitors were selective against L858R or L858R/T790M, although JBJ-09-063 only showed ~2-fold selectivity due to its strong potency against WT EGFR. Taken together, the first three generations of EGFR TKIs were not selective for exon 20 insertion variants, while 4th generation EGFR TKIs and allosteric inhibitors displayed limited potency.

Table 3.1

Inhibitor profiles across EGFR variants

		IC_{50} (nM)					
		WT	**L858R**	**L858R/ T790M**	**insASV**	**insSVD**	**insNPG**
1st Gen	**Gefitinib**	>1000	40	>1000	>1000	>1000	>1000
	Lapatinib	0.5	3.5	>1000	136	212	9.9
2nd Gen	**Afatinib**	2.6	0.7	3.0	>1000	>1000	267
	Neratinib	0.5	0.3	14	33	348	11
3rd Gen	**Osimertinib**	328	37	22	362	409	297
	Olmutinib	73	25	1.2	69	122	44
	WZ4002	63	16	2.6	105	126	68
	Naquotinib	79	13	1.2	133	409	74
	Mavelertinib	4.3	3.8	3.2	14	24	11
	Nazartinib	56	19	6.2	150	304	78
4th	**BI-4020**	>1000	69	0.4	>1000	>1000	427
Allosteric	**EAI045**	>1000	9.3	10	>1000	>1000	>1000
	EAI002	>1000	861	167	>1000	>1000	>1000
	JBJ-09-063	1.5	0.8	0.9	28	192	20
Exon 20	**BAY-33**	>1000	4.5	227	2.4	3.8	2.1
	BAY-568	>1000	1.4	72	0.6	0.3	0.3
	TAS6417	34	3.2	1.1	16	6.0	4.1
	Poziotinib	0.3	0.7	0.9	0.6	0.9	0.3
	TAK-788	4.4	0.6	0.6	1.4	1.7	1.3
	BDTX-189	0.3	0.3	1.5	2.1	55	0.8

Figure 3.2

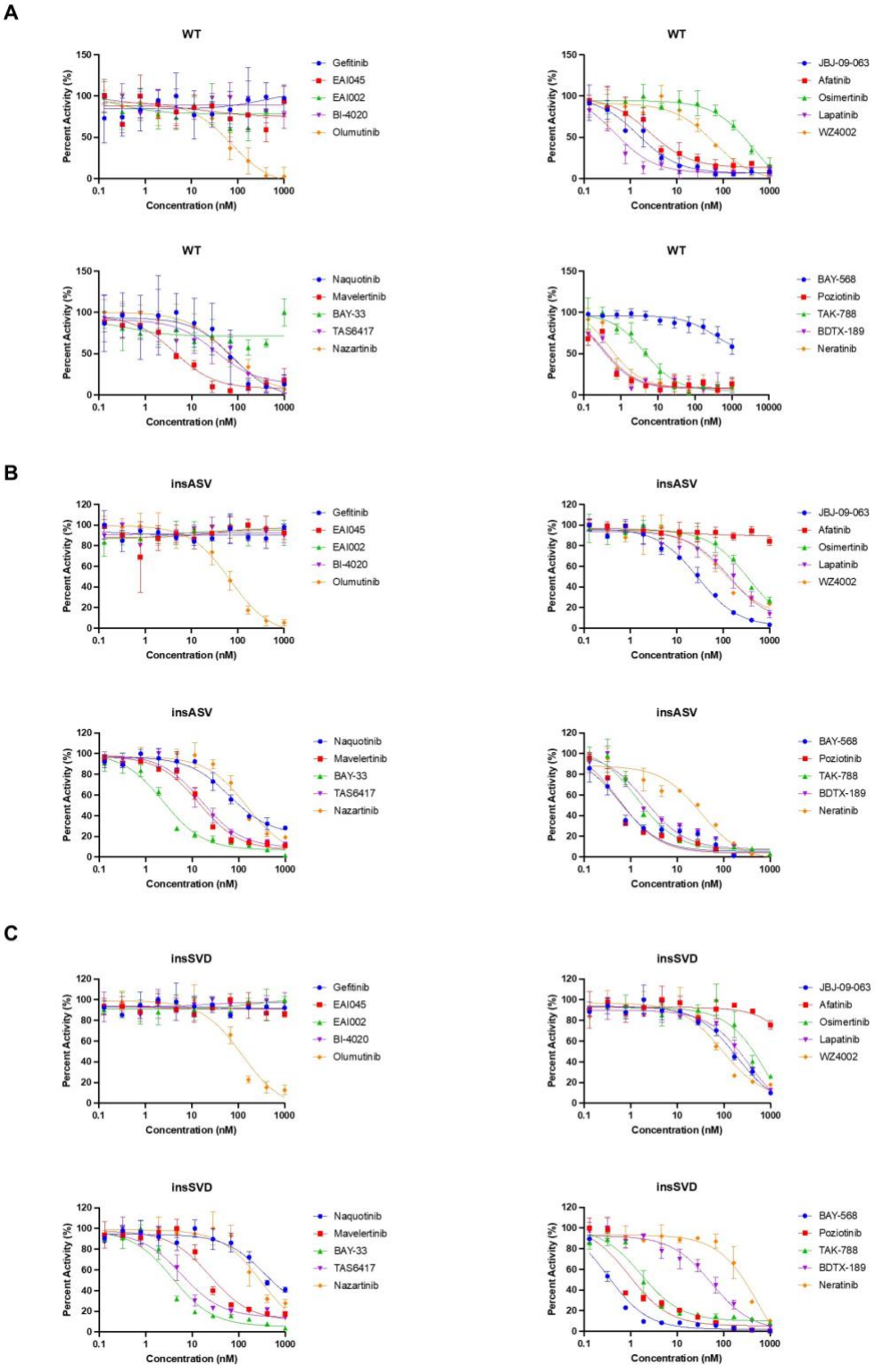

A
WT
Percent Activity (%)
Concentration (nM)
Gefitinib
EAI045
EAI002
BI-4020
Olumutinib
WT
JBJ-09-063
Afatinib
Osimertinib
Lapatinib
WZ4002
WT
Naquotinib
Mavelertinib
BAY-33
TAS6417
Nazartinib
WT
BAY-568
Poziotinib
TAK-788
BDTX-189
Neratinib
B
insASV
Percent Activity (%)
Concentration (nM)
Gefitinib
EAI045
EAI002
BI-4020
Olumutinib
insASV
JBJ-09-063
Afatinib
Osimertinib
Lapatinib
WZ4002
insASV
Naquotinib
Mavelertinib
BAY-33
TAS6417
Nazartinib
insASV
BAY-568
Poziotinib
TAK-788
BDTX-189
Neratinib
C
insSVD
Percent Activity (%)
Concentration (nM)
Gefitinib
EAI045
EAI002
BI-4020
Olumutinib
insSVD
JBJ-09-063
Afatinib
Osimertinib
Lapatinib
WZ4002
insSVD
Naquotinib
Mavelertinib
BAY-33
TAS6417
Nazartinib
insSVD
BAY-568
Poziotinib
TAK-788
BDTX-189
Neratinib

Figure 3.2 (continued)

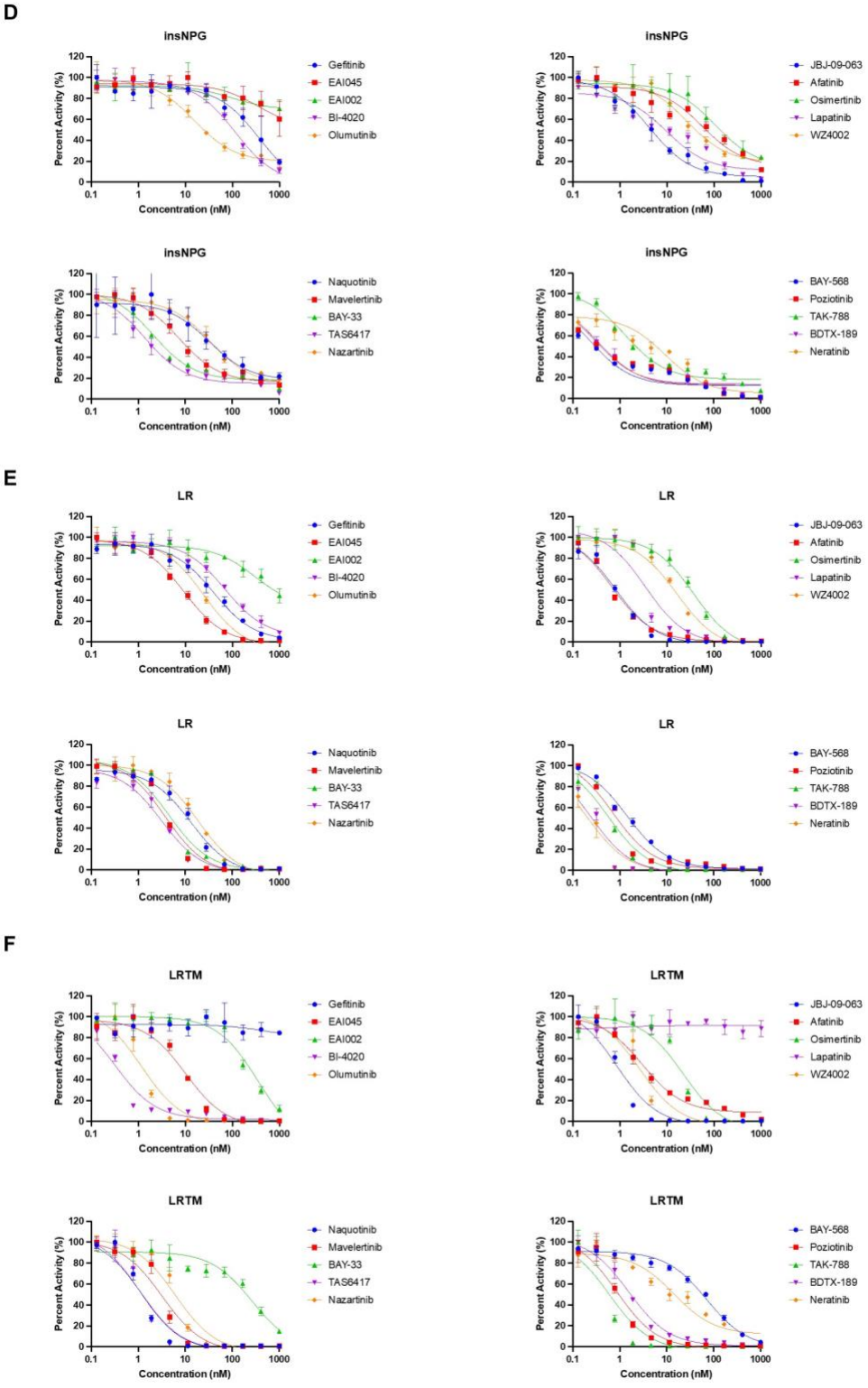


Figure 3.2. Concentration-response curves of EGFR inhibitors studied here against WT EGFR, insASV, insSVD, insNPG, L858R, or L858R/T790M. A-F, Inhibition curves with Y-axis showing percent activity (%), where 100% activity corresponds to no inhibitor present while 0% corresponds to only background signal, and X-axis showing concentrations of inhibitors used. Error bars indicate SD (N=3).

Mutant-selectivity of EGFR exon 20 insertion-targeted TKIs

To facilitate direct comparison of the potency of EGFR exon 20 insertion-targeted TKIs on wild-type versus mutant EGFR, selected dose-response data from Figure 3.2 are replotted with inhibition curves clustered by compounds, rather than by EGFR variants (Figure 3.3). Several observations are immediately apparent from inspection of these plots. First, poziotinib is potent but entirely non-selective (Figure 3.3, F). It inhibits WT EGFR and all five variants tested with essentially equivalent potency (0.3-0.9 nM, see also Table 3.1). Second, the Bayer compounds exhibit clear sparing of WT EGFR, while potently inhibiting all three exon 20 insertion variants as well as L858R (Figure 3.3, A-B). However, the potency of BAY-33 or BAY-568 (a close analog of BAY-33) is markedly reduced by the T790M mutation (Figure 3.3, A-B, see L858R/T790M curves). Finally, TAS6417 and TAK-788 exhibit intermediate degree of mutant selectivity, with ~5-fold selectivity for TAS6417 and ~3-fold selectivity for TAK-788 (Figure 3.3, C-D; Table 3.1). Generally, compounds that inhibited exon 20 insertion variants did so with comparable potency against all three exon 20 insertion variants we tested here.

Taken together, among EGFR exon 20 insertion-targeted TKIs, poziotinib and BDTX-189 are potent but not selective against any variant, while the other inhibitors are potent and selective against all EGFR mutants, except that the Bayer compounds show compromised activity against T790M. The only insertion-selective inhibitors are BAY-568, BAY-33, TAK-788 (mobocertinib), and TAS6417 (CLN-081), with BAY-568 being the most selective.

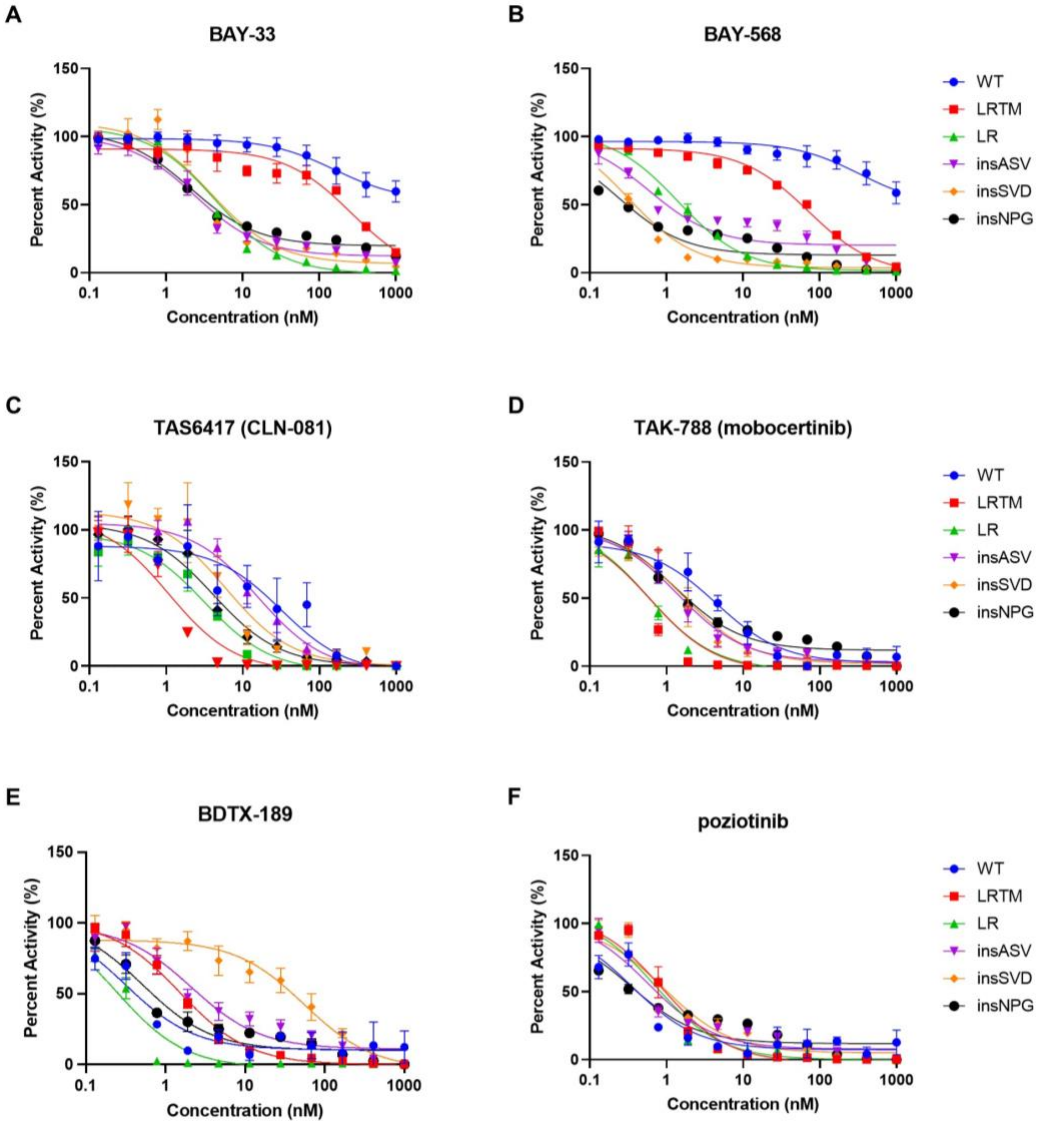


Figure 3.3. Biochemical inhibition curves of EGFR exon 20 insertion-targeted TKIs. A-F, Inhibition curves of exon 20 insertion-targeted inhibitors against wild-type EGFR, L858R, L858R/T790M, insASV, insSVD or insNPG. Data were normalized to enzyme activity when no inhibitor is present. Note that this is a re-plot of data in Figure 3.2. (N=3; points are shown as mean ± SD)

Exon 20 insertions differ in drug sensitivity but not inherently resistant

As the inhibition data above reveal, a subset of inhibitors are quite potent against insASV, insSVD and insNPG, and a general trend of similar potencies against these mutants is apparent in Table 3.1. To better understand the correlation of drug sensitivity across EGFR variants, we plotted comparisons of inhibitor IC_{50}s for selected pairs of EGFR variants (Figure 3.4). For exon 20 insertions compared with WT EGFR, these plots reveal no obvious correlation (Figure 3.4, A-B). By contrast, there is clear correlation in sensitivities across the three exon 20 insertion variants (Figure 3.4, D-F). In particular, inhibitors are nearly equipotent against insASV and insNPG (Figure 3.4, E).

Interestingly, there is also clear correlation in sensitivity between L858R and WT EGFR, but most compounds are more potent against L858R than WT (Figure 3.4, C). Almost all of the compounds we tested are potent against the traditionally drug-sensitive L858R variant. We expect that this is a manifestation of the drug-sensitizing effect of the markedly increased $K_{m,\ ATP}$ of this mutant. By contrast, insertion mutants effectively act more like different kinases as compared to WT EGFR (Figure 3.4, A-B). Based on the selectivity plots, we can conclude that EGFR exon 20 insertions are not inherently drug-resistant, but rather differ in their inhibitor sensitivity from WT EGFR.

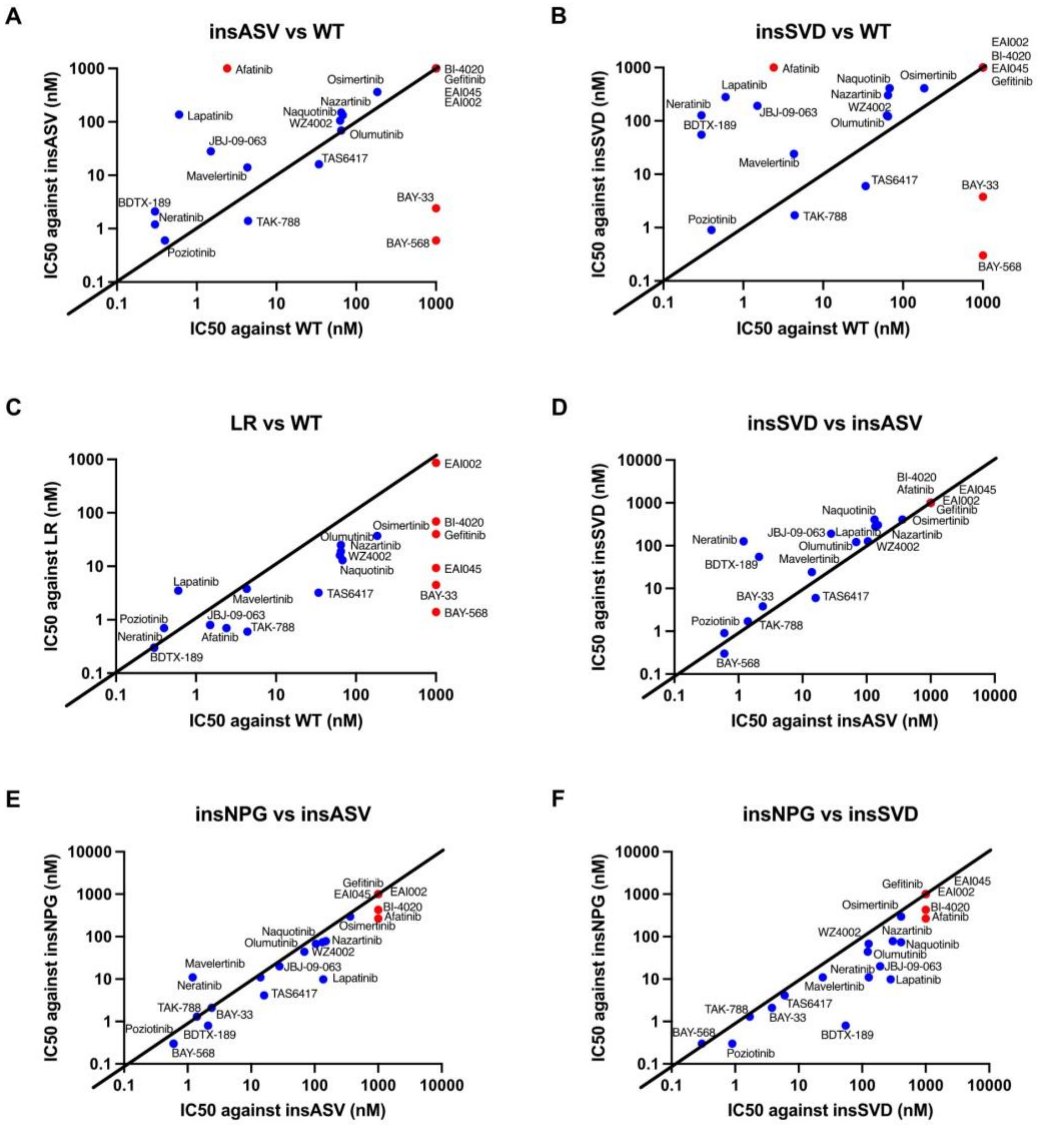


Figure 3.4. Comparisons of inhibitor sensitivity between selected EGFR variants. A-F, Selectivity plots showing insASV, insSVD, or LR vs WT, insSVD vs insASV, and insNPG vs insASV or insSVD, with X- or Y-axis being biochemical IC_{50} values of EGFR inhibitors. The line y=x is shown for easier identification of mutant-selective compounds. The red dots indicate that the compounds tested have IC_{50} values beyond the maximum concentration used in the study (each point corresponds to the average of three data points).

Cellular potencies of EGFR exon 20 inhibitors against insertion variants

We sought to validate our biochemical results by treating Ba/F3 cells expressing wild-type EGFR, EGFR insASV, and insSVD with EGFR exon 20 insertion-targeted inhibitors to determine their cellular potency and mutant selectivity (Figure 3.5, A-F). Ba/F3 cells have been widely used in research on EGFR since they are easy to culture and highly sensitive to EGFR signaling. Normal Ba/F3 cells do not express EGFR and need interleukin-3 (IL-3) to survive and proliferate, but wild-type EGFR-transfected Ba/F3 cells can use EGF to replace IL-3 for short-term proliferation[254]. The cancer-derived mutant EGFR-transformed Ba/F3 cells are independent of IL-3 to grow, which makes those cells entirely dependent on EGFR activity for survival in the absence of IL-3[255]. As a result, anti-proliferation assays are commonly used to examine EGFR inhibition by inhibitors.

We assessed cell viability in the presence of exon 20 inhibitors and observed that BAY-568, TAK-788, and TAS6417 were selective for insASV and insSVD compared to wild-type EGFR, which displayed similar selectivity trends as in the biochemical studies (Figure 3.5, A-C). Poziotinib showed stronger potency against wild-type EGFR than against insASV or insSVD, while BDTX-189 displayed similar potency against wild-type and against the insertion variants. Notably, potency of BAY-33 in cell-based assays (~2 μM against exon 20 insertions) did not match biochemical results (~2 nM IC_{50}), possibly indicating issues with cell permeability and/or stability of the compound. BAY-568, the fluorine-substituted analog of BAY-33, is both potent and selective for insASV or insSVD (>6-fold) (Figure 3.5, A-C; Table 3.2). Western blots of the EGFR signaling cascade revealed phosphorylation of EGFR, ERK, and Akt was not fully inhibited by 100 nM BAY-568 in cells expressing wild-type EGFR but were for insASV and insSVD (Figure 3.5, D-F). The inhibition of downstream signaling confirmed that the cell-killing was on-target, which was

a result of EGFR inhibition. Taken together, cell-based assays agree with our biochemical results in terms of potency and mutant selectivity, except BAY-33.

Table 3.2

Cellular potencies of EGFR exon 20 insertion-targeted TKIs

	IC_{50} (nM)		
	WT	**insASV**	**insSVD**
poziotinib	0.8	2.7	2.4
TAK-788	53	11	12
BDTX-189	21	27	25
TAS6417	101	42	41
BAY-33	7860	2360	1598
BAY-568	128	24	20

Figure 3.5

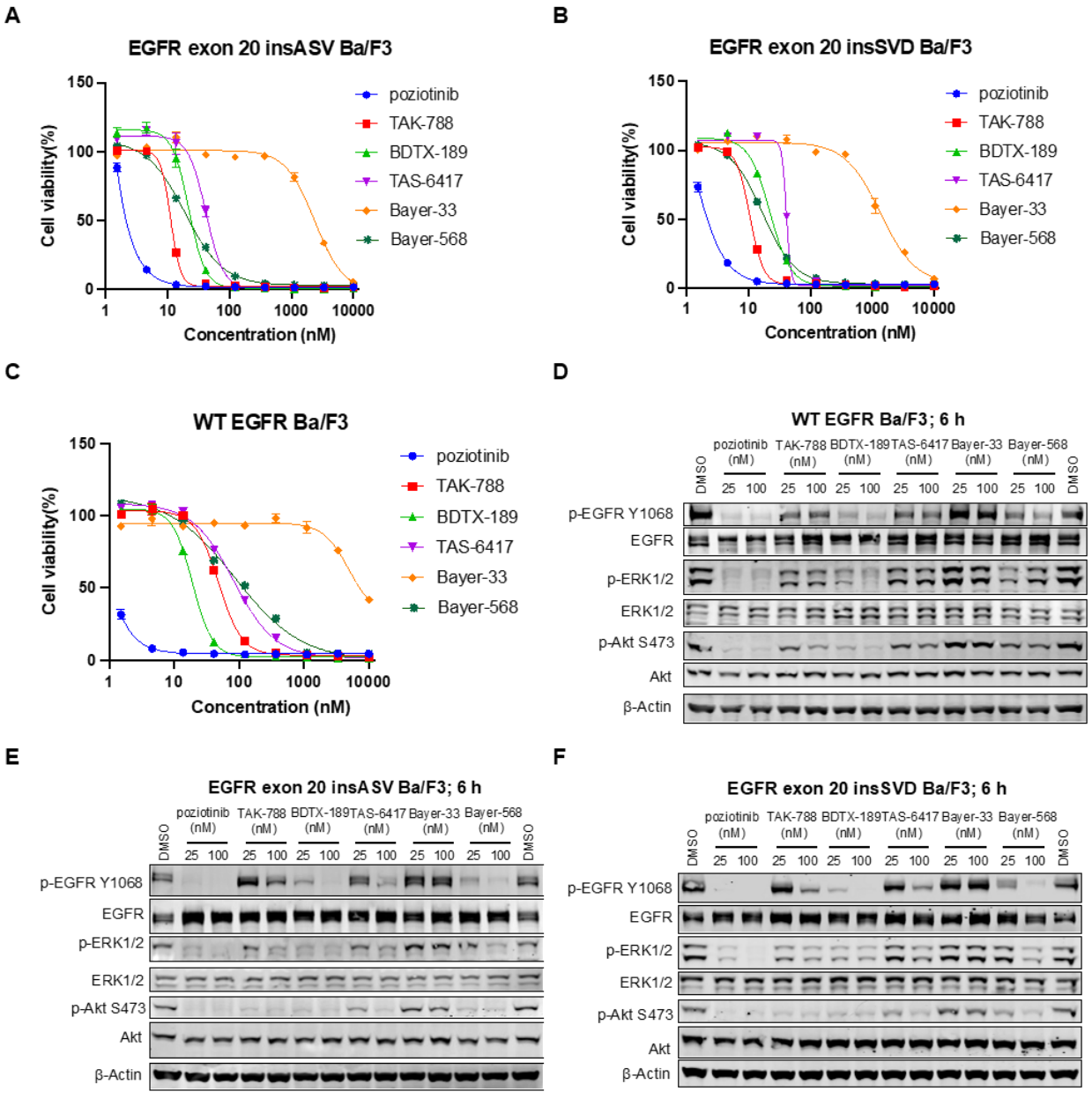

A
EGFR exon 20 insASV Ba/F3
Cell viability(%)
Concentration (nM)
poziotinib
TAK-788
BDTX-189
TAS-6417
Bayer-33
Bayer-568
B
EGFR exon 20 insSVD Ba/F3
C
WT EGFR Ba/F3
D
WT EGFR Ba/F3; 6 h
E
EGFR exon 20 insASV Ba/F3; 6 h
F
EGFR exon 20 insSVD Ba/F3; 6 h
p-EGFR Y1068
EGFR
p-ERK1/2
ERK1/2
p-Akt S473
Akt
β-Actin
DMSO

Figure 3.5. Cellular activity of exon 20 inhibitors in Ba/F3 model. A-C, Dose-dependent cell growth inhibition of Ba/F3 cells expressing EGFR D770_N771insNPG, D770_N771insSVD or wild-type EGFR in the presence of 50 ng/mL EGF. Ba/F3 cells expressing aforementioned EGFR mutations were treated with the indicated doses of exon 20 inhibitors for 72 hours. Cell survival was measured using a CellTiter Aqueous One Solution Cell Proliferation Assay. Error bars indicate SD (N = 3). **D-F,** Dose response of exon 20 inhibitors in Ba/F3 cells expressing EGFR D770_N771insNPG, D770_N771insSVD or wild-type EGFR in the presence of 50 ng/mL EGF. The cells were treated with indicated doses of inhibitors for 6 hours. Phosphorylation of EGFR, AKT, and ERK proteins was detected by immunoblotting.

3.4: Discussion

EGFR exon 20 insertion mutations represent a unique set of EGFR-activating mutations, which are generally described as insensitive to clinically available EGFR TKIs[29,195,256]. Recently, two therapeutic agents received accelerated approval for treatment of advanced NSCLC patients with EGFR exon 20 insertion mutations, a bispecific antibody targeting EGFR-MET, amivantamab[226], and a small molecule TKI, TAK-788 (mobocertinib)[231]. However, both of these agents have considerable liabilities. Amivantamab eliminates cancer cells by inducing antibody-dependent cell-mediated cytotoxicity (ADCC)[227], which does not differentiate cancer cells from normal cells. As a result, amivantamab has a narrow therapeutic window and limited efficacy[226,230]. Furthermore, most patients in the TAK-788 clinical trials experienced severe side effects, which could be due to its strong potency against wild-type EGFR and low (3-fold) mutant selectivity I report here. Therefore, the discovery and development of potent, exon 20 insertion-selective EGFR inhibitors remains a high priority.

In the present report, I showed that BAY-568, BAY-33, TAK-788 and TAS6417, are mutant selective in biochemical and cellular assays, with >1000-fold selectivity for BAY-568 and >300-fold selectivity for BAY-33. Although BAY-33 is not as potent or mutant selective in cellular assays as in biochemical assays possibly due to issues with cell permeability or metabolic stability[257], the fluorine-substituted analog, BAY-568, retains favorable profile in cellular assays. However, the Bayer compounds show compromised activity against T790M, which raises the possibility that T790M could confer resistance in the presence of exon 20 insertions. Therefore, it might be beneficial in the long run to further improve the Bayer compounds or screen for inhibitors against exon 20 insertions regardless of T790M. Observation of unique inhibitor sensitivity of exon 20

insertion variants in Figure 3.4 suggests that there may be additional scaffolds that are selective for exon 20 insertions. Thus, screening a larger and more diverse collection of kinase inhibitor scaffolds may provide new approaches to selective inhibition. This is the topic of Chapter 5 of my thesis.

Additionally, I observed some mutant-selectivity for TAK-788 and TAS6417, whereas two other clinical candidates developed for exon 20 insertion mutations, poziotinib and BDTX-189, were not mutant-selective. The selectivity plots show that exon 20 insertion variants and WT EGFR have very different inhibitor sensitivity, which suggests a need for different scaffolds against exon 20 insertions from previous generations of EGFR TKIs. Therefore, it might not be surprising to see "repurposed" 2nd generation EGFR TKIs, including poziotinib and BDTX-189, fail clinical trials to treat exon 20 insertions[206,208]. On the other hand, Bayer compounds were deliberately developed from an initial screen against exon 20 insertions (counter-screened against WT EGFR for mutant selectivity), so it makes sense to see that they are selective against exon 20 insertions[218].

Although the most recent allosteric inhibitors, such as JBJ-09-063[177], have been proved successful against many EGFR variants[174,176,177], they were found not selective for insertion variants in my studies. One hypothesis is that exon 20 insertion mutants have heavily shifted equilibrium towards the active conformation, which is not easily accessible for these allosteric inhibitors. Similarly, type-1.5 inhibitors that bind the inactive conformation of EGFR kinase, including lapatinib and neratinib, were not selective for exon 20 insertions, possibly due to the same reason. BI-4020, a recently developed 4th generation EGFR TKI that was designed to overcome the T790M and C797S resistance mutations, was found not potent against EGFR insASV, insSVD or insNPG, most likely due to a two-fold effect: loss of binding affinity due to the

lack of M790, and maintenance of strong competition from ATP due to low $K_{m, ATP}$ values of these exon 20 insertion mutants.

3.5: Contributions

I expressed and purified all EGFR proteins and performed the biochemical assays. Dr. Jie Jiang from the Jänne lab performed the cellular experiments. Prof. Michael J. Eck and Prof. Pasi A. Jänne supervised the research.

Chapter 4

Structural and mechanistic investigations of mutant selectivity

4.1: Introduction

To further understand the molecular mechanisms of mutant selectivity of BAY-33, BAY-568, TAS6417, and TAK-788, I performed a series of biochemical experiments to examine the kinetics of inhibitor binding, which could be used to dissect the individual steps of inhibition[258]. For a typical covalent inhibitor, the inhibition can be divided into two steps: reversible binding, and then covalent bond formation[259] (Scheme 4.1). The reversible binding affinity (Ki) is defined by k_{off}/k_{on}, while the rate of covalent bond formation (k_{inact}) is a first-order rate constant that measures how fast an irreversible inhibitor can covalently bind to the enzyme. The overall inactivation efficiency (k_{inact}/K_I) accounts for both steps, where K_I is the inhibition constant defined by the inhibitor concentration giving half-maximal rate for covalent bond formation[258,260]. Although K_I is not equal to Ki ($K_I = [k_{off} + k_{inact}]/k_{on}$, ≈ Ki when $k_{inact} << k_{off}$), it is often used to describe the potency of the reversible binding step[260]. Importantly, unlike IC_{50} values of irreversible inhibitors, k_{inact}/K_I values are not time dependent, thus can be used to compare the overall potency of irreversible inhibitors more accurately than IC_{50} values[261]. In addition, by analyzing each step of covalent inhibition, I could compare them to wild-type EGFR and gain insights into the source of mutant selectivity, ultimately facilitating better decision making on inhibitor development. Specifically, for TAK-788 and TAS6417, I was interested to determine whether the reversible binding step or the covalent bond formation step differs more between wild-type and insertion mutants.

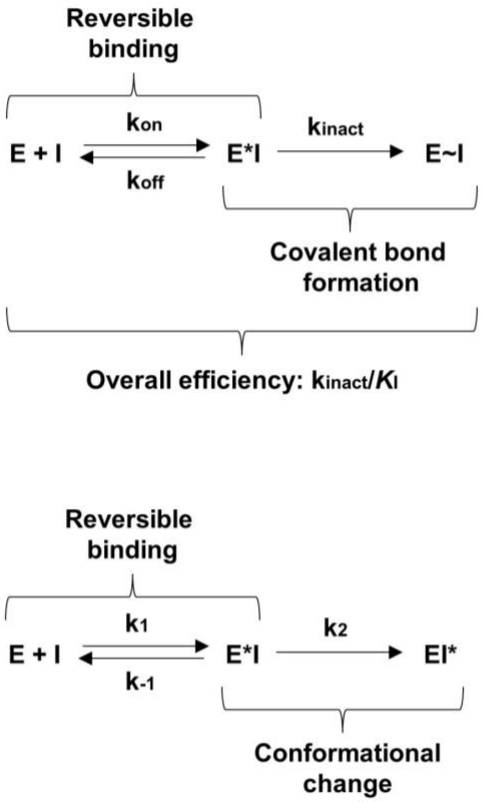
Similarly, even though reversible inhibitors cannot covalently bind to the kinase, in some cases, they could also have a two-step mechanism of inhibition, where the reversible inhibitor first binds to the target protein and then changes the protein conformation so that the inhibitor is "locked" in the protein-inhibitor complex or has a very slow off-rate[262–265] (Scheme 4.1). For

reversible inhibitors with a one-step mechanism of inhibition, the k_{obs} values derived from the reaction progress curves at different inhibitor concentrations are expected to be in a linear relationship with the inhibitor concentrations, while for inhibitors with a two-step mechanism, the k_{obs} values versus inhibitor concentrations are hyperbolic[262]. These potent reversible inhibitors with a two-step mechanism, including the widely used chemotherapy agent methotrexate, behave almost like irreversible inhibitors[264–266]. Therefore, I was also interested to see if BAY-33 and BAY-568, as potent reversible mutant-selective inhibitors, exhibit such a two-step mechanism of inhibition. If they indeed inhibit EGFR kinases via a two-step mechanism, I would like to examine which step contributes most to the potency differences between wild-type and insertion mutants.

In addition to the kinetics studies of those mutant-selective inhibitors, structural studies of those mutant-selective inhibitors could also provide insight into the underlying source of selectivity for the exon20 insertion variants. By analyzing inhibitor binding modes in the kinase-inhibitor complexes, I hoped to find any unique interactions that are not seen in most of EGFR TKI-kinase complex structures that might explain mutant-selective inhibition. In the long term, understanding the structural basis for selectivity could guide development of even more selective agents in the future. Unfortunately, only structures of WT EGFR in complex with TKIs were determined, which limited my analysis, but I was still able to find some interesting interactions with these TKI-kinase complex structures.

In this chapter, I present my results from kinetics studies of inhibition. Mutant-selective covalent inhibitors, such as TAK-788, are examined in the context of insASV, insSVD, and wild-type EGFR, and are compared with the inhibition properties of non-mutant-selective covalent inhibitors. BAY-33, a reversible mutant-selective inhibitor, is tested as well to see which mechanisms of inhibition it adopts, and to identify what contributes to its mutant selectivity. Second, I will discuss my findings from the structural studies of EGFR kinases in complex with

exon 20 insertion inhibitors and provide insights into structural factors that may contribute to mutant selectivity.



Scheme 4.1 – General kinetic mechanisms of covalent inactivation and reversible inactivation

4.2: Materials and Methods

Kinetics of inhibition by EGFR TKIs

Inhibitor kinetic assays were performed using the HTRF KinEASE tyrosine kinase assay kit (Cisbio) as described in the previous chapter, but reactions were quenched at different time points. Inhibitors (10 mM DMSO stocks) were dispensed into black 384-well plates using an HP D300e dispenser (Hewlett-Packard), normalized to a 1% final DMSO concentration, and 100 µM ATP was dispensed using a Multidrop Combi dispenser (ThermoFisher). For inhibitor kinetic experiments, assay buffer containing purified EGFR at a final concentration of 5 nM (wild-type EGFR) or 0.1 nM (D770_N771insSVD and V769_D770insASV) was dispensed to initiate the reactions. The reactions were quenched at different time points at room temperature using the detection reagent from the HTRF KinEASE assay kit. In a typical experiment, each column of wells was quenched sequentially with an interval of 2 minutes using the HP D300e dispenser. The FRET signal ratio was measured at 665 and 620 nm using a PHERAstar microplate reader (BMG LABTECH). Data were processed using GraphPad Prism and fit to a one-phase association model to determine k_{obs}.

Crystallization and structure determination

Purified EGFR kinase domain with affinity tags cleaved were crystallized via the hanging drop method at room temperature. EGFR(T790M/V948R) at ~3 mg/mL was supplemented with 1 mM adenylyl-imidodiphosphate (AMP-PNP) and 10 mM $MgCl_2$ and drops containing 1 µL protein+1 µL well solution were set above wells containing 0.1 M Bis-Tris pH=5.7 and 25-30%

(w/v) PEG 3,350. WT EGFR was similarly crystallized in the presence of 1 mM AMP-PNP and 10 mM $MgCl_2$ over wells containing 0.1 M MES pH 6.5 and 0.9-1.2 M sodium citrate. Inhibitors were soaked into crystals overnight at a concentration of 1 mM except for BAY-33 for which solid compound was added to the drop containing crystals. Crystals were cryoprotected in well solution supplemented with 20% ethylene glycol prior to flash freezing in liquid nitrogen.

Diffraction data were collected at the Advanced Photon Source at the Argonne National Laboratory on NE-CAT beamlines 24-ID-C and 24-ID-E at 100 K. Data were indexed, integrated, and scaled using Dials and xia2 compiled by SBGrid[267–269]. Structures were phased via molecular replacement and refined using Phenix with iterative rounds of manual model building in Coot[270,271]. Ligand restraints were generated using eLBOW in Phenix using the AM1 quantum mechanical optimization method[272]. Structures have been deposited in the Protein Data Bank (PDB) with the accession codes 8F1H (WT:TAS6417), 8F1W (TMVR:poziotinib), 8F1X (WT:TAK-788), 8F1Y (WT: poziotinib), and 8F1Z (WT:BAY-33).

4.3: Results

Kinetics of inhibition by EGFR TKIs

To gain further mechanistic insight into inhibitor binding and factors that could potentially contribute to mutant selectivity, I determined the inhibitor concentration giving half-maximal rate for covalent bond formation (K_I), an approximate measure of reversible binding affinity, and covalent inactivation rate (k_{inact}) for the covalent inhibitors, afatinib, poziotinib, TAK-788 and TAS6417. I also examined whether the reversible inhibitor BAY-33 inhibits EGFR kinase via a simple one-step mechanism with on and off rates, or a two-step mechanism, where the inhibitor concentration giving half-maximal inactivation rate (K_I) and the rate constant (k_2) could be determined (Scheme 4.1). EGFR kinase activity over time was monitored with different concentrations of representative covalent TKIs using an HTRF assay, and k_{obs} vs inhibitor concentration for wild-type EGFR, insASV, and insSVD were plotted to derive K_I and k_{inact} values for the covalent inhibitors (Figure 4.1) and k_2 for the reversible inhibitor BAY-33 (Figure 4.2). In addition, I used afatinib, which is more potent against wild-type EGFR than against insASV and insSVD, as a control.

In the kinetics of inhibition experiment, reaction progress curves of EGFR variants at various inhibitor concentrations were monitored. The k_{obs} values were derived from the progress curves and then plotted against inhibitor concentrations to obtain k_{inact} and K_I values. Consistent with the biochemical potency determination, afatinib has higher inactivation efficiency (k_{inact}/K_I) against wild-type EGFR than against insASV or insSVD, and poziotinib has similar inactivation efficiency against wild-type EGFR compared to against insASV or insSVD (Table 4.1). TAK-788 displays higher inactivation efficiency against insASV or insSVD than against wild-type EGFR

(k_{inact}/K_I values are ~20-40-fold greater). Interestingly, K_I values for TAK-788 are similar between the insertion mutants and wild-type EGFR, indicating similar reversible binding rates. However, k_{inact} values for the insertion mutants are at least 10 times higher than that of the wild-type, which indicates a faster rate of covalent bond formation against the mutant than the wild-type (Table 4.1; Figure 4.1).

BAY-33, which is a reversible EGFR inhibitor, is selective against the insertion variants in biochemical assays as shown in chapter 4. In the absence of a k_{inact} step, I sought to determine its source of mutant selectivity and asked whether BAY-33 binding occurs in multiple steps associated with a protein conformational change that may affect selectivity (Figure 4.2). Interestingly, I observed the k_{obs} vs [inhibitor] plots of BAY-33 were not linear but hyperbolic, which suggests inhibition via a multi-step mechanism. It also implies that the inhibitor does not have a weak reversible binding affinity like osimertinib[258,262]. Similar to TAK-788 where the second binding step affects selectivity, BAY-33 displays a 30-fold greater k_2/K_I value for the insertion variants than wild-type EGFR with the greatest difference in k_2. This suggests that binding might be accompanied by a significant secondary conformational change that "locks" BAY-33 in complex with EGFR kinase, and that insertion mutants could more easily adopt the conformational change than wild-type EGFR (Table 4.1). Collectively, kinetic analysis identified key differences between inhibitor binding and covalent adduct formation and offers insights that can guide inhibitor development and optimization.

Table 4.1

Kinetics of inactivation by exon 20 inhibitors

	WT			insASV			insSVD		
	k_{inact}/K_I ($s^{-1}\mu M^{-1}$)	K_I (nM)	k_{inact} (s^{-1})	k_{inact}/K_I ($s^{-1}\mu M^{-1}$)	K_I (nM)	k_{inact} (s^{-1})	k_{inact}/K_I ($s^{-1}\mu M^{-1}$)	K_I (nM)	k_{inact} (s^{-1})
Afatinib	0.75	6.6	0.0050	0.16	7.3	0.00091	0.26	9.4	0.0017
Poziotinib	4.3	0.28	0.0012	1.7	1.5	0.0025	9.7	0.46	0.0045
TAK-788	0.0090	24	0.00022	0.16	15	0.0023	0.36	24	0.0084
TAS6417	N.D.*	N.D.*	N.D.*	0.026	5.6	0.00015	0.036	8.5	0.00030
	k_2/K_I ($s^{-1}\mu M^{-1}$)	K_I (nM)	k_2 (s^{-1})	k_2/K_I ($s^{-1}\mu M^{-1}$)	K_I (nM)	k_2 (s^{-1})	k_2/K_I ($s^{-1}\mu M^{-1}$)	K_I (nM)	k_2 (s^{-1})
BAY-33	0.0025	157	0.00039	0.082	45	0.0036	0.091	52	0.0046

*Wild-type EGFR was not inhibited by TAS-6417 even at 2 μM in the presence of 100 μM ATP.

Figure 4.1

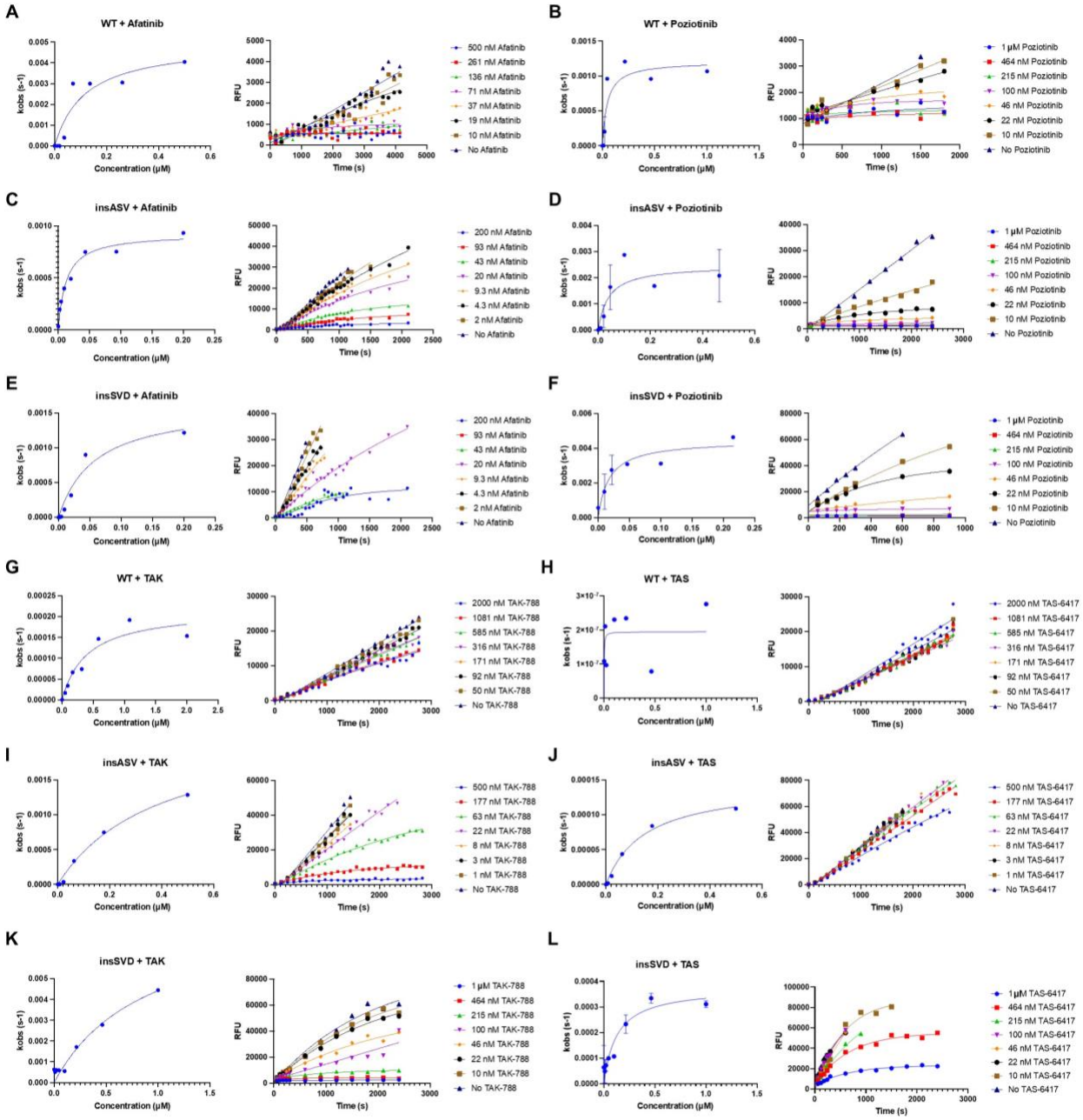

A
WT + Afatinib
kobs (s-1)
Concentration (μM)
RFU
Time (s)
500 nM Afatinib
261 nM Afatinib
136 nM Afatinib
71 nM Afatinib
37 nM Afatinib
19 nM Afatinib
10 nM Afatinib
No Afatinib
B
WT + Poziotinib
1 μM Poziotinib
464 nM Poziotinib
215 nM Poziotinib
100 nM Poziotinib
46 nM Poziotinib
22 nM Poziotinib
10 nM Poziotinib
No Poziotinib
C
insASV + Afatinib
200 nM Afatinib
93 nM Afatinib
43 nM Afatinib
20 nM Afatinib
9.3 nM Afatinib
4.3 nM Afatinib
2 nM Afatinib
No Afatinib
D
insASV + Poziotinib
E
insSVD + Afatinib
F
insSVD + Poziotinib
G
WT + TAK
2000 nM TAK-788
1081 nM TAK-788
585 nM TAK-788
316 nM TAK-788
171 nM TAK-788
92 nM TAK-788
50 nM TAK-788
No TAK-788
H
WT + TAS
2000 nM TAS-6417
1081 nM TAS-6417
585 nM TAS-6417
316 nM TAS-6417
171 nM TAS-6417
92 nM TAS-6417
50 nM TAS-6417
No TAS-6417
I
insASV + TAK
500 nM TAK-788
177 nM TAK-788
63 nM TAK-788
22 nM TAK-788
8 nM TAK-788
3 nM TAK-788
1 nM TAK-788
No TAK-788
J
insASV + TAS
500 nM TAS-6417
177 nM TAS-6417
63 nM TAS-6417
22 nM TAS-6417
8 nM TAS-6417
3 nM TAS-6417
1 nM TAS-6417
No TAS-6417
K
insSVD + TAK
1 μM TAK-788
464 nM TAK-788
215 nM TAK-788
100 nM TAK-788
46 nM TAK-788
22 nM TAK-788
10 nM TAK-788
No TAK-788
L
insSVD + TAS
1 μM TAS-6417
464 nM TAS-6417
215 nM TAS-6417
100 nM TAS-6417
46 nM TAS-6417
22 nM TAS-6417
10 nM TAS-6417
No TAS-6417

Figure 4.1. Kinetics of covalent inhibition. A-L, Reaction time course curves of EGFR wild-type, insASV or insSVD with different concentrations of inhibitors on the right side (RFU = relative fluorescence unit). Graphs on the left side show k_{obs} vs [inhibitor] to derive K_I and k_{inact} values. Error bars indicate SD (N=2). Data were processed using GraphPad Prism and fit to a one-phase association model to determine k_{obs}. Some points were excluded so the model could be fitted in the software. Note that wild-type EGFR was not inhibited by TAS-6417 even at 2 μM in the presence of 100 μM ATP.

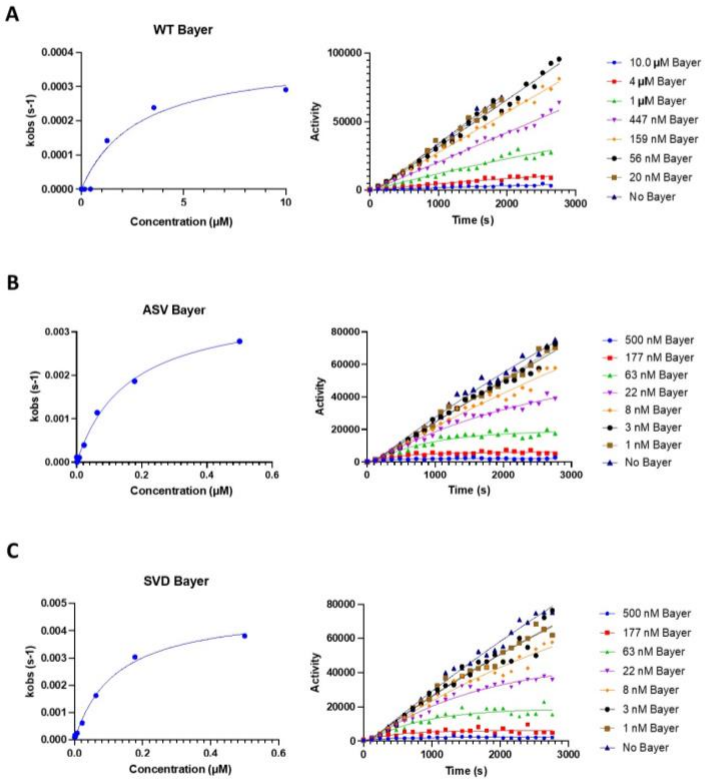


Figure 4.2. Kinetics of inhibition by BAY-33. A-C, Reaction time course curves of EGFR wild-type, insASV or insSVD with different concentrations of BAY-33 (right) and k_{obs} vs [inhibitor] (left) to derive K_I and k_2 values (activity shown as RFU = relative fluorescence unit). Data were processed using GraphPad Prism and fit to a one-phase association model to determine k_{obs}.

Crystal structures of EGFR kinase in complex with exon 20 inhibitors

To better interpret our potency and kinetic binding data, as well as to understand the molecular basis for mutant selectivity, we determined the crystal structures of EGFR kinase in complex with BAY-33 (PDB: 8F1Z), TAK-788 (PDB: 8F1X), TAS6417 (PDB: 8F1H), and poziotinib (PDB: 8F1Y for wild-type or 8F1W for T790M/V948R) (Figure 4.3; Figure 4.4; Table 4.2). Poziotinib, which has been used in clinical trials to treat patients with EGFR exon 20 insertions, represents a non-selective inhibitor that can be compared with mutant-selective inhibitors to understand mechanisms of selectivity. Unfortunately, I have not been able to crystallize insASV or insSVD due to tendency of aggregation of these two variants during concentrating efforts for crystallization. This limited my interpretation since I would not know if any additional interactions with mutant-selective inhibitors were present in insASV or insSVD compared to WT EGFR. However, even with WT EGFR structures, I was able to find some interesting binding features of mutant-selective inhibitors that were not commonly present in other non-mutant-selective inhibitors.

The crystal structures revealed that BAY-33 and TAS6417 induced F723 in the phosphate-binding loop (P-loop) to flip in and stack on top of the inhibitors to make favorable hydrophobic interactions with the TKIs, which makes the kinase adopts a “bent” P-loop conformation not commonly seen in other EGFR-TKI structures (Figure 4.5, B-C). Similar to other protein kinase co-crystal structures where the conserved P-loop aromatic ring contacts the ligand to achieve significant selectivity, such as imatinib bound to Abl and SU5402 bound to FGFR1, this “flipped” F723 conformation has been proposed to enhance the binding of osimertinib to EGFR and may also affect the potency or selectivity of these inhibitors[94,273–275]. Both BAY-33 and TAS6417 make hydrogen bonding interactions with the hinge region and T854 preceding the DFG motif, and TAS6417 is covalently bonded to C797 (Figure 4.4, A-B, F).

TAK-788 is a structural analog of osimertinib with an additional isopropyl ester attached to its pyrimidine (Figure 1.6). In the crystal structure, the core of TAK-788 binds similarly to osimertinib, with the additional isopropyl ester simultaneously hydrogen bonding with T790 and T854 (Figure 4.4, E). A common observation with the mutant-selective inhibitors BAY-33, TAS6417, and TAK-788 is the simultaneous placement of a hydrophobic group adjacent to the gatekeeper 790 position and a hydrogen bond to T854, the latter of which is not observed in structures with the non-selective inhibitor poziotinib (Figure 4.4, B-F).

Table 4.2

Data collection and refinement statistics

	BAY-33	poziotinib (TMVR)	poziotinib (WT)	TAK-788 (mobocertinib)	TAS6417 (CLN-081)
PDB Accession	8F1Z	8F1W	8F1Y	8F1X	8F1H
Resolution range	46.27 - 2.8 (2.9 - 2.8)	69.16 - 3.2 (3.31 - 3.2)	59.43 - 2.75 (2.85 - 2.75)	51.42 - 2.3 (2.38 - 2.3)	46.27 - 2.8 (2.9 - 2.8)
Space group	*I* 2 3	*P* 1 2_1 1	*I* 2 3	*I* 2 3	*I* 2 3
Unit cell *a*, *b*, *c* (Å) a, b, γ (°)	146.32 146.32 146.32 90 90 90	70.79 102.53 87.33 90 102.31 90	145.57 145.57 145.57 90 90 90	145.43 145.43 145.43 90 90 90	146.32 146.32 146.32 90 90 90
Total reflections	66621 (6882)	60156 (6113)	67815 (6927)	117782 (11837)	66621 (6882)
Unique reflections	12960 (1303)	19809 (1962)	13469 (1330)	22536 (2256)	12960 (1303)
Multiplicity	5.1 (5.3)	3.0 (3.1)	5.0 (5.2)	5.2 (5.2)	5.1 (5.3)
Completeness (%)	99.55 (99.85)	97.18 (97.14)	99.65 (100.00)	98.52 (100.00)	99.55 (99.85)
Mean I/sigma(I)	10.16 (1.44)	3.30 (0.75)	8.31 (1.22)	12.78 (1.14)	10.16 (1.44)
Wilson B-factor	52.75	55.77	72.37	53.80	52.75
R-merge	0.1305 (0.6979)	0.2148 (0.584)	0.1401 (1.193)	0.07429 (0.7518)	0.1305 (0.6979)
R-meas	0.1452 (0.7747)	0.2605 (0.705)	0.1567 (1.329)	0.08233 (0.836)	0.1452 (0.7747)
R-pim	0.06271 (0.3325)	0.1454 (0.3908)	0.06895 (0.5791)	0.03453 (0.3587)	0.06271 (0.3325)
CC1/2	0.994 (0.753)	0.96 (0.796)	0.988 (0.585)	0.998 (0.689)	0.994 (0.753)
Reflections used in refinement	12956 (1303)	19674 (1939)	13462 (1330)	22532 (2256)	12956 (1303)
Reflections used for R-free	613 (55)	1025 (100)	639 (66)	1131 (132)	613 (55)
R-work	0.1865 (0.2823)	0.2311 (0.3154)	0.2063 (0.3386)	0.2052 (0.2871)	0.1865 (0.2823)
R-free	0.2085 (0.3376)	0.2721 (0.3768)	0.2308 (0.4567)	0.2212 (0.3109)	0.2085 (0.3376)
Number of non-hydrogen atoms	2514	9658	2391	2639	2514
macromolecules	2460	9526	2338	2551	2460
ligands	43	132	46	43	43
solvent	11	0	7	45	11
Protein residues	307	1185	293	317	307
RMS(bonds)	0.005	0.004	0.003	0.176	0.005
RMS(angles)	0.70	0.71	0.64	2.64	0.70

Ramachandran favored (%)	95.38	93.55	94.04	94.21	95.38
Ramachandran allowed (%)	3.96	5.67	5.26	4.50	3.96
Ramachandran outliers (%)	0.66	0.77	0.70	1.29	0.66
Rotamer outliers (%)	3.33	4.76	3.91	5.00	3.33
Clashscore	5.20	7.95	5.87	5.60	5.20
Average B-factor	55.09	56.83	75.61	69.57	55.09
macromolecules	55.06	56.92	75.55	69.91	55.06
ligands	58.97	50.62	81.21	64.49	58.97
Solvent	46.33		61.43	55.33	46.33

Statistics for the highest-resolution shell are shown in parentheses.

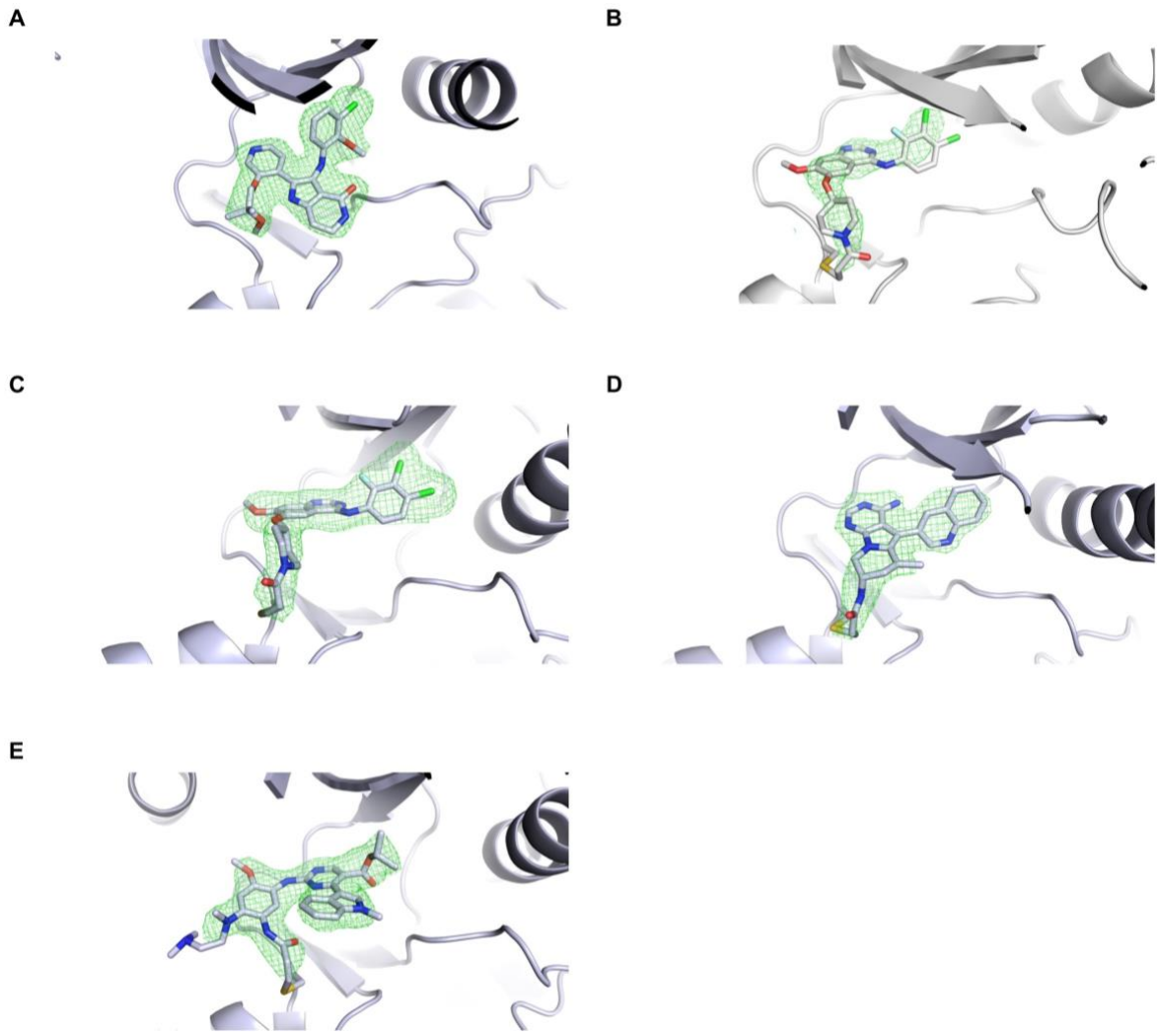


Figure 4.3. Ligand densities for inhibitor crystal structures. Positive F_o-F_c density at 3 σ is shown for **A,** BAY-33 (PDB 8F1Z), **B,** poziotinib (EGFR T790M/V948R, PDB 8F1W), **C,** poziotinib (WT EGFR, 8F1Y), **D,** TAS6417 (8F1H), and **E,** TAK-788 (PDB 8F1X).

Figure 4.4

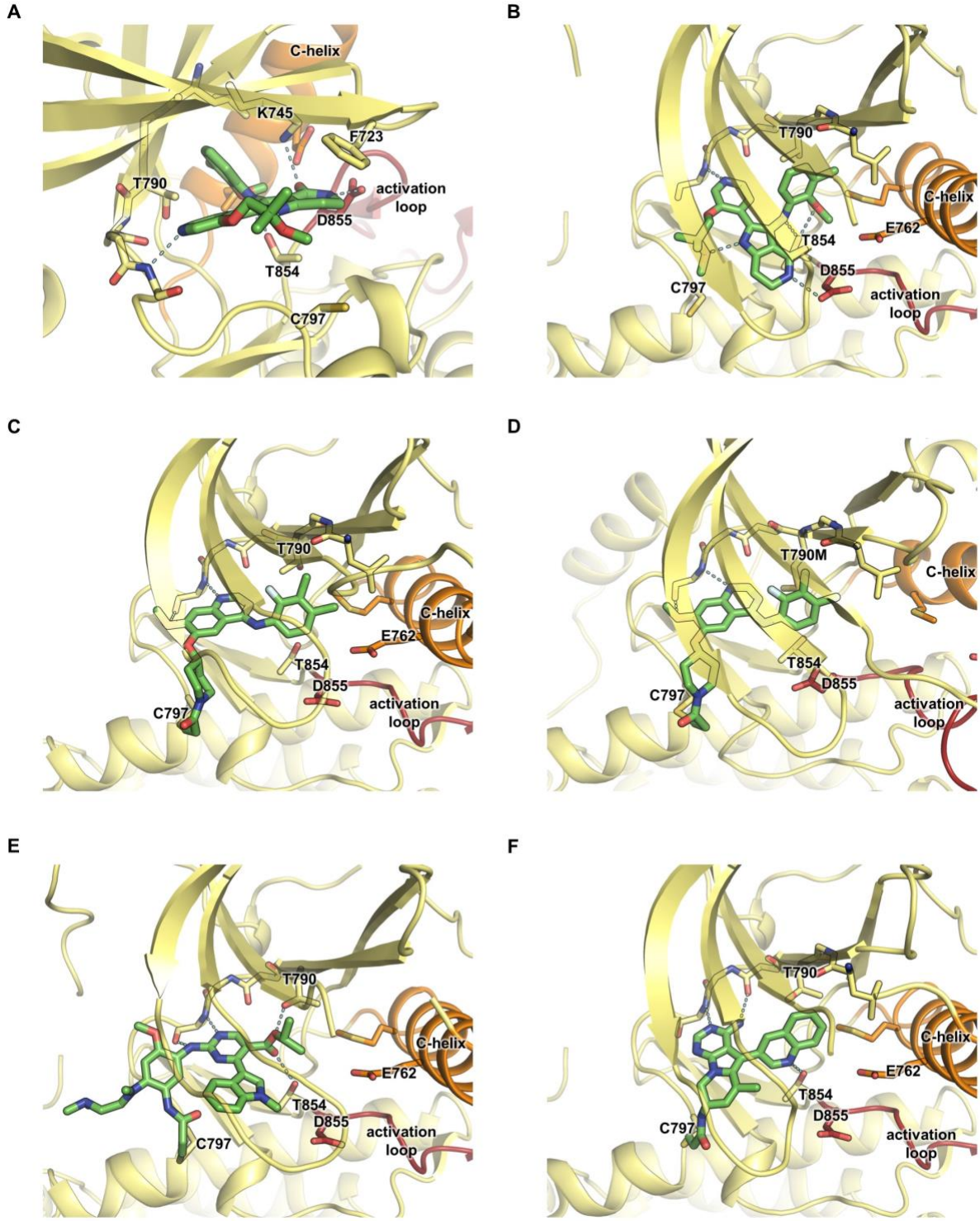

A
C-helix
K745
F723
T790
activation loop
D855
T854
C797
B
T790
C-helix
E762
T854
D855
C797
activation loop
C
T790
C-helix
E762
T854
D855
C797
activation loop
D
T790M
C-helix
T854
D855
C797
activation loop
E
T790
C-helix
E762
T854
D855
C797
activation loop
F
T790
C-helix
E762
T854
D855
C797
activation loop

Figure 4.4. Crystal structures of EGFR kinase in complex with exon 20 inhibitors. A-B, Two alternative views of active site for wild-type EGFR in complex with BAY-33 (PDB 8F1Z). **A,** Side view of active site for wild-type EGFR in complex with BAY-33 showing interactions with the hinge region, C-helix, as well as F723 (PDB 8F1Z). **B,** Alternative top view of BAY-33 showing the complete compound structure and its interactions with the protein (PDB 8F1Z). **C,** Top view of active site for wild-type EGFR in complex with poziotinib (PDB 8F1Y). **D,** Top view of active site for EGFR T790M/V948R in complex with poziotinib (PDB 8F1W). **E,** Top view of active site for wild-type EGFR in complex with TAK-788 (PDB 8F1X). **F,** Top view of active site for wild-type EGFR in complex with TAS6417 (PDB 8F1H).

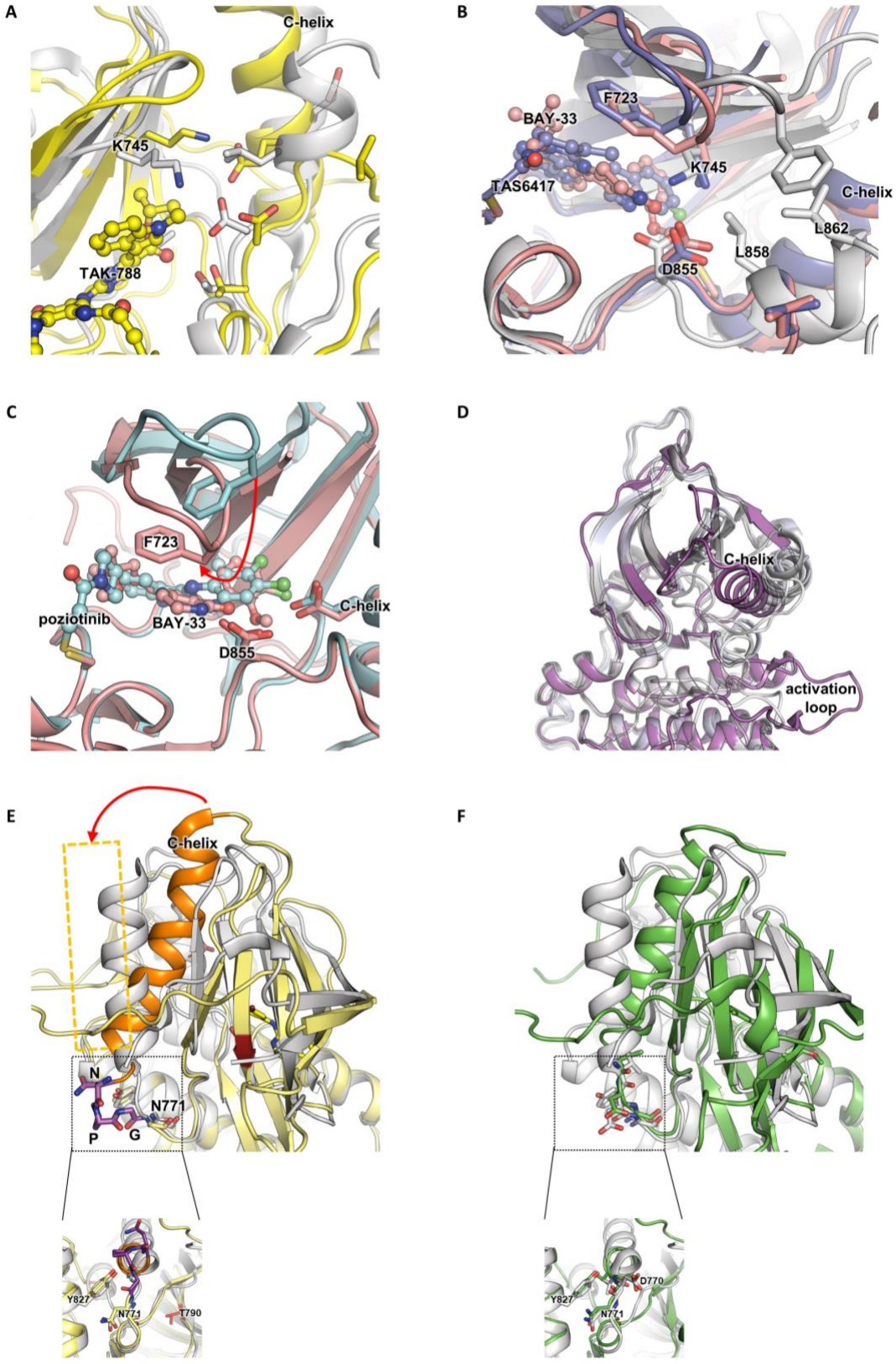

A
C-helix
K745
TAK-788
B
F723
BAY-33
K745
TAS6417
C-helix
L862
L858
D855
C
F723
poziotinib
BAY-33
C-helix
D855
D
C-helix
activation
loop
E
C-helix
N
N771
P
G
Y827
N771
T790
F
Y827
D770
N771

Figure 4.5. Structural insights for EGFR kinase in complex with exon 20 inhibitors. Crystal structures of EGFR kinase were superimposed to illustrate that TAK-788, BAY-33 and TAS6417 prefer to bind the kinase in the active conformation, and that the insertion variants of EGFR predominantly adopt the active conformation. **A,** Structure of wild-type EGFR in the active form in complex with TAK-788 (in yellow, PDB: 8F1X) superimposed onto wild-type EGFR in the inactive form (in grey, PDB: 4HJO). **B,** Structure of wild-type EGFR in complex with BAY-33 (in pink, PDB: 8F1Z) and wild-type EGFR in complex with TAS6417 (in purple, PDB: 8F1H) superimposed onto wild-type EGFR in the inactive form (in grey, PDB: 4HJO). **C,** Structure of wild-type EGFR in complex with BAY-33 (in pink, PDB: 8F1Z) superimposed onto wild-type EGFR in complex with poziotinib (in cyan, PDB: 8F1Y). **D,** Structure of insNPG/V948R (in purple, PDB: 7LGS) superimposed onto EGFR V948R (in light grey, PDB: 4HJO), L858R/V948R (in dark grey, PDB: 7K1I), T790M/V948R (in white, PDB: 8F1W), L858R/T790M/V948R (in light blue, PDB: 4I22). **E,** Structure of active insNPG (in yellow with C-helix highlighted in orange, PDB: 4LRM) superimposed onto inactive wild-type (in grey, PDB: 4HJO). The box with dashed lines indicates the expected position of C-helix in a hypothetical inactive conformation of insNPG. Note that the position of the box is parallel but to the left of the C-helix of inactive wild-type. **F,** Structure of active wild-type EGFR (in green, PDB: 2ITX) superimposed onto inactive wild-type (in grey, PDB: 4HJO). The zoomed-in images for **E** and **F** offer alternative views of Y827 “locking” the position of the inserted residues in insNPG.

4.4: Discussion

Structural insights into mutant selectivity of TAK-788, BAY-33, and TAS6417

Upon investigation of the active conformation of wild-type EGFR structure in complex with TAK-788 superimposed with the inactive conformation of wild-type EGFR, I noticed that it is unfavorable for the inactive conformation of EGFR kinase to accommodate TAK-788 in the same binding mode as seen in the co-crystal structure of active WT with TAK-788 (Figure 4.5, A). Compared to osimertinib, TAK-788 has an additional isopropyl ester, which makes the compound bulkier and harder to accommodate. In the inactive conformation, V726, K745, and the backbone of I744 are closer to the active site, which would make them too close to the isopropyl ester of TAK-788. The isopropyl ester might have to rotate and adopt a less favorable conformation as seen previously[276] (PDB: 7A6K), to avoid steric clash, so that it would lose hydrogen bonding interactions with T790 and T854 (Figure 4.5, A). Therefore, TAK-788 favors the active conformation of EGFR kinase, much more than osimertinib, which makes TAK-788 selective against insertion mutants as they have a heavily shifted equilibrium towards the active conformation. As a result, TAK-788 would more easily adopt its preferred binding mode to form covalent bond with C797 in the insertion mutants, which illustrates the differences in rates of covalent bond formation between wild-type and insASV or insSVD mutant.

Structural analysis of wild-type EGFR in complex with BAY-33 or TAS6417 revealed F723 in the turn of the P-loop flips and makes stacking interactions with the pyrrolopyridinone of BAY-33 or methyl-tetrahydropyridine of TAS6417 (Figure 4.5, B), which is not as favorable in the inactive conformation as F723 mostly makes hydrophobic interactions with L858 or L862 of the folded activation loop. In addition, compared to TAS6417, the pyrrolopyridinone of BAY-33 makes

greater hydrophobic contact with F723 as the phenylalanine ring is directly on top (face-to-face) of the pyrrolopyridinone of BAY-33 with a distance ~3.4 Å, while it is only partially on top (offset, or staggered) of methyl-tetrahydropyridine of TAS6417 with the closest distance ~3.8 Å. Non-aromatic ring stacking has been previously described as a strong interaction that favors short-distance face-to-face stacking[277,278]. Therefore, it suggests that BAY-33 might favor the active conformation of EGFR even more than TAS6417, which could explain the greater mutant selectivity of BAY-33 in the biochemical setting. Similar to the case of TAK-788, it would be easier for BAY-33 to induce the additional conformational change (movement of F723) in the insertion mutants as they are mostly in the active conformation, which is consistent with the inhibitor kinetics data. Based on the differences in rate of covalent bond formation for TAK-788 and differences in rates of conformational change induced by BAY-33, it is likely that an important feature of exon 20 insertion-selective inhibitor is their ability to form a covalent bond or induce a conformational change after binding with mutant kinase faster than with wild-type kinase. This may be accomplished with an inhibitor that prefers binding to the active conformation of the kinase or could favorably interact with F723 as I observe in our crystal structures (Figure 4.5, C).

Speculation about conformational dynamics on inhibitor selectivity

Previous work showed that the oncogenic nature of EGFR L858R is dimerization dependent, while exon 20 insertion variants are not[115]. Based on our kinetic data and available structures, I hypothesize that insertion mutants with insertions in the loop following C-helix, such as insASV, insSVD, and insNPG, have elevated basal activity due to a large shift in equilibrium towards the active conformation. Selectivity for the active versus inactive conformation of EGFR has been previously invoked as a potential mechanism of mutant-selectivity for EGFR inhibitor,

including gefitinib[16]. I suspect that this mechanism may also apply to the mutant-selective compounds identified here.

With the dimer-disrupting mutation V948R, the equilibrium for EGFR kinase is heavily shifted towards the inactive conformation. Regardless of this, the crystal structure of EGFR insNPG/V948R is uniquely in the active conformation while all other available crystal structures of EGFR V948R are in the inactive conformation (Figure 4.5, D). In the insNPG/V948R or previously determined insNPG structures, the Y827 interacts with the inserted loop by forming hydrogen bonds to the carbonyl of V769 and the amino of the inserted glycine at the end of C-helix, which limits the space that the inserted residues could reside in and the residues following N771 all align very well with wild-type kinase. Comparison of active versus inactive wild-type EGFR structures reveals that the C-helix rotates around the pivot point at its C-terminal end, which suggests the same for insertion mutants (Figure 4.5, E-F). In order for insertion mutants to adopt the inactive conformation, the αC-β3 loop at the N-terminal end of C-helix could be stretched to an unfavorable conformation due to a shift in the position of C-helix compared to WT, which significantly alters the equilibrium in the direction favoring the active conformation (Figure 4.5, E).

Speculation about the role of inactive conformation on inhibitor binding

My hypothesis is that in the active conformation, inhibitors would more easily dissociate from the kinase than in the inactive conformation, possibly due to lack of the helical turns formed by the activation loop that interacts with F723 to partially block the access to the active site. Even though the inactive conformation of EGFR kinase might not be easily accessible for most of the TKIs, it could play a role in “locking” the target compound in the ATP binding pocket due to formation of helical turns of the activation loop, which gives the compound more time to adopt a

favorable binding mode with kinase[70,97,102]. The C-helix “out” conformation could also give the compound more space to adjust its binding mode after initial binding. This might be able to explain why 1st – 3rd generations of TKIs are more potent against wild-type EGFR than insertion mutants, or at least equipotent, even though insertion mutants have weaker affinities to ATP compared to wild-type EGFR.

If the kinase adopts the active conformation for the vast majority of time, inhibitors might not have enough time before dissociation to adopt a favorable binding mode, or in the case of irreversible inhibitors, to form a covalent bond with the kinase, compared to the mostly inactive WT. The L858R mutation is an exception since R858 could not be accommodated in the helical turns formed by the activation loop as it would be placed in a hydrophobic core, so the activation loop would be in an extended form even in the inactive conformation[185,279,280]. However, the L858R mutant has a much weaker affinity to ATP than WT EGFR, which makes it sensitive to TKIs[16,127]. Unlike L858R, exon 20 insertion variants have only modestly lower ATP affinities than WT EGFR and do not rely on dimerization to be active, which suggests that they adopt the C-helix “in” conformation as the default mode. As a result, many ATP competitive inhibitors are not selective against exon 20 insertion variants, while inhibitors that preferably bind to the inactive conformation, including type-1.5 inhibitors and allosteric inhibitors, have much less potency against insertion variants than WT EGFR.

4.5: Contributions

I expressed and purified all EGFR proteins, performed the inhibitor kinetics assays, and did structural analysis. Dr. Tyler S. Beyett expressed and purified wild-type proteins for structural studies, crystallized proteins, and determined crystal structures. Jaimin K. Rana, Ilse K. Schaeffner, and Jhasmer Santana crystallized proteins. Prof. Michael J. Eck supervised the research.

Chapter 5

Biochemical screening for mutant-selective inhibitors

5.1: Introduction

EGFR exon 20 insertions represent a unique class of EGFR mutations in NSCLC, which rarely respond to EGFR TKIs[128,199]. Recently, amivantamab, a bispecific antibody capable of targeting EGFR and MET, received the accelerated approval from the FDA for the treatment of EGFR exon 20 insertions, soon followed by TAK-788, which is a novel EGFR TKI designed based on osimertinib, the current standard of care for treating NSCLC patients with classical EGFR mutations[201,226,231]. However, based on clinical trial results, neither of these treatments has a wide enough therapeutic window to minimize toxicities and dose interruptions or discontinuation, which limits their clinical efficacies[213]. Therefore, there is still an unmet need for a treatment with a wider therapeutic window against EGFR exon 20 insertions. In addition, it is almost inevitable for the resistance mutations, such as C797S, to occur with treatment of potent covalent inhibitors[194,198,202]. As a result, it is important to develop additional exon 20 insertion mutant-specific inhibitors, in particular with a reversible mechanism of action, to expand the arsenal for treating patients whose lung tumors are driven by EGFR exon 20 insertions.

To find such novel inhibitors that are insertion mutant selective, and preferably reversible inhibitors instead of covalent ones, I set up a large screen of approximately six thousand compounds using the focused kinase inhibitor library from the Gray lab at Stanford University. In this chapter, I will present my results from the large screen and the follow-up studies of selected hits based on cheminformatics to show that a lead reversible compound, YKL-04-125, is selective against insASV or insSVD. I will later discuss the importance of these findings and conclude with possible directions in further inhibitor development.

5.2: Materials and Methods

Inhibitors

Unless otherwise noted, inhibitors used in this screening were from the Gray focused kinase inhibitor library developed at Dana-Farber Cancer Institute. Compound stocks of 10 mM were pin transferred to new plates, diluted in DMSO to 5 mM intermediate stocks and stored at -20 ºC. These ~6000 compounds are diverse inhibitor scaffolds explored in the Gray lab over the past two decades, for mostly kinase targets including EGFR and CDK7[152,281].

EGFR inhibitor screening and IC_{50} determination

Inhibition assays were performed using the HTRF KinEASE tyrosine kinase assay kit (Cisbio) according to the manufacturer's protocol, as previously described. Inhibitors from the intermediate stocks were normalized to a 1% final DMSO concentration in 384-well plates with each well representing a different inhibitor. For screening experiments, assay buffer containing purified EGFR at a final concentration of 5 nM was dispensed using a Multidrop Combi dispenser (ThermoFisher) and incubated with compounds at room temperature for 30 min. Reactions were initiated with 1 mM ATP and allowed to proceed for 30 min (Wild-type) or 10 min (mutants) at room temperature before being quenched using the detection reagent from the HTRF KinEASE assay kit. The FRET signal ratio was measured at 665 and 620 nm using a PHERAstar microplate reader (BMG LABTECH). The experiment was performed in duplicate, and the hits were selected via cheminformatics analysis. Follow-up IC_{50} determination of selected hits was performed using the previously described HTRF assay with increasing inhibitor concentrations and a constant ATP

concentration of 1 mM. Data of follow-up IC_{50} determination were processed using GraphPad Prism and fit to a three-parameter dose-response model with a Hill slope constrained to -1.

Scaffold-driven ontology-based pattern identification (OPI) analysis

In the scaffold-driven OPI analysis, numerical representations of compounds were first generated using the extended connectivity fingerprint (ECFP), and then compounds were clustered into chemical scaffold families whose members shared a structure similarity greater than a Tanimoto value of 0.9, since the Gray library is primarily kinase inhibitors with limited and biased chemical space[282,283]. The enrichment for each compound cluster was calculated for their activities in insASV, insSVD, and WT EGFR. With the statistical measures of their enrichments, compounds that are enriched in insASV or insSVD but not in WT EGFR were selected. Finally, the selections were further filtered based on percent inhibition, where the inhibition for insertion mutants needs to be more than 50% with less than 10% inhibition for WT EGFR.

5.3: Results

A high-throughput screen for inhibitors selective for insASV and insSVD

To find novel scaffolds that might present stronger selectivity towards EGFR exon 20 insertion mutants than wild-type EGFR, I devised a screen of approximately 6,000 compounds in the focused kinase inhibitor library from the Gray lab. The screening was performed at a single compound concentration of 5 μM, with 5 nM EGFR kinase and 1 mM ATP, which is approximately the cellular concentration of ATP under physiological conditions[284]. Using the previously described HTRF assay, I screened the library against my purified EGFR insASV and insSVD, and then counter-screened against purified wild-type EGFR kinase (Figure 5.1).

A holistic view of the screening results showed that, consistent with my selectivity plots in Chapter 3, inhibitor sensitivities of insASV and insSVD are largely correlated, although there are outliers that are more potent in one variant than the other (Figure 5.2, A). In contrary to the correlation between insSVD and insASV, inhibitor sensitivities of WT EGFR are much less correlated with those of insASV, with compounds scattered all over the selectivity plot, further confirming that insASV and insSVD are not inherently harder to target than wild-type EGFR but rather prefer different scaffolds (Figure 5.2, B).

Further cheminformatics analysis of the screening results identified eight hits that inhibit more than 50% of kinase activity in insASV and insSVD while inhibiting WT EGFR for less than 10%. Close examination of these selected hits showed that five of them are reversible inhibitors, YKL-04-125, YKL-05-069, YKL-05-152, YKL-05-179, and YKL-06-051, while the remaining three, HG-14-130-02, HG-14-133-01, and WZ 8-057, are covalent inhibitors, with most of the hits belonging to two structural clusters except HG-14-130-02 (Figure 5.3; Figure 5.4, A). The

distribution of inhibitor potencies in insASV, insSVD, and WT EGFR illustrated that these eight selected hits target insASV and insSVD while sparing WT EGFR (Figure 5.4, B-D).

Follow-up IC_{50} determination of selected hits

To confirm the potency and selectivity of these eight selected hits, I determined their biochemical IC_{50} values against insASV and insSVD using the HTRF assay described in Chapter 3, and compared these to the wild-type EGFR (Figure 5.5; Table 5.1). I used 1 mM ATP, same as the screening condition, to obtain dose-response curves for these selected hits. The results showed that although all of the hits displayed some mutant selectivity, YKL-04-125 and HG-14-130-02 are more potent and mutant selective than the others, with >100-fold selectivity against WT EGFR. None of the selected hits showed significant activity towards WT even at 10 μM inhibitor concentration. Consistent with the trend seen in large-scale screening, these eight selected hits inhibited insASV and insSVD with comparable potency (Figure 5.5, C-F; Table 5.1).

Notably, the most mutant-selective hit, YKL-04-125, is a reversible inhibitor that makes it more valuable for further development. All of its analogs, YKL-05-069, YKL-05-152, YKL-05-179, and YKL-06-051, are also reversible inhibitors, with YKL-05-069 being the closest analog of YKL-04-125 among the selected hits. Interestingly, apart from a slight position shift of substituent groups on the phenyl ring, the only difference between YKL-05-069 and YKL-04-125 is that, in the chemical structure of YKL-04-125, the isopropyl ether on a phenyl ring is replaced by a trifluoromethyl group, which leads to a significant increase (~100 fold) in potency against insASV and insSVD (Figure 5.3; Table 5.1).

Figure 5.1

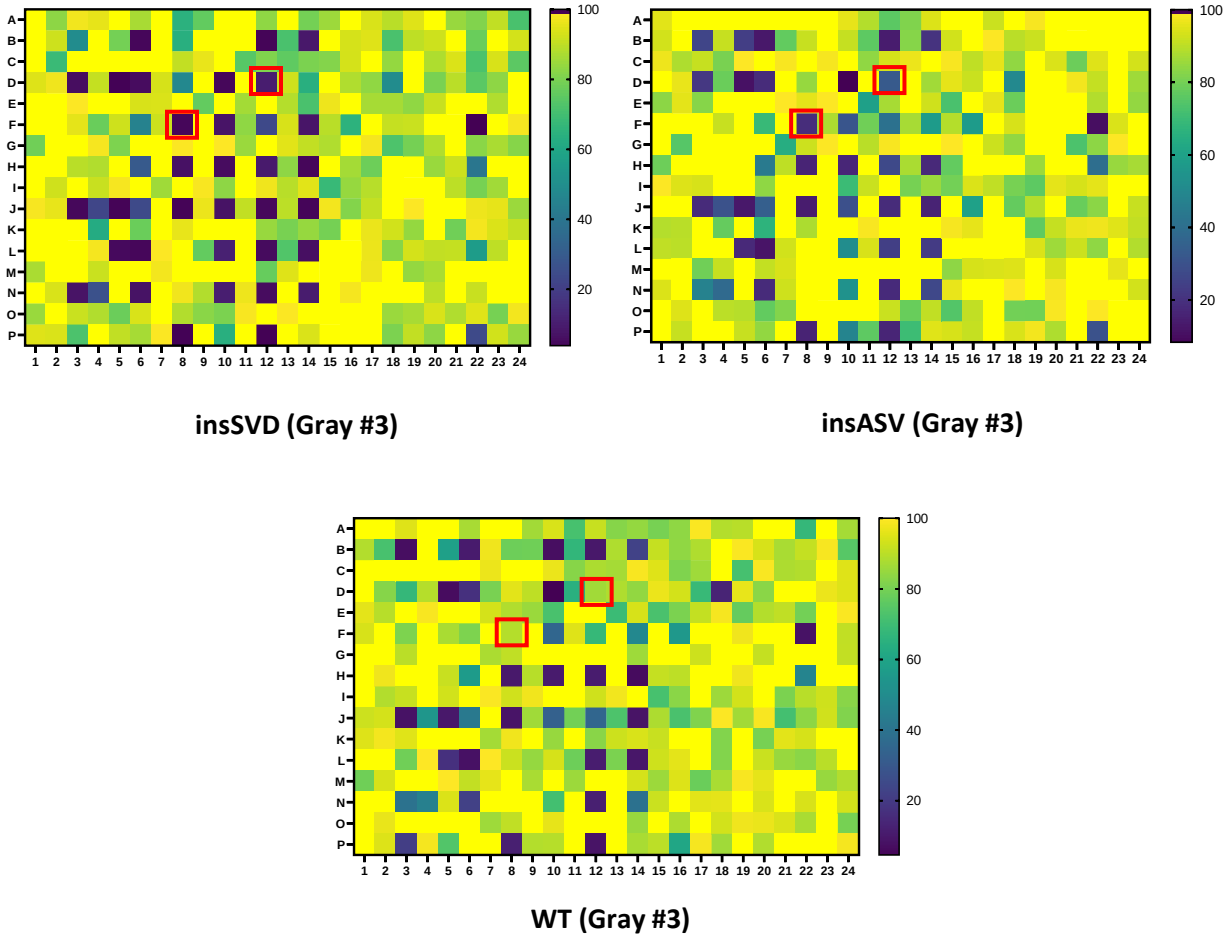
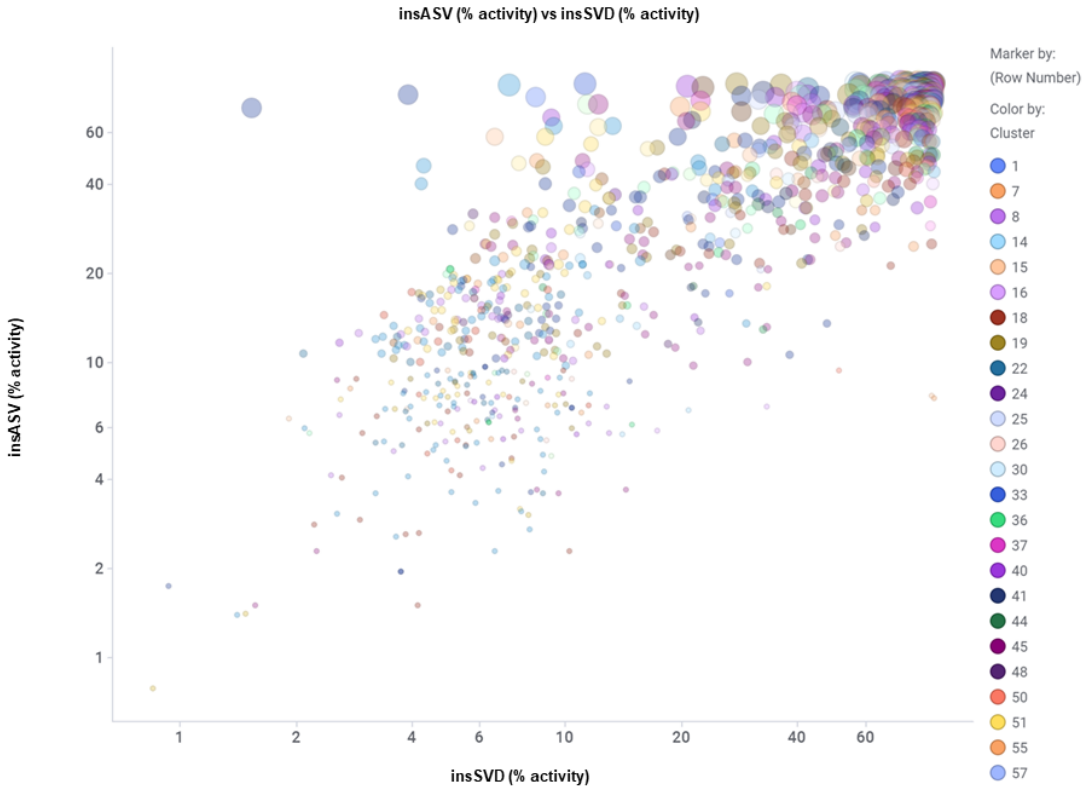

A B C D E F G H I J K L M N O P
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24
100 80 60 40 20
insSVD (Gray #3)
insASV (Gray #3)
WT (Gray #3)

Figure 5.1. Representative 384-well plate with kinase inhibitors screened against insSVD, insASV, and wild-type EGFR. A total of 16 unique 384-well plates with kinase inhibitors synthesized in the Gray lab were screened. Each plate was analyzed as a heatmap with spectrum indicating the activity level of each well, with bright yellow showing highest activity while dark blue showing lowest activity. The activity of each well, measured as product formation within 20 minutes by HTRF assay, was normalized to the average of those wells with no inhibitor presence for each plate. As the representative plate, Gray #3, showed, the vast majority of compounds have similar potencies against insASV, insSVD, and WT EGFR kinase. However, inhibitors with some selectivity, such as the ones in F8 and D12 (highlighted by red boxes), could be visually identified by comparing these heatmaps of different EGFR variants.

A


insASV (% activity) vs insSVD (% activity)
insASV (% activity)
insSVD (% activity)
1
2
4
6
10
20
40
60
Marker by:
(Row Number)
Color by:
Cluster
1
7
8
14
15
16
18
19
22
24
25
26
30
33
36
37
40
41
44
45
48
50
51
55
57

B

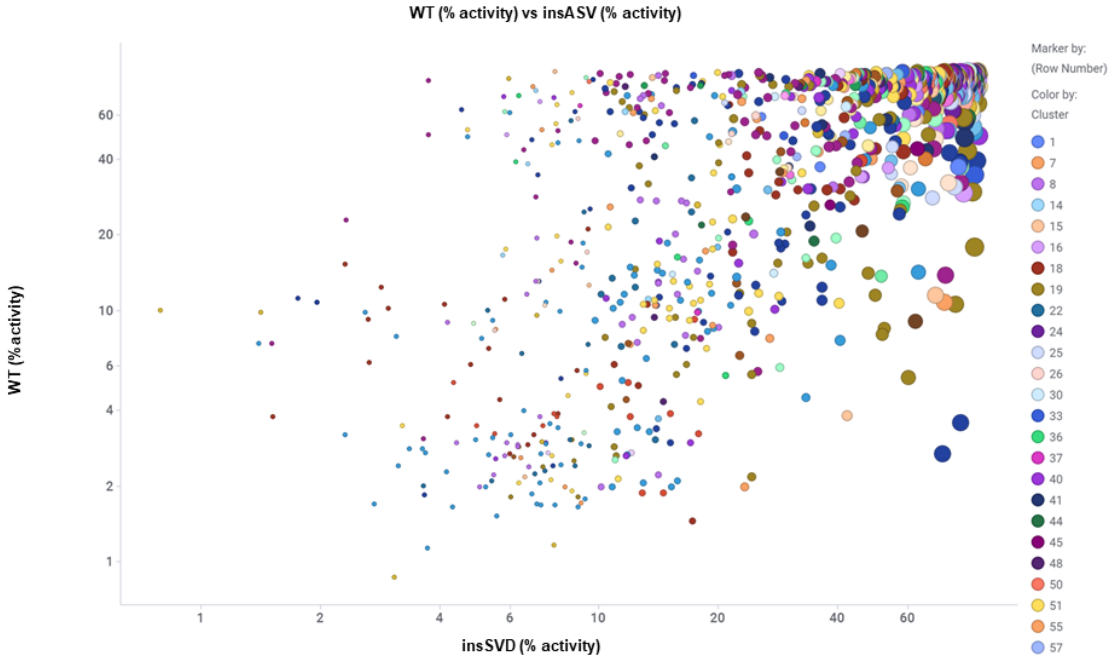

WT (% activity) vs insASV (% activity)
WT (% activity)
insSVD (% activity)
1
2
4
6
10
20
40
60
Marker by:
(Row Number)
Color by:
Cluster
1
7
8
14
15
16
18
19
22
24
25
26
30
33
36
37
40
41
44
45
48
50
51
55
57

Figure 5.2. Correlation of inhibitor sensitivities of EGFR variants. A, Correlation plot of inhibitor sensitivities of insASV vs insSVD shown. Although there are outliers that are more potent in one variant than the other, the inhibitor sensitivities of insASV and insSVD are largely correlated, with most compounds showing similar inhibition in both variants. **B,** Correlation plot of inhibitor sensitivities of WT vs insASV shown. Compared to the insASV vs insSVD plot, insASV and WT EGFR are much less correlated, with compounds scattered all over the plot.

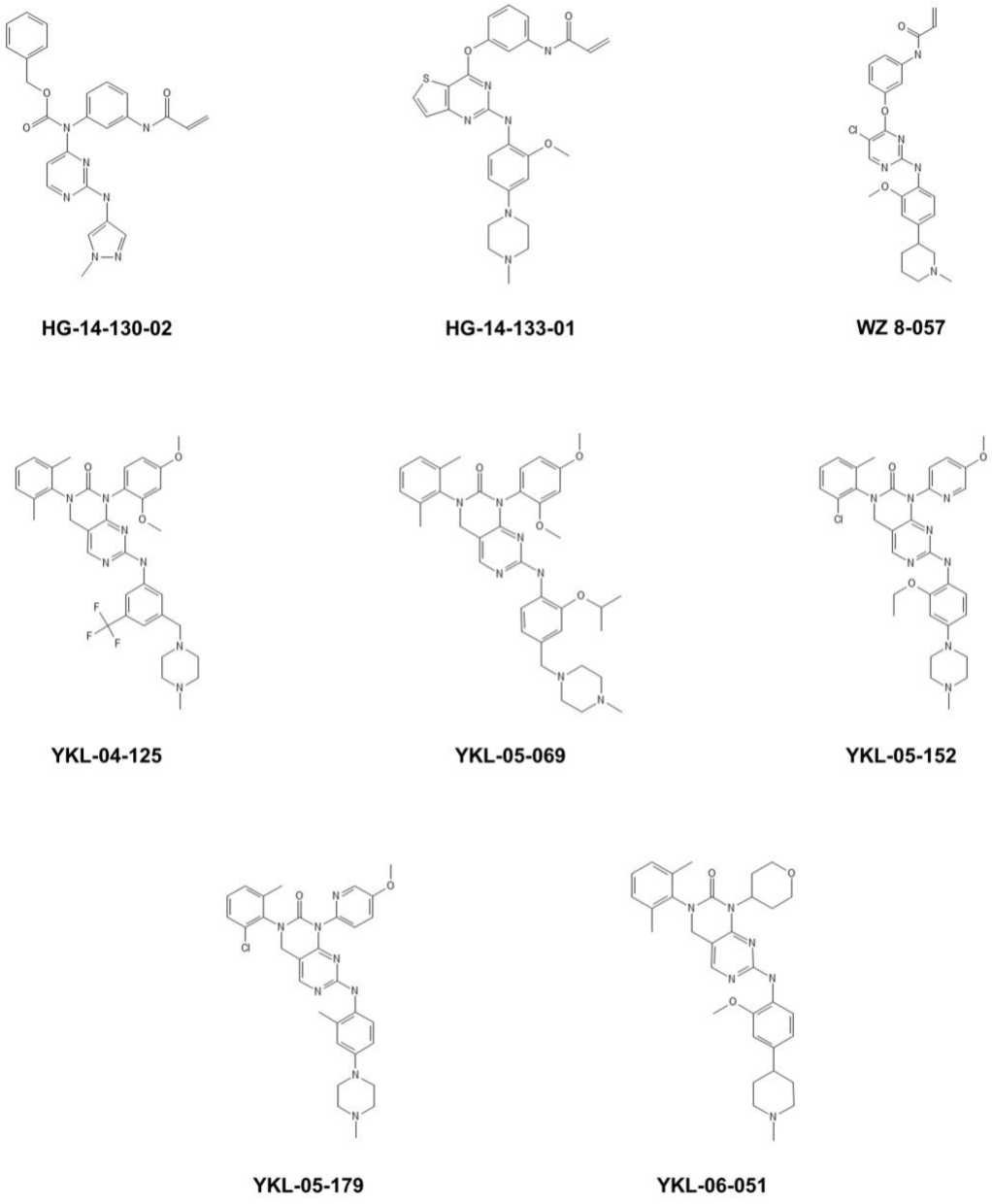


Figure 5.3. Chemical structures of selected hits. The three covalent inhibitors are in the first row, while the five reversible inhibitors are in the second and third rows.

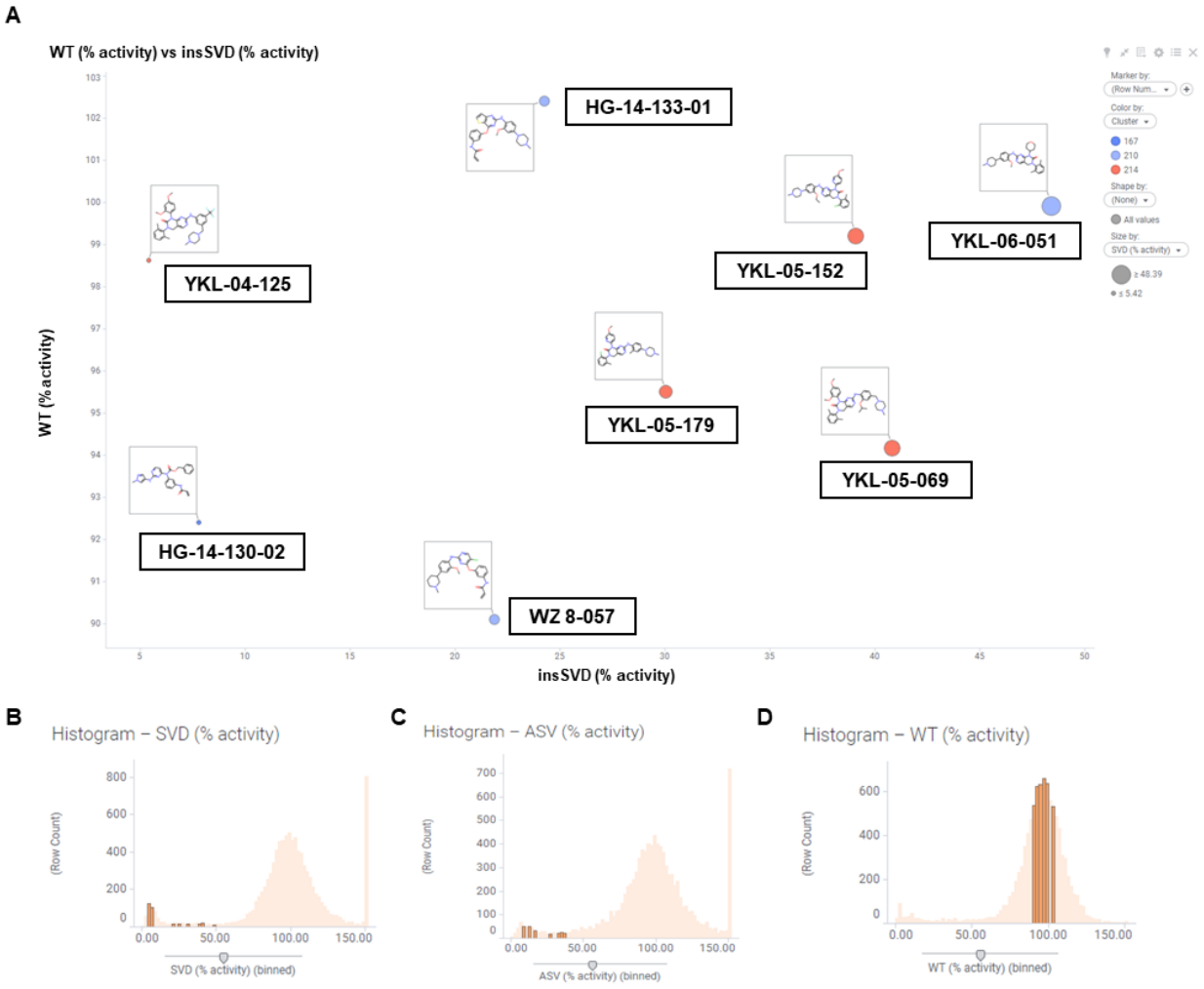


Figure 5.4. Cheminformatics analysis of selected hits. A, Zoomed-in selectivity plot showing selected hits with names and chemical structures. The Y-axis shows the percent activity of WT, and the X-axis shows the percent activity of insSVD. **B-D,** Histogram showing the distribution of compounds inhibiting insSVD, insASV, or WT, respectively, with the Y-axis showing the number of compounds while the X-axis showing the percent activity. The selected hits are highlighted in bright orange.

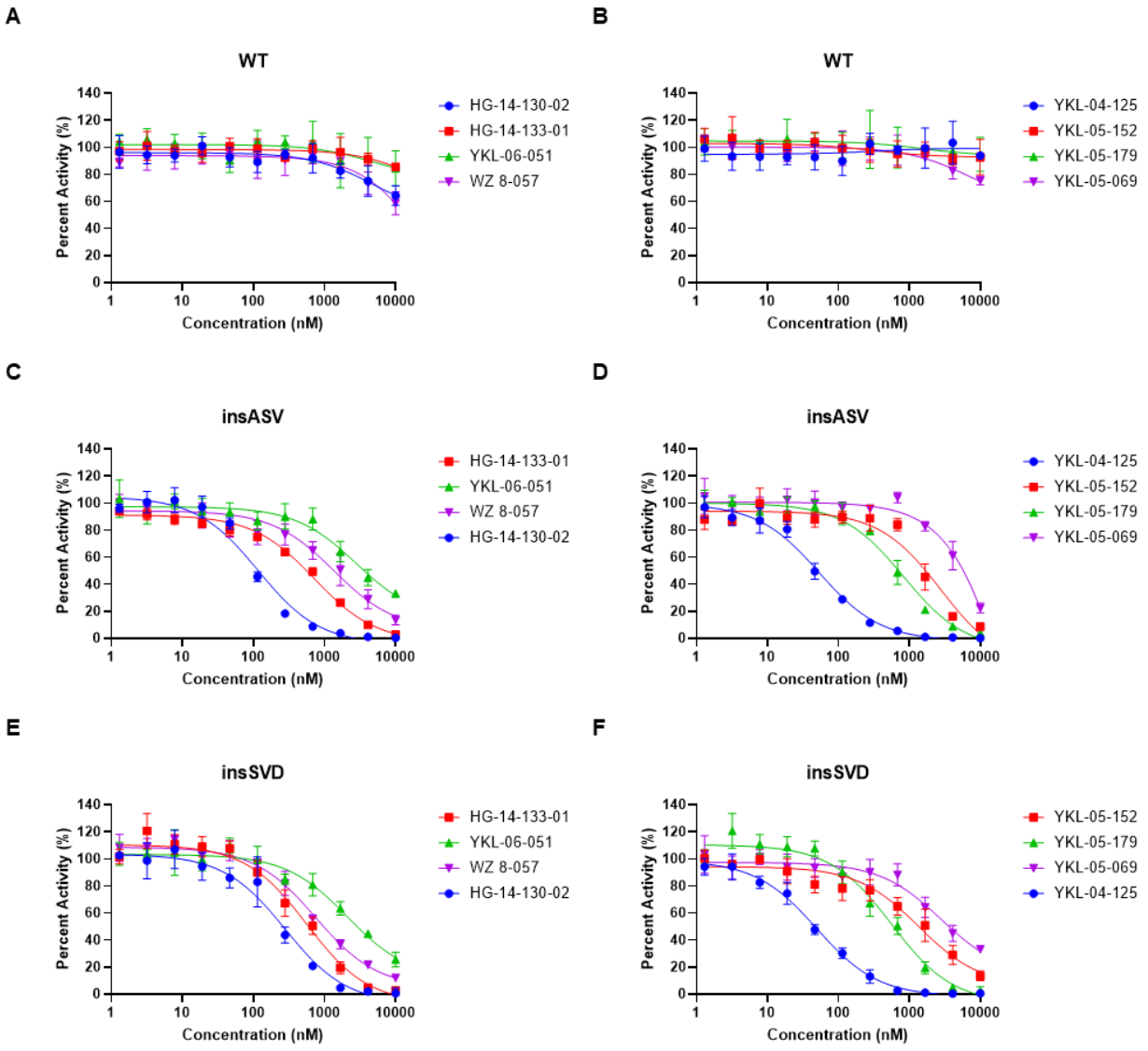


Figure 5.5. Follow-up IC_{50} determination of selected hits. A-B, Dose response curves showing selected hits against WT EGFR, with the Y-axis showing percent activity, which is normalized based on the average of signals without inhibitor presence, and the X-axis showing inhibitor concentrations. **C-D,** Dose response curves showing selected hits against insASV. **E-F,** Dose response curves showing selected hits against insSVD.

Table 5.1

Biochemical potencies of selected hits

	IC_{50} (nM)		
	WT	**insASV**	**insSVD**
HG-14-130-02	>10000	116	280
HG-14-133-01	>10000	719	563
WZ 8-057	>10000	1359	1021
YKL-04-125	>10000	53	50
YKL-05-069	>10000	5467	3961
YKL-05-152	>10000	1560	1359
YKL-05-179	>10000	830	590
YKL-06-051	>10000	2825	2989

5.4: Discussion

Potent inhibitors are not necessarily the most effective treatment for patients, with many examples of failed clinical trials like the ones with poziotinib or BDTX-189[206,208]. For these potent EGFR inhibitors, they not only kill cancer cells by inhibiting the functions of mutant proteins, but they also kill normal cells with strong potency against wild-type proteins. Therefore, there is not a wide therapeutic window to treat patients using these potent inhibitors, which ineitably causes severe side effects, such as skin rashes and diarrhera, and frequent treatment-related dose interruptions or discontinuation, limiting their clinical efficacies. As a result, mutant-selective inhibitors are in huge demand; a great example is that osimertinib treating NSCLC patients with activating EGFR mutations has become the standard of care[157,158,172]. However, the acquired resistance mutation C797S often occurs after extensive osimertinib treatment, which prevents osimertinib from covalently binding to EGFR kinase[186]. In fact, this C797S secondary mutation is not unique for osimertinib, which is observed in many 2nd and 3rd generation TKIs as well as some covalent exon 20 insertion inhibitors[166,171,198,202]. Therefore, reversible inhibitors with great mutant-selectivity have been an urgent unmet need.

To find such scaffolds with great mutant selectivity, I performed the large scale biochemical screen using purified EGFR kinases, and Jianwei Che helped me select eight hits using cheminformatics analysis. The follow-up IC_{50} determination showed that the most potent and mutant-selective inhibitor from these hits is YKL-04-125, which is fortunately a reversible compound. Further evaluation of off-target inhibition and cellular potency of this promising candidate is needed, but this scaffold could provide valuable insights on specific interactions required for mutant selectivity. For example, YKL-05-069, an isopropyl ether-substituted analog

of YKL-04-125, is 100-fold less mutant selective than YKL-04-125, highlighting the importance of the trifluoromethyl group on the phenyl ring. Co-crystal structures of these two hits in complex with insASV or insSVD kinase would greatly benefit our understanding of the mechanism of mutant selectivity, but since these two proteins are very hard to crystallize, I would focus on obtaining co-crystal structures of these two hits in complex with insNPG or WT EGFR in the near term. Based on the co-crystal structures, further modifications of the lead compound, YKL-04-125, should be attempted using medicinal chemistry methods to improve its mutant selectivity and drug properties.

5.5: Contributions

I expressed and purified all EGFR proteins, performed the inhibitor screening assays, and did follow-up IC_{50} determination of lead compounds. Dr. Jianwei Che did cheminformatics analysis of screening results. Prof. Michael J. Eck supervised the research.

Chapter 6

Conclusions and future directions

6.1: Conclusions

In my thesis, I have focused on and described results of a systematic biochemical and structural analysis of the two most frequent exon 20 insertions, EGFR insASV and insSVD, since they are responsible for about 40% of patients harboring insertions in exon 20[107,197]. With approximately 95% of exon 20 insertions occurring in the same region in the loop following C-helix[29], insASV and insSVD are good representatives to study, and clinical data indicate that they are both not sensitive to prior generations of TKIs[146,197]. As the mechanisms of oncogenic activation and TKI insensitivity are still poorly understood in EGFR exon 20 insertion mutants, we kinetically characterized EGFR insASV and insSVD kinases and assessed their sensitivity to a broad panel of EGFR inhibitors in biochemical and cellular assays. We utilized a series of crystal structures of EGFR in complex with exon 20 insertion targeting inhibitors to identify features that may contribute to potency and mutant selectivity. The systematic nature of our approach and the kinetic and structural insights we obtained allow us to propose potential directions for developing the next generation inhibitors selective against EGFR exon 20 insertion variants.

Extensive trials of optimizing expression and purification of EGFR insASV and insSVD kinases revealed that an inclusion of 20% glycerol in the purification buffer could maintain insASV or insSVD in a monomeric state and thus prevent aggregation. Kinetic characterization of insASV and insSVD showed that these insertion mutants have $K_{m,\ ATP}$ values much lower than the classical EGFR mutants and at a level similar to the drug-resistant L858R/T790M mutant, which partially explains why EGFR exon 20 insertion mutants are generally resistant to ATP competitive inhibitors. Examination of biochemical inhibitor sensitivities of a structurally diverse selection of EGFR inhibitors revealed that the 1st-4th generation EGFR TKIs are either not potent against exon

20 insertion mutants or not mutant selective, and that only four EGFR inhibitors, BAY-33, BAY-568, TAS6417, and TAK-788, displayed strong potency and mutant selectivity biochemically. Cellular studies of EGFR exon 20 insertion inhibitors confirmed my biochemical findings except BAY-33, possibly due to stability or cell permeability issues. Further studies on kinetics of inhibition indicated that the initial reversible binding is not a significant contributor of mutant selectivity, but the subsequent covalent bond formation, or the kinase conformational change, in the case of BAY-33, determines potency differences between WT EGFR and insertion mutants. Co-crystal structures of EGFR kinase in complex with exon 20 insertion inhibitors suggested that the conserved F723 in the P-loop and the T854 preceding the DFG motif could play important roles in selectively targeting exon 20 insertion mutants since they are associated with the active conformation of EGFR kinase, which is predominantly adopted by the insertion mutants while WT EGFR is in the inactive conformation by default.

In parallel with the above studies, I also performed a large-scale screening of approximately six thousand compounds against insASV, insSVD, and WT EGFR, using the focused kinase inhibitor library from the Gray lab at Stanford University. Facilitated by cheminformatics analysis, a total of eight hits were identified, five of which are reversible inhibitors. Further follow-up biochemical IC_{50} determination showed that YKL-04-125 and HG-14-130-02 are the most potent and mutant selective (~200-fold and ~50-fold, respectively). The most promising candidate, YKL-04-125, is a reversible inhibitor with >200-fold mutant selectivity, which deserves further evaluation of off-target inhibition (e.g., kinome profiling) and cellular potency. Overall, the large-scale screening showed that inhibitor sensitivities of EGFR insASV and insSVD are largely correlated, unlike their poor correlation with WT EGFR, which suggests that insASV and insSVD are not inherently more difficult to target than wild-type EGFR but instead favoring different scaffolds. Notably, the mutant-selective Bayer compounds were initially identified from a

large-scale biochemical screen similar to the one that I performed, further highlighting the importance of different scaffolds against exon 20 insertion mutants[218]. Whereas the inhibitor sensitivity of L858R is correlated with WT EGFR, this appears not to be the case with insASV or insSVD. They behave more like different kinases. This implies that different EGFR mutants may need different drugs; maybe we have not understood the precise inhibitor sensitivity mechanism.

6.2: Future directions

Our findings provide directions for future development of kinase inhibitors for EGFR exon 20 insertion variants. A co-crystal structure of either insASV or insSVD kinase in complex with a mutant-selective inhibitor like BAY-568 would provide additional structural evidence on how exon 20 insertion variants structurally differ from the wild-type EGFR, shedding light on mechanisms of oncogenic activation, as well as how inhibitors like BAY-568 achieve mutant selectivity. This could help validate our mechanism of where mutant selectivity comes from and give us insights on further inhibitor development for even more selective inhibitors, which could greatly benefit patients harboring these exon 20 insertion mutations. Progress in structural studies of these mutants has been limited so far largely due to tendency of aggregation of insASV or insSVD during concentrating efforts for crystallization. Further efforts could be focused on expressing and purifying the V948R variants of insASV and insSVD, which could not dimerize and thus might help prevent aggregation; a similar version of insNPG was successfully crystallized by Gonzalvez et al., which suggests the same for insASV and insSVD[201]. Alternatively, since insNPG displays a similar inhibition profile as insASV or insSVD and has the same insertion site with insSVD, we believe that co-crystal structures of insNPG with different inhibitors[128,201] coupled with our biochemical characterization could also help elucidate mechanisms of mutant selectivity and provide insight on future inhibitor development that would ultimately benefit patients harboring these oncogenic mutations.

However, there are still some other questions remain unanswered. First, how important are the roles of F723 and T854 in mutant selectivity? To illustrate the importance of these two amino acid residues, I would perform mutagenesis to mutate F723 to alanine and in parallel

mutating T854 to valine in insASV or insSVD and WT EGFR to see whether mutant-selective inhibitors, including BAY-33, BAY-568, TAS6417, and TAK-788, could still selectively target insertion mutants, if these newly generated mutants could properly fold and were catalytically active. Second, why does BAY-33 show great potency and mutant selectivity in our biochemical assay but is not potent in the cellular assay? To test if it is indeed due to stability or cell permeability issues, I would expose BAY-33 and BAY-568 to various chemical reagents or solvents followed by mass-spectrometry analysis, and also measure the cell permeability by LC-MS. Third, as I showed in Chapter 2, EGFR variants showed very different affinities for the peptide substrate poly(Glu_4,Tyr_1). How different are affinities for endogenous peptide substrates among EGFR variants and how would kinetic parameters change with these endogenous peptides? To test this, I would purify C-terminal tail of EGFR to substitute the use of poly(Glu_4,Tyr_1) in the coupled enzyme assay to compare the k_{cat} and K_m values.

Finally, what is the kinome specificity for the lead compound YKL-04-125 that was identified from my high-throughput screen? To examine this, I would perform kinome profiling and/or computational docking of this compound to different kinases. Interestingly, YKL-04-125 and YKL-05-069 are structurally similar but have 100-fold difference in mutant selectivity, so does the substitution of a trifluoromethyl group, or the position shift of substituent groups on the phenyl ring make the difference in potency and mutant selectivity between YKL-04-125 and YKL-05-069? To answer this question, I would make a trifluoromethyl-substituted analog of YKL-05-069 that has the substituent groups occupying the same positions on the phenyl ring to compare its potency and mutant selectivity with the two existing hits. It would be easier to understand the mechanism of mutant selectivity with co-crystal structures of these inhibitors in complex with EGFR kinase, preferably insASV or insSVD. Based on the differences in drug sensitivity between WT EGFR and insASV or insSVD, a broader inhibitor screen targeting EGFR exon 20 insertion

variants can potentially find more mutant-selective inhibitors. Nevertheless, the identification of YKL-04-125 as a promising candidate presents us a unique scaffold for further development of a reversible mutant-selective inhibitor with a wide therapeutic window.

References

1. Wee, P. & Wang, Z. Epidermal Growth Factor Receptor Cell Proliferation Signaling Pathways. *Cancers (Basel)* **9**, (2017).

2. Liu, Q. *et al.* EGFR-TKIs resistance via EGFR-independent signaling pathways. *Mol Cancer* **17**, (2018).

3. Ellis, L. M. Epidermal growth factor receptor in tumor angiogenesis. *Hematol Oncol Clin North Am* **18**, 1007–1021 (2004).

4. Sharma, S. v., Bell, D. W., Settleman, J. & Haber, D. A. Epidermal growth factor receptor mutations in lung cancer. *Nature Reviews Cancer 2007 7:3* **7**, 169–181 (2007).

5. Sung, H. *et al.* Global Cancer Statistics 2020: GLOBOCAN Estimates of Incidence and Mortality Worldwide for 36 Cancers in 185 Countries. *CA Cancer J Clin* **71**, 209–249 (2021).

6. Subramanian, J. & Govindan, R. Lung cancer in never smokers: A review. *Journal of Clinical Oncology* **25**, 561–570 (2007).

7. Molina, J. R. *et al.* Non-Small Cell Lung Cancer: Epidemiology, Risk Factors, Treatment, and Survivorship. *Mayo Clin Proc* **83**, 584–594 (2008).

8. Navada, S., Lai, P., Schwartz, A. G. & Kalemkerian, G. P. Temporal trends in small cell lung cancer: Analysis of the national Surveillance, Epidemiology, and End-Results (SEER) database. *https://doi.org/10.1200/jco.2006.24.18_suppl.7082* **24**, 7082–7082 (2006).

9. Siegel, R. L., Miller, K. D., Fuchs, H. E. & Jemal, A. Cancer Statistics, 2021. *CA Cancer J Clin* **71**, 7–33 (2021).

10. Lynch, T. J. *et al.* Activating Mutations in the Epidermal Growth Factor Receptor Underlying Responsiveness of Non–Small-Cell Lung Cancer to Gefitinib. *New England Journal of Medicine* **350**, 2129–2139 (2004).

11. Fukuoka, M. *et al.* Multi-institutional randomized phase II trial of gefitinib for previously treated patients with advanced non-small-cell lung cancer. *Journal of Clinical Oncology* **21**, 2237–2246 (2003).

12. Vivanco, I. Targeting molecular addictions in cancer. *British Journal of Cancer 2014 111:11* **111**, 2033–2038 (2014).

13. Arteaga, C. L. ErbB-targeted therapeutic approaches in human cancer. *Exp Cell Res* **284**, 122–130 (2003).

14. Cohen, M. H., Williams, G. A., Sridhara, R., Chen, G. & Pazdur, R. FDA Drug Approval Summary: Gefitinib (ZD1839) (Iressa®) Tablets. *Oncologist* **8**, 303–306 (2003).

15. Pao, W. *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* **101**, 13306 (2004).

16. Yun, C. H. *et al.* Structures of Lung Cancer-Derived EGFR Mutants and Inhibitor Complexes: Mechanism of Activation and Insights into Differential Inhibitor Sensitivity. *Cancer Cell* **11**, 217–227 (2007).

17. Carey, K. D. *et al.* Kinetic analysis of epidermal growth factor receptor somatic mutant proteins shows increased sensitivity to the epidermal growth factor receptor tyrosine kinase inhibitor, erlotinib. *Cancer Res* **66**, 8163–8171 (2006).

18. Paez, J. G. *et al.* EGFR mutations in lung, cancer: Correlation with clinical response to gefitinib therapy. *Science (1979)* **304**, 1497–1500 (2004).

19. Maemondo, M. *et al.* Gefitinib or Chemotherapy for Non–Small-Cell Lung Cancer with Mutated EGFR. *New England Journal of Medicine* **362**, 2380–2388 (2010).

20. Hirsch, F. R. & Bunn, P. A. EGFR testing in lung cancer is ready for prime time. *Lancet Oncol* **10**, 432–433 (2009).

21. Hirsch, F. R. *et al.* Epidermal growth factor receptor in non-small-cell lung carcinomas: correlation between gene copy number and protein expression and impact on prognosis. *J Clin Oncol* **21**, 3798–3807 (2003).

22. Shigematsu, H. *et al.* Clinical and Biological Features Associated With Epidermal Growth Factor Receptor Gene Mutations in Lung Cancers. *JNCI: Journal of the National Cancer Institute* **97**, 339–346 (2005).

23. Huang, L. T., Zhang, S. L., Han, C. B. & Ma, J. T. Impact of EGFR exon 19 deletion subtypes on clinical outcomes in EGFR-TKI-Treated advanced non-small-cell lung cancer. *Lung Cancer* **166**, 9–16 (2022).

24. Sordella, R., Bell, D. W., Haber, D. A. & Settleman, J. Gefitinib-sensitizing EGFR mutations in lung cancer activate anti-apoptotic pathways. *Science (1979)* **305**, 1163–1167 (2004).

25. Sordella, R., Bell, D. W., Haber, D. A. & Settleman, J. Gefitinib-sensitizing EGFR mutations in lung cancer activate anti-apoptotic pathways. *Science (1979)* **305**, 1163–1167 (2004).

26. Tumbrink, H. L., Heimsoeth, A. & Sos, M. L. The next tier of EGFR resistance mutations in lung cancer. *Oncogene 2020 40:1* **40**, 1–11 (2020).

27. Harrison, P. T., Vyse, S. & Huang, P. H. Rare epidermal growth factor receptor (EGFR) mutations in non-small cell lung cancer. *Semin Cancer Biol* **61**, 167 (2020).

28. Eck, M. J. & Yun, C. H. Structural and mechanistic underpinnings of the differential drug sensitivity of EGFR mutations in non-small cell lung cancer. *Biochimica et Biophysica Acta - Proteins and Proteomics* vol. 1804 559–566 Preprint at https://doi.org/10.1016/j.bbapap.2009.12.010 (2010).

29. Yasuda, H., Kobayashi, S. & Costa, D. B. EGFR exon 20 insertion mutations in non-small-cell lung cancer: Preclinical data and clinical implications. *Lancet Oncol* **13**, e23–e31 (2012).

30. Oxnard, G. R. *et al.* Natural history and molecular characteristics of lung cancers harboring EGFR exon 20 insertions. *J Thorac Oncol* **8**, 179 (2013).

31. Arcila, M. E. *et al.* EGFR Exon 20 Insertion Mutations in Lung Adenocarcinomas: Prevalence, Molecular Heterogeneity, and Clinicopathologic Characteristics. *Mol Cancer Ther* **12**, 220 (2013).

32. Burnett, H. *et al.* Epidemiological and clinical burden of EGFR Exon 20 insertion in advanced non-small cell lung cancer: A systematic literature review. *PLoS One* **16**, e0247620 (2021).

33. Huang, Y. *et al.* Molecular basis for multimerization in the activation of the epidermal growth factor receptor. *Elife* **5**, (2016).

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

35. Ullrich, A. & Schlessinger, J. Signal transduction by receptors with tyrosine kinase activity. *Cell* **61**, 203–212 (1990).

36. Robinson, D. R., Wu, Y. M. & Lin, S. F. The protein tyrosine kinase family of the human genome. *Oncogene 2000 19:49* **19**, 5548–5557 (2000).

37. Carpenter, G., King, L. & Cohen, S. Epidermal growth factor stimulates phosphorylation in membrane preparations in vitro. *Nature 1978 276:5686* **276**, 409–410 (1978).

38. Cohen, S., Ushiro, H., Stoscheck, C. & Chinkers, M. A native 170,000 epidermal growth factor receptor-kinase complex from shed plasma membrane vesicles. *Journal of Biological Chemistry* **257**, 1523–1531 (1982).

39. Yewale, C., Baradia, D., Vhora, I., Patil, S. & Misra, A. Epidermal growth factor receptor targeting in cancer: A review of trends and strategies. *Biomaterials* **34**, 8690–8707 (2013).

40. Yarden, Y. & Sliwkowski, M. X. Untangling the ErbB signalling network. *Nature Reviews Molecular Cell Biology 2001 2:2* **2**, 127–137 (2001).

41. Jones, S. & Rappoport, J. Z. Interdependent epidermal growth factor receptor signalling and trafficking. *Int J Biochem Cell Biol* **51**, 23–28 (2014).

42. Holbro, T., Civenni, G. & Hynes, N. E. The ErbB receptors and their role in cancer progression. *Exp Cell Res* **284**, 99–110 (2003).

43. Bublil, E. M. & Yarden, Y. The EGF receptor family: spearheading a merger of signaling and therapeutics. *Curr Opin Cell Biol* **19**, 124–134 (2007).

44. Yarden, Y. The EGFR family and its ligands in human cancer: Signalling mechanisms and therapeutic opportunities. *Eur J Cancer* **37**, 3–8 (2001).

45. Soltoff, S. P., Carraway Iii, K. L., Prigent, S. A., Gullick, W. G. & Cantley', L. C. ErbB3 is involved in activation of phosphatidylinositol 3-kinase by epidermal growth factor. *Mol Cell Biol* **14**, 3550–3558 (1994).

46. Burgess, A. W. *et al.* An Open-and-Shut Case? Recent Insights into the Activation of EGF/ErbB Receptors. *Mol Cell* **12**, 541–552 (2003).

47. Cho, H. S. *et al.* Structure of the extracellular region of HER2 alone and in complex with the Herceptin Fab. *Nature 2003 421:6924* **421**, 756–760 (2003).

48. Shi, F., Telesco, S. E., Liu, Y., Radhakrishnan, R. & Lemmona, M. A. ErbB3/HER3 intracellular domain is competent to bind ATP and catalyze autophosphorylation. *Proc Natl Acad Sci U S A* **107**, 7692–7697 (2010).

49. Citri, A., Skaria, K. B. & Yarden, Y. The deaf and the dumb: the biology of ErbB-2 and ErbB-3. *Exp Cell Res* **284**, 54–65 (2003).

50. Boudeau, J., Miranda-Saavedra, D., Barton, G. J. & Alessi, D. R. Emerging roles of pseudokinases. *Trends Cell Biol* **16**, 443–452 (2006).

51. Uribe, M. L., Marrocco, I. & Yarden, Y. EGFR in Cancer: Signaling Mechanisms, Drugs, and Acquired Resistance. *Cancers (Basel)* **13**, (2021).

52. Kovacs, E., Zorn, J. A., Huang, Y., Barros, T. & Kuriyan, J. A Structural Perspective on the Regulation of the EGF Receptor. *Annu Rev Biochem* **84**, 739 (2015).

53. Lemmon, M. A., Schlessinger, J. & Ferguson, K. M. The EGFR family: Not so prototypical receptor tyrosine kinases. *Cold Spring Harb Perspect Biol* **6**, (2014).

54. Normanno, N. *et al.* Epidermal growth factor receptor (EGFR) signaling in cancer. *Gene* **366**, 2–16 (2006).

55. Ullrich, A. *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* **309**, 418–425 (1984).

56. Lax, I. *et al.* Domain deletion in the extracellular portion of the EGF-receptor reduces ligand binding and impairs cell surface expression. *Cell Regul* **1**, 173 (1990).

57. Grasso, A. W. *et al.* ErbB kinases and NDF signaling in human prostate cancer cells. *Oncogene 1997 15:22* **15**, 2705–2716 (1997).

58. Arteaga, C. L. & Engelman, J. A. ERBB receptors: from oncogene discovery to basic science to mechanism-based cancer therapeutics. *Cancer Cell* **25**, 282–303 (2014).

59. Schneider, M. R. & Wolf, E. The epidermal growth factor receptor ligands at a glance. *J Cell Physiol* **218**, 460–466 (2009).

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

61. Carraway, K. L. & Cantley, L. C. A neu acquaintance for ErbB3 and ErbB4: A role for receptor heterodimerization in growth signaling. *Cell* **78**, 5–8 (1994).

62. Freed, D. M. *et al.* EGFR ligands differentially stabilize receptor dimers to specify signaling kinetics. *Cell* **171**, 683 (2017).

63. Sinclair, J. K. L., Walker, A. S., Doerner, A. E. & Schepartz, A. Mechanism of allosteric coupling into and through the plasma membrane by EGFR. *Cell Chem Biol* **25**, 857 (2018).

64. Lu, C. *et al.* Structural Evidence for Loose Linkage between Ligand Binding and Kinase Activation in the Epidermal Growth Factor Receptor. *Mol Cell Biol* **30**, 5432 (2010).

65. Scheck, R. A., Lowder, M. A., Appelbaum, J. S. & Schepartz, A. Bipartite Tetracysteine Display Reveals Allosteric Control of Ligand-Specific EGFR Activation. *ACS Chem Biol* **7**, 1367 (2012).

66. Wilson, K. J., Gilmore, J. L., Foley, J., Lemmon, M. A. & Riese, D. J. Functional Selectivity of EGF Family Peptide Growth Factors: Implications for Cancer. *Pharmacol Ther* **122**, 1 (2009).

67. Huang, Y. *et al.* A molecular mechanism for the generation of ligand-dependent differential outputs by the epidermal growth factor receptor. *Elife* **10**, (2021).

68. Sinclair, J. K. L., Walker, A. S., Doerner, A. E. & Schepartz, A. Mechanism of Allosteric Coupling into and through the Plasma Membrane by EGFR. *Cell Chem Biol* **25**, 857-870.e7 (2018).

69. Lu, C. *et al.* Structural Evidence for Loose Linkage between Ligand Binding and Kinase Activation in the Epidermal Growth Factor Receptor. *Mol Cell Biol* **30**, 5432–5443 (2010).

70. Zhang, X., Gureasko, J., Shen, K., Cole, P. A. & Kuriyan, J. An Allosteric Mechanism for Activation of the Kinase Domain of Epidermal Growth Factor Receptor. *Cell* **125**, 1137–1149 (2006).

71. Gotoh, N., Tojo, A., Hino, M., Yazaki, Y. & Shibuya, M. A highly conserved tyrosine residue at codon 845 within the kinase domain is not required for the transforming activity of human epidermal growth factor receptor. *Biochem Biophys Res Commun* **186**, 768–774 (1992).

72. Hubbard, S. R. The juxtamembrane region of EGFR takes center stage. *Cell* **137**, 1181–1183 (2009).

73. Maeda, R., Sato, T., Okamoto, K., Yanagawa, M. & Sako, Y. Lipid-Protein Interplay in Dimerization of Juxtamembrane Domains of Epidermal Growth Factor Receptor. *Biophys J* **114**, 893–903 (2018).

74. Jura, N. *et al.* Mechanism for activation of the EGF receptor catalytic domain by the juxtamembrane segment. *Cell* **137**, 1293–1307 (2009).

75. Arkhipov, A. *et al.* Architecture and Membrane Interactions of the EGF Receptor. *Cell* **152**, 557–569 (2013).

76. Endres, N. F. *et al.* Conformational Coupling across the Plasma Membrane in Activation of the EGF Receptor. *Cell* **152**, 543–556 (2013).

77. Downward, J., Parker, P. & Waterfield, M. D. Autophosphorylation sites on the epidermal growth factor receptor. *Nature* **311**, 483–485 (1984).

78. Schulze, W. X., Deng, L. & Mann, M. Phosphotyrosine interactome of the ErbB-receptor kinase family. *Mol Syst Biol* **1**, (2005).

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

80. Grøvdal, L. M., Stang, E., Sorkin, A. & Madshus, I. H. Direct interaction of Cbl with pTyr 1045 of the EGF receptor (EGFR) is required to sort the EGFR to lysosomes for degradation. *Exp Cell Res* **300**, 388–395 (2004).

81. Oda, K., Matsuoka, Y., Funahashi, A. & Kitano, H. A comprehensive pathway map of epidermal growth factor receptor signaling. *Mol Syst Biol* **1**, 2005.0010 (2005).

82. Downward, J. Targeting RAS signalling pathways in cancer therapy. *Nature Reviews Cancer 2003 3:1* **3**, 11–22 (2003).

83. Yoshizawa, R. *et al.* p52Shc regulates the sustainability of ERK activation in a RAF-independent manner. *Mol Biol Cell* **32**, 1838–1848 (2021).

84. Park, E. *et al.* Architecture of autoinhibited and active BRAF–MEK1–14-3-3 complexes. *Nature 2019 575:7783* **575**, 545–550 (2019).

85. Rodrigues, G. A., Falasca, M., Zhang, Z., Ong, S. H. & Schlessinger, J. A Novel Positive Feedback Loop Mediated by the Docking Protein Gab1 and Phosphatidylinositol 3-Kinase in Epidermal Growth Factor Receptor Signaling. *Mol Cell Biol* **20**, 1448 (2000).

86. Cully, M., You, H., Levine, A. J. & Mak, T. W. Beyond PTEN mutations: the PI3K pathway as an integrator of multiple inputs during tumorigenesis. *Nature Reviews Cancer 2006 6:3* **6**, 184–192 (2006).

87. Mendoza, M. C., Er, E. E. & Blenis, J. The Ras-ERK and PI3K-mTOR Pathways: Cross-talk and Compensation. *Trends Biochem Sci* **36**, 320 (2011).

88. Tia, N. *et al.* Role of Forkhead Box O (FOXO) transcription factor in aging and diseases. *Gene* **648**, 97–105 (2018).

89. de Larco, J. E. & Todaro, G. J. Epithelioid and fibroblastic rat kidney cell clones: Epidermal growth factor (EGF) receptors and the effect of mouse sarcoma virus transformation. *J Cell Physiol* **94**, 335–342 (1978).

90. Downward, J. *et al.* Close similarity of epidermal growth factor receptor and v-erb-B oncogene protein sequences. *Nature 1984 307:5951* **307**, 521–527 (1984).

91. Johnson, L. N., Lowe, E. D., Noble, M. E. M. & Owen, D. J. The structural basis for substrate recognition and control by protein kinases. *FEBS Lett* **430**, 1–11 (1998).

92. Taylor, S. S. & Kornev, A. P. Protein kinases: evolution of dynamic regulatory proteins. *Trends Biochem Sci* **36**, 65–77 (2011).

93. Modi, V. & Dunbrack, R. L. Defining a new nomenclature for the structures of active and inactive kinases. *Proc Natl Acad Sci U S A* **116**, 6818–6827 (2019).

94. Huse, M. & Kuriyan, J. The conformational plasticity of protein kinases. *Cell* **109**, 275–282 (2002).

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

96. Kanev, G. K. *et al.* The Landscape of Atypical and Eukaryotic Protein Kinases. *Trends Pharmacol Sci* **40**, 818–832 (2019).

97. Park, J. H., Liu, Y., Lemmon, M. A. & Radhakrishnan, R. Erlotinib binds both inactive and active conformations of the EGFR tyrosine kinase domain. *Biochem J* **448**, 417–423 (2012).

98. Shan, Y., Arkhipov, A., Kim, E. T., Pan, A. C. & Shawa, D. E. Transitions to catalytically inactive conformations in EGFR kinase. *Proc Natl Acad Sci U S A* **110**, 7270–7275 (2013).

99. Vijayan, R. S. K. *et al.* Conformational analysis of the DFG-out kinase motif and biochemical profiling of structurally validated type II inhibitors. *J Med Chem* **58**, 466–479 (2015).

100. Peng, Y. H. *et al.* Protein kinase inhibitor design by targeting the Asp-Phe-Gly (DFG) motif: the role of the DFG motif in the design of epidermal growth factor receptor inhibitors. *J Med Chem* **56**, 3889–3903 (2013).

101. Hasenahuer, M. A., Barletta, G. P., Fernandez-Alberti, S., Parisi, G. & Fornasari, M. S. Pockets as structural descriptors of EGFR kinase conformations. *PLoS One* **12**, (2017).

102. Wood, E. R. *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* **64**, 6652–6659 (2004).

103. Stamos, J., Sliwkowski, M. X. & Eigenbrot, C. Structure of the epidermal growth factor receptor kinase domain alone and in complex with a 4-anilinoquinazoline inhibitor. *Journal of Biological Chemistry* **277**, 46265–46272 (2002).

104. Jeffrey, P. D. *et al.* Mechanism of CDK activation revealed by the structure of a cyclinA-CDK2 complex. *Nature 1995 376:6538* **376**, 313–320 (1995).

105. Kumar, V., Abbas, A. & Aster, J. Robbins Basic Pathology E-Book. *Elsevier Health Sciences* 133–142 (2017).

106. Jänne, P. A. *et al.* Effect of Epidermal Growth Factor Receptor Tyrosine Kinase Domain Mutations on the Outcome of Patients with Non–Small Cell Lung Cancer Treated with Epidermal Growth Factor Receptor Tyrosine Kinase Inhibitors. *Clinical Cancer Research* **12**, 4416s–4420s (2006).

107. Kosaka, T. *et al.* Response heterogeneity of EGFR and HER2 exon 20 insertions to covalent EGFR and HER2 inhibitors. *Cancer Res* **77**, 2712–2721 (2017).

108. Holbro, T. & Hynes, N. E. ErbB receptors: directing key signaling networks throughout life. *Annu Rev Pharmacol Toxicol* **44**, 195–217 (2004).

109. Hynes, N. E. & Lane, H. A. ERBB receptors and cancer: the complexity of targeted inhibitors. *Nature Reviews Cancer 2005 5:5* **5**, 341–354 (2005).

110. Chan, B. A. & Hughes, B. G. M. Targeted therapy for non-small cell lung cancer: current standards and the promise of the future. *Transl Lung Cancer Res* **4**, 36 (2015).

111. Xu, M. J., Johnson, D. E. & Grandis, J. R. EGFR-targeted therapies in the post-genomic era. *Cancer Metastasis Rev* **36**, 463–473 (2017).

112. Mendelsohn, J. Antibody-mediated EGF receptor blockade as an anticancer therapy: from the laboratory to the clinic. *Cancer Immunology, Immunotherapy 2003 52:5* **52**, 342–346 (2003).

113. Cai, W. Q. *et al.* The Latest Battles Between EGFR Monoclonal Antibodies and Resistant Tumor Cells. *Front Oncol* **10**, 1249 (2020).

114. Martinelli, E., de Palma, R., Orditura, M., de Vita, F. & Ciardiello, F. Anti-epidermal growth factor receptor monoclonal antibodies in cancer therapy. *Clin Exp Immunol* **158**, 1 (2009).

115. Cho, J. *et al.* Cetuximab response of lung cancer-derived EGF receptor mutants is associated with asymmetric dimerization. *Cancer Res* **73**, 6770–6779 (2013).

116. Yuan, M., Huang, L. L., Chen, J. H., Wu, J. & Xu, Q. The emerging treatment landscape of targeted therapy in non-small-cell lung cancer. *Signal Transduction and Targeted Therapy 2019 4:1* **4**, 1–14 (2019).

117. Majeed, U., Manochakian, R., Zhao, Y. & Lou, Y. Targeted therapy in advanced non-small cell lung cancer: current advances and future trends. *Journal of Hematology & Oncology 2021 14:1* **14**, 1–20 (2021).

118. Tian, X., Gu, T., Lee, M. H. & Dong, Z. Challenge and countermeasures for EGFR targeted therapy in non-small cell lung cancer. *Biochim Biophys Acta Rev Cancer* **1877**, (2022).

119. Abourehab, M. A. S., Alqahtani, A. M., Youssif, B. G. M. & Gouda, A. M. Globally Approved EGFR Inhibitors: Insights into Their Syntheses, Target Kinases, Biological Activities, Receptor Interactions, and Metabolism. *Molecules* **26**, (2021).

120. Cohen, M. H., Johnson, J. R., Chen, Y.-F., Sridhara, R. & Pazdur, R. FDA drug approval summary: erlotinib (Tarceva) tablets. *Oncologist* **10**, 461–466 (2005).

121. Kobayashi, S. *et al.* EGFR mutation and resistance of non-small-cell lung cancer to gefitinib. *N Engl J Med* **352**, 786–792 (2005).

122. Johnson, J. R. *et al.* Approval Summary for Erlotinib for Treatment of Patients with Locally Advanced or Metastatic Non–Small Cell Lung Cancer after Failure of at Least One Prior Chemotherapy Regimen. *Clinical Cancer Research* **11**, 6414–6421 (2005).

123. Ryan, Q. *et al.* FDA drug approval summary: lapatinib in combination with capecitabine for previously treated metastatic breast cancer that overexpresses HER-2. *Oncologist* **13**, 1114–1119 (2008).

124. Gilmer, T. M. *et al.* Impact of Common Epidermal Growth Factor Receptor and HER2 Variants on Receptor Activity and Inhibition by Lapatinib. *Cancer Res* **68**, 571–579 (2008).

125. Johnston, S. R. D. & Leary, A. Lapatinib: A novel EGFR/HER2 tyrosine kinase inhibitor for cancer. *Drugs of Today* **42**, 441–453 (2006).

126. Roskoski, R. Classification of small molecule protein kinase inhibitors based upon the structures of their drug-enzyme complexes. *Pharmacol Res* **103**, 26–48 (2016).

127. Yun, C. H. *et al.* The T790M mutation in EGFR kinase causes drug resistance by increasing the affinity for ATP. *Proc Natl Acad Sci U S A* **105**, 2070–2075 (2008).

128. Yasuda, H. *et al.* Structural, biochemical and clinical characterization of epidermal growth factor receptor (EGFR) exon 20 insertion mutations in lung cancer. *Sci Transl Med* **5**, 216ra177 (2013).

129. Kosaka, T. *et al.* Analysis of epidermal growth factor receptor gene mutation in patients with non-small cell lung cancer and acquired resistance to gefitinib. *Clin Cancer Res* **12**, 5764–5769 (2006).

130. Balak, M. N. *et al.* Novel D761Y and common secondary T790M mutations in epidermal growth factor receptor-mutant lung adenocarcinomas with acquired resistance to kinase inhibitors. *Clin Cancer Res* **12**, 6494–6501 (2006).

131. Yu, H. A. *et al.* Analysis of tumor specimens at the time of acquired resistance to EGFR-TKI therapy in 155 patients with EGFR-mutant lung cancers. *Clin Cancer Res* **19**, 2240–2247 (2013).

132. Pao, W. *et al.* Acquired Resistance of Lung Adenocarcinomas to Gefitinib or Erlotinib Is Associated with a Second Mutation in the EGFR Kinase Domain. *PLoS Med* **2**, e73 (2005).

133. Pao, W. & Chmielecki, J. Rational, biologically based treatment of EGFR-mutant non-small-cell lung cancer. *Nat Rev Cancer* **10**, 760 (2010).

134. Takezawa, K. *et al.* HER2 amplification: a potential mechanism of acquired resistance to EGFR inhibition in EGFR-mutant lung cancers that lack the second-site EGFRT790M mutation. *Cancer Discov* **2**, 922–933 (2012).

135. Sequist, L. v. *et al.* Genotypic and histological evolution of lung cancers acquiring resistance to EGFR inhibitors. *Sci Transl Med* **3**, (2011).

136. Engelman, J. A. *et al.* MET amplification leads to gefitinib resistance in lung cancer by activating ERBB3 signaling. *Science (1979)* **316**, 1039–1043 (2007).

137. Dungo, R. T. & Keating, G. M. Afatinib: First Global Approval. *Drugs 2013 73:13* **73**, 1503–1515 (2013).

138. Shirley, M. Dacomitinib: First Global Approval. *Drugs* **78**, 1947–1953 (2018).

139. Kwak, E. L. *et al.* Irreversible inhibitors of the EGF receptor may circumvent acquired resistance to gefitinib. *Proc Natl Acad Sci U S A* **102**, 7665–7670 (2005).

140. Stasi, I. & Cappuzzo, F. Second generation tyrosine kinase inhibitors for the treatment of metastatic non-small-cell lung cancer. *Transl Respir Med* **2**, (2014).

141. Wu, Y. L. *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* **18**, 1454–1466 (2017).

142. Song, Z. *et al.* Efficacy and safety of pyrotinib in advanced lung adenocarcinoma with HER2 mutations: a multicenter, single-arm, phase II trial. *BMC Med* **20**, 1–10 (2022).

143. Blair, H. A. Pyrotinib: First Global Approval. *Drugs* **78**, 1751–1755 (2018).

144. Singh, H. *et al.* U.S. Food and drug administration approval: Neratinib for the extended adjuvant treatment of early-stage HER2-positive breast cancer. *Clinical Cancer Research* **24**, 3486–3491 (2018).

145. Deeks, E. D. Neratinib: First Global Approval. *Drugs* **77**, 1695–1704 (2017).

146. Yang, J. C. H. *et al.* Clinical activity of afatinib in patients with advanced non-small-cell lung cancer harbouring uncommon EGFR mutations: A combined post-hoc analysis of LUX-Lung 2, LUX-Lung 3, and LUX-Lung 6. *Lancet Oncol* **16**, 830–838 (2015).

147. Yap, T. A. *et al.* Phase I trial of the irreversible EGFR and HER2 kinase inhibitor BIBW 2992 in patients with advanced solid tumors. *Journal of Clinical Oncology* **28**, 3965–3972 (2010).

148. Dziadziuszko, R. *et al.* Afatinib in NSCLC With HER2 Mutations: Results of the Prospective, Open-Label Phase II NICHE Trial of European Thoracic Oncology Platform (ETOP). *Journal of Thoracic Oncology* **14**, 1086–1094 (2019).

149. Fan, Y. *et al.* Afatinib in patients with advanced non-small cell lung cancer harboring HER2 mutations, previously treated with chemotherapy: A phase II trial. *Lung Cancer* **147**, 209–213 (2020).

150. Song, Z. *et al.* Efficacy and Resistance of Afatinib in Chinese Non-Small Cell Lung Cancer Patients With HER2 Alterations: A Multicenter Retrospective Study. *Front Oncol* **11**, 1403 (2021).

151. Katzung, B. G. *et al.* Basic & Clinical Pharmacology Fourteenth Edition a LANGE medical book. (2018).

152. Zhou, W. *et al.* Novel mutant-selective EGFR kinase inhibitors against EGFR T790M. *Nature* **462**, 1070–1074 (2009).

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

154. Pawara, R. *et al.* Design and synthesis of the novel, selective WZ4002 analogue as EGFR-L858R/T790M tyrosine kinase inhibitors for targeted drug therapy in non-small-cell lung cancer (NSCLC). *J Mol Struct* **1254**, 132313 (2022).

155. Finlay, M. R. v. *et al.* Discovery of a potent and selective EGFR inhibitor (AZD9291) of both sensitizing and T790M resistance mutations that spares the wild type form of the receptor. *J Med Chem* **57**, 8249–8267 (2014).

156. Greig, S. L. Osimertinib: First Global Approval. *Drugs 2015 76:2* **76**, 263–273 (2016).

157. Ramalingam, S. S. *et al.* Osimertinib As First-Line Treatment of EGFR Mutation-Positive Advanced Non-Small-Cell Lung Cancer. *J Clin Oncol* **36**, 841–849 (2018).

158. Koch, A. L. *et al.* FDA Approval Summary: Osimertinib for Adjuvant Treatment of Surgically Resected Non-Small Cell Lung Cancer, a Collaborative Project Orbis Review. *Clin Cancer Res* **27**, 6638–6643 (2021).

159. Soria, J.-C. *et al.* Osimertinib in Untreated EGFR-Mutated Advanced Non-Small-Cell Lung Cancer. *N Engl J Med* **378**, 113–125 (2018).

160. Colclough, N. *et al.* Preclinical comparison of the blood-brain barrier permeability of osimertinib with other EGFR TKIs. *Clinical Cancer Research* **27**, 189–201 (2021).

161. Zhao, Y. *et al.* Overall survival benefit of osimertinib and clinical value of upfront cranial local therapy in untreated EGFR-mutant nonsmall cell lung cancer with brain metastasis. *Int J Cancer* **150**, 1318–1328 (2022).

162. Barnes, T. A., O'Kane, G. M., Vincent, M. D. & Leighl, N. B. Third-Generation Tyrosine Kinase Inhibitors Targeting Epidermal Growth Factor Receptor Mutations in Non-Small Cell Lung Cancer. *Front Oncol* **7**, 113 (2017).

163. Wright, N. M. A. & Goss, G. D. Third-generation epidermal growth factor receptor tyrosine kinase inhibitors for the treatment of non-small cell lung cancer. *Transl Lung Cancer Res* **8**, S247–S264 (2019).

164. Thress, K. S. *et al.* Acquired EGFR C797S mutation mediates resistance to AZD9291 in non–small cell lung cancer harboring EGFR T790M. *Nature Medicine 2015 21:6* **21**, 560–562 (2015).

165. Niederst, M. J. *et al.* The allelic context of the C797S mutation acquired upon treatment with third-generation EGFR inhibitors impacts sensitivity to subsequent treatment strategies. *Clinical Cancer Research* **21**, 3924–3933 (2015).

166. Ercan, D. *et al.* EGFR mutations and resistance to irreversible pyrimidine-based EGFR inhibitors. *Clinical Cancer Research* **21**, 3913–3923 (2015).

167. Song, H. N. *et al.* Acquired C797S mutation upon treatment with a T790M-specific third-generation EGFR inhibitor (HM61713) in non-small cell lung cancer. *Journal of Thoracic Oncology* **11**, e45–e47 (2016).

168. Janne, P. A. *et al.* Incidence and heterogeneity of C797S and other EGFR resistance mutations on routine comprehensive genomic profiling (CGP). *Journal of Clinical Oncology* **39**, 9101–9101 (2021).

169. Wang, S., Tsui, S. T., Liu, C., Song, Y. & Liu, D. EGFR C797S mutation mediates resistance to third-generation inhibitors in T790M-positive non-small cell lung cancer. *J Hematol Oncol* **9**, (2016).

170. Tripathi, S. K. & Biswal, B. K. Allosteric mutant-selective fourth-generation EGFR inhibitors as an efficient combination therapeutic in the treatment of non-small cell lung carcinoma. *Drug Discov Today* **26**, 1466–1472 (2021).

171. Du, X. *et al.* Acquired resistance to third-generation EGFR-TKIs and emerging next-generation EGFR inhibitors. *The Innovation* **2**, 100103 (2021).

172. Piper-Vallillo, A. J., Sequist, L. v. & Piotrowska, Z. Emerging treatment paradigms for EGFR-mutant lung cancers progressing on osimertinib: A review. *Journal of Clinical Oncology* **38**, 2926–2936 (2020).

173. Tian, X., Gu, T., Lee, M. H. & Dong, Z. Challenge and countermeasures for EGFR targeted therapy in non-small cell lung cancer. *Biochimica et Biophysica Acta (BBA) - Reviews on Cancer* **1877**, 188645 (2022).

174. Jia, Y. *et al.* Overcoming EGFR T790M and C797S resistance with mutant-selective allosteric inhibitors. *Nature* **534**, 129 (2016).

175. de Clercq, D. J. H. *et al.* Discovery and Optimization of Dibenzodiazepinones as Allosteric Mutant-Selective EGFR Inhibitors. *ACS Med Chem Lett* **10**, 1549–1553 (2019).

176. To, C. *et al.* Single and dual targeting of mutant EGFR with an allosteric inhibitor. *Cancer Discov* **9**, 926 (2019).

177. To, C. *et al.* An allosteric inhibitor against the therapy-resistant mutant forms of EGFR in non-small cell lung cancer. *Nature Cancer 2022 3:4* **3**, 402–417 (2022).

178. Eno, M. S. *et al.* Discovery of BLU-945, a Reversible, Potent, and Wild-Type-Sparing Next-Generation EGFR Mutant Inhibitor for Treatment-Resistant Non-Small-Cell Lung Cancer. *J Med Chem* **65**, 9662–9677 (2022).

179. Kashima, K. *et al.* CH7233163 overcomes osimertinib-resistant EGFR-Del19/T790M/C797S Mutation. *Mol Cancer Ther* **19**, 2288–2297 (2020).

180. Lu, X. *et al.* Discovery of JND3229 as a New EGFRC797S Mutant Inhibitor with in Vivo Monodrug Efficacy. *ACS Med Chem Lett* **9**, 1123–1127 (2018).

181. Shen, J. *et al.* Structure-based design of 5-Methylpyrimidopyridone derivatives as new wild-type sparing inhibitors of the epidermal growth factor receptor triple mutant (EGFRL858R/T790M/C797S). *J Med Chem* **62**, 7302–7308 (2019).

182. Engelhardt, H. *et al.* Start Selective and Rigidify: The Discovery Path toward a Next Generation of EGFR Tyrosine Kinase Inhibitors. *J Med Chem* **62**, 10272–10293 (2019).

183. Uchibori, K. *et al.* Brigatinib combined with anti-EGFR antibody overcomes osimertinib resistance in EGFR-mutated non-small-cell lung cancer. *Nature Communications 2017 8:1* **8**, 1–16 (2017).

184. Wang, S., Song, Y. & Liu, D. EAI045: The fourth-generation EGFR inhibitor overcoming T790M and C797S resistance. *Cancer Lett* **385**, 51–54 (2017).

185. Beyett, T. S. *et al.* Molecular basis for cooperative binding and synergy of ATP-site and allosteric EGFR inhibitors. *Nature Communications 2022 13:1* **13**, 1–11 (2022).

186. Ríos-Hoyo, A., Moliner, L. & Arriola, E. Acquired Mechanisms of Resistance to Osimertinib—The Next Challenge. *Cancers (Basel)* **14**, (2022).

187. Liu, Y. *et al.* LS-106, a novel EGFR inhibitor targeting C797S, exhibits antitumor activities both in vitro and in vivo. *Cancer Sci* **113**, 709–720 (2022).

188. Lim, S. M. *et al.* MA07.09 BBT-176, a 4th generation EGFR TKI, for Progressed NSCLC after EGFR TKI Therapy: PK, Safety and Efficacy from Phase 1 Study. *Journal of Thoracic Oncology* **17**, S70–S71 (2022).

189. Lim, S. M. *et al.* 1365TiP A phase I/II, open-label study of BBT-176, a triple mutation targeting EGFR TKI, in patients with NSCLC who progressed after prior EGFR TKI therapy. *Annals of Oncology* **32**, S1035 (2021).

190. Liu, L. *et al.* Abstract 5462: BPI-361175, a 4th generation EGFR-TKI for the treatment of non-small cell lung cancer (NSCLC). *Cancer Res* **82**, 5462–5462 (2022).

191. Jang, J. *et al.* Discovery of a Highly Potent and Broadly Effective Epidermal Growth Factor Receptor and HER2 Exon 20 Insertion Mutant Inhibitor. *Angewandte Chemie - International Edition* **57**, 11629–11633 (2018).

192. Vyse, S. & Huang, P. H. Targeting EGFR exon 20 insertion mutations in non-small cell lung cancer. *Signal Transduction and Targeted Therapy 2019 4:1* **4**, 1–10 (2019).

193. Friedlaender, A. *et al.* EGFR and HER2 exon 20 insertions in solid tumours: from biology to treatment. *Nature Reviews Clinical Oncology 2021 19:1* **19**, 51–69 (2021).

194. Robichaux, J. P. *et al.* Mechanisms and clinical activity of an EGFR and HER2 exon 20–selective kinase inhibitor in non–small cell lung cancer. *Nature Medicine 2018 24:5* **24**, 638–646 (2018).

195. Yang, M. *et al.* NSCLC harboring EGFR exon-20 insertions after the regulatory C-helix of kinase domain responds poorly to known EGFR inhibitors. *Int J Cancer* **139**, 171–176 (2016).

196. Suda, K. *et al.* EP08.02-085 In vitro Activity and Potential Resistance Mutations Against BI-4020, a 4th-generation EGFR-TKI. *Journal of Thoracic Oncology* **17**, S442 (2022).

197. Naidoo, J. *et al.* Epidermal growth factor receptor exon 20 insertions in advanced lung adenocarcinomas: Clinical outcomes and response to erlotinib. *Cancer* **121**, 3212–3220 (2015).

198. Elamin, Y. Y. *et al.* Poziotinib for EGFR exon 20-mutant NSCLC: Clinical efficacy, resistance mechanisms, and impact of insertion location on drug sensitivity. *Cancer Cell* **40**, 754-767.e6 (2022).

199. Hou, J. *et al.* EGFR exon 20 insertion mutations in advanced non-small-cell lung cancer: current status and perspectives. *Biomark Res* **10**, 1–12 (2022).

200. Hasako, S. *et al.* TAS6417, a novel EGFR inhibitor targeting exon 20 insertion mutations. *Mol Cancer Ther* **17**, 1648–1658 (2018).

201. Gonzalvez, F. *et al.* Mobocertinib (Tak-788): A targeted inhibitor of egfr exon 20 insertion mutants in non–small cell lung cancer. *Cancer Discov* **11**, 1672–1687 (2021).

202. Nishino, M. *et al.* Activity of tarloxotinib-E in cells with EGFR exon-20 insertion mutations and mechanisms of acquired resistance. *Thorac Cancer* **12**, 1511–1516 (2021).

203. Prelaj, A. *et al.* Poziotinib for EGFR and HER2 exon 20 insertion mutation in advanced NSCLC: Results from the expanded access program. *Eur J Cancer* **149**, 235–248 (2021).

204. Cha, M. Y. *et al.* Antitumor activity of HM781-36B, a highly effective pan-HER inhibitor in erlotinib-resistant NSCLC and other EGFR-dependent cancer models. *Int J Cancer* **130**, 2445–2454 (2012).

205. Le, X. *et al.* Poziotinib in Non–Small-Cell Lung Cancer Harboring HER2 Exon 20 Insertion Mutations After Prior Therapies: ZENITH20-2 Trial. *Journal of Clinical Oncology* **40**, 710–718 (2022).

206. Ryan, C. FDA Issues Complete Response Letter to Poziotinib for Metastatic NSCLC Harboring HER2 Exon 20 Mutations. *OncLive* https://www.onclive.com/view/fda-issues-complete-response-letter-to-poziotinib-for-metastatic-nsclc-harboring-her2-exon-20-mutations (2022).

207. Bie, E. Spectrum lung cancer drug denied US FDA breakthrough tag | S&P Global Market Intelligence. *S&P Global Market Intelligence* https://www.spglobal.com/marketintelligence/en/news-insights/trending/xYHw56lV7eZYaM7kwUbqpw2 (2018).

208. Buntz, B. Black Diamond Therapeutics slashes workforce by 30% while discontinuing development of BDTX-189. *Drug Discovery & Development* https://www.drugdiscoverytrends.com/black-diamond-therapeutics-slashes-workforce-by-30-while-discontinuing-development-of-bdtx-189/ (2022).

209. Schram, A. M. *et al.* Safety and preliminary efficacy from the phase 1 portion of MasterKey-01: A First-in-human dose-escalation study to determine the recommended phase 2 dose (RP2D), pharmacokinetics (PK) and preliminary antitumor activity of BDTX-189, an inhibitor of allosteric ErbB mutations, in patients (pts) with advanced solid malignancies. *Journal of Clinical Oncology* **39**, 3086–3086 (2021).

210. Hasako, S. *et al.* TPC-064, a novel mutant-selective EGFR inhibitor, targets NSCLC driven by EGFR exon 20 insertion mutations. *Eur J Cancer* **69**, S125–S126 (2016).

211. Udagawa, H. *et al.* TAS6417/CLN-081 is a pan-mutation-selective EGFR tyrosine kinase inhibitor with a broad spectrum of preclinical activity against clinically-relevant EGFR mutations. *Mol Cancer Res* **17**, 2233 (2019).

212. Rosa, K. CLN-081 Granted FDA Breakthrough Therapy Designation Status for EGFR-Mutated NSCLC. *OncLive* https://www.onclive.com/view/cln-081-granted-fda-breakthrough-therapy-designation-status-for-egfr-mutated-nsclc (2022).

213. Wang, M. *et al.* Sunvozertinib, a Selective EGFR Inhibitor for Previously Treated Non–Small Cell Lung Cancer with EGFR Exon 20 Insertion Mutations. *Cancer Discov* **12**, 1676–1689 (2022).

214. Yang, J. C.-H. *et al.* Preliminary safety and efficacy results from phase 1 studies of DZD9008 in NSCLC patients with EGFR Exon20 insertion mutations. *Journal of Clinical Oncology* **39**, 9008–9008 (2021).

215. Tucker, N. FDA Grants Breakthrough Therapy Designation to DZD9008 for EGFR Exon20Ins-Positive NSCLC. *Targeted Oncology* https://www.targetedonc.com/view/fda-grants-breakthrough-therapy-designation-to-dzd9008-for-egfr-exon20ins-positive-nsclc (2022).

216. Remon, J., Hendriks, L. E. L., Cardona, A. F. & Besse, B. EGFR exon 20 insertions in advanced non-small cell lung cancer: A new history begins. *Cancer Treat Rev* **90**, 102105 (2020).

217. Siegel, F. *et al.* BAY 2927088: The first non-covalent, potent, and selective tyrosine kinase inhibitor targeting EGFR exon 20 insertions and C797S resistance mutations in NSCLC. *Eur J Cancer* **174**, S9–S10 (2022).

218. Siegel, F. *et al.* Abstract 1470: Preclinical activity of the first reversible, potent and selective inhibitor of EGFR exon 20 insertions. *Cancer Res* **81**, 1470–1470 (2021).

219. SIEGEL, S. *et al.* 4H-PYRROLO[3,2-C]PYRIDIN-4-ONE DERIVATIVES. *WIPO Patent No. 2019081486* (2019).

220. de Pas, T. *et al.* Activity of Epidermal Growth Factor Receptor-Tyrosine Kinase Inhibitors in Patients with Non-small Cell Lung Cancer Harboring Rare Epidermal Growth Factor Receptor Mutations. *Journal of Thoracic Oncology* **6**, 1895–1901 (2011).

221. Wang, F., Li, C., Wu, Q. & Lu, H. EGFR exon 20 insertion mutations in non-small cell lung cancer. *Transl Cancer Res* **9**, 2982–2991 (2020).

222. Vasconcelos, P. E. N. S. *et al.* EGFR-A763_Y764insFQEA Is a Unique Exon 20 Insertion Mutation That Displays Sensitivity to Approved and In-Development Lung Cancer EGFR Tyrosine Kinase Inhibitors. *JTO Clin Res Rep* **1**, 100051 (2020).

223. Robichaux, J. P. *et al.* Structure-based classification predicts drug response in EGFR-mutant NSCLC. *Nature 2021 597:7878* **597**, 732–737 (2021).

224. van Alderwerelt van Rosenburgh, I. K. *et al.* Biochemical and structural basis for differential inhibitor sensitivity of EGFR with distinct exon 19 mutations. *Nature Communications 2022 13:1* **13**, 1–16 (2022).

225. Cho, J. H. *et al.* Osimertinib for patients with non-small-cell lung cancer harboring uncommon EGFR mutations: A multicenter, open-label, phase II trial (KCSG-Lu15-09). *Journal of Clinical Oncology* **38**, 488–495 (2020).

226. Syed, Y. Y. Amivantamab: First Approval. *Drugs 2021 81:11* **81**, 1349–1353 (2021).

227. Yun, J. *et al.* Antitumor activity of amivantamab (Jnj-61186372), an egfr–met bispecific antibody, in diverse models of egfr exon 20 insertion–driven nsclc. *Cancer Discov* **10**, 1194–1209 (2020).

228. Neijssen, J. *et al.* Discovery of amivantamab (JNJ-61186372), a bispecific antibody targeting EGFR and MET. *Journal of Biological Chemistry* **296**, 100641 (2021).

229. Vijayaraghavan, S. *et al.* Amivantamab (JNJ-61186372), an Fc enhanced EGFR/cMet bispecific antibody, induces receptor downmodulation and antitumor activity by monocyte/macrophage trogocytosis. *Mol Cancer Ther* **19**, 2044–2056 (2020).

230. Amivantamab OK'd for EGFR-Mutant NSCLC. *Cancer Discov* **11**, 1604 (2021).

231. Markham, A. Mobocertinib: First Approval. *Drugs 2021 81:17* **81**, 2069–2074 (2021).

232. Riely, G. J. *et al.* Activity and Safety of Mobocertinib (TAK-788) in Previously Treated Non-Small Cell Lung Cancer with EGFR Exon 20 Insertion Mutations from a Phase I/II Trial. *Cancer Discov* **11**, 1688–1699 (2021).

233. Zhou, C. *et al.* Treatment Outcomes and Safety of Mobocertinib in Platinum-Pretreated Patients With EGFR Exon 20 Insertion–Positive Metastatic Non–Small Cell Lung Cancer: A Phase 1/2 Open-label Nonrandomized Clinical Trial. *JAMA Oncol* **7**, e214761–e214761 (2021).

234. Qin, Y. R. *et al.* Variability of EGFR exon 20 insertions in 24 468 Chinese lung cancer patients and their divergent responses to EGFR inhibitors. *Mol Oncol* **14**, 1695–1704 (2020).

235. Costa, D. B. *et al.* BIM mediates EGFR tyrosine kinase inhibitor-induced apoptosis in lung cancers with oncogenic EGFR mutations. *PLoS Med* **4**, 1669–1680 (2007).

236. Cox, M. M. J. Innovations in the Insect Cell Expression System for Industrial Recombinant Vaccine Antigen Production. *Vaccines (Basel)* **9**, (2021).

237. Kitts, P. A. & Possee, R. D. A method for producing recombinant baculovirus expression vectors at high frequency. *Biotechniques* **14**, 810–817 (1993).

238. Barker, S. C. *et al.* Characterization of pp60c'irc Tyrosine Kinase Activities Using a Continuous Assay: Autoactivation of the Enzyme Is an Intermolecular Autophosphorylation Process1. *Biochemistry* **34**, 14843–14851 (1995).

239. McFarlane, C. & Murray, J. A Sensitive Coupled Enzyme Assay for Measuring Kinase and ATPase Kinetics Using ADP-Specific Hexokinase. *Bio Protoc* **10**, (2020).

240. Red Brewer, M. *et al.* The Juxtamembrane Region of the EGF Receptor Functions as an Activation Domain. *Mol Cell* **34**, 641–651 (2009).

241. Zhao, X., Li, G. & Liang, S. Several affinity tags commonly used in chromatographic purification. *J Anal Methods Chem* **2013**, (2013).

242. Maertens, B., Spriestersbach, A., Kubicek, J. & Schäfer, F. Strep-Tagged Protein Purification. *Methods Enzymol* **559**, 53–69 (2015).

243. Booth, W. T. *et al.* Impact of an N-terminal polyhistidine tag on protein thermal stability. *ACS Omega* **3**, 760–768 (2018).

244. Schmidt, T. G. M. *et al.* Development of the Twin-Strep-tag® and its application for purification of recombinant proteins from cell culture supernatants. *Protein Expr Purif* **92**, 54–61 (2013).

245. Young, C. L., Britton, Z. T. & Robinson, A. S. Recombinant protein expression and purification: A comprehensive review of affinity tags and microbial applications. *Biotechnol J* **7**, 620–634 (2012).

246. Grace, M. R., Walsh, C. T. & Cole, P. A. Divalent ion effects and insights into the catalytic mechanism of protein tyrosine kinase Csk. *Biochemistry* **36**, 1874–1881 (1997).

247. Sequist, L. v., Bell, D. W., Lynch, T. J. & Haber, D. A. Molecular predictors of response to epidermal growth factor receptor antagonists in non-small-cell lung cancer. *Journal of Clinical Oncology* **25**, 587–595 (2007).

248. Mok, T. S. *et al.* Gefitinib or Carboplatin–Paclitaxel in Pulmonary Adenocarcinoma. *New England Journal of Medicine* **361**, 947–957 (2009).

249. Wakeling, A. E. *et al.* ZD1839 (Iressa): An Orally Active Inhibitor of Epidermal Growth Factor Signaling with Potential for Cancer Therapy. *Cancer Res* **62**, 5749–5754 (2002).

250. Mitsudomi, T. & Yatabe, Y. Mutations of the epidermal growth factor receptor gene and related genes as determinants of epidermal growth factor receptor tyrosine kinase inhibitors sensitivity in lung cancer. *Cancer Sci* **98**, 1817–1824 (2007).

251. Graham, J., Muhsin, M. & Kirkpatrick, P. Cetuximab. *Nat Rev Drug Discov* **3**, 549–550 (2004).

252. Lacouture, M. E. Mechanisms of cutaneous toxicities to EGFR inhibitors. *Nature Reviews Cancer 2006 6:10* **6**, 803–812 (2006).

253. Degorce, F. *et al.* HTRF: A technology tailored for drug discovery - A review of theoretical aspects and recent applications. *Curr Chem Genomics* **3**, 22–32 (2009).

254. Collins, M. K. L. *et al.* Transfer of functional EGF receptors to an IL3-dependent cell line. *J Cell Physiol* **137**, 293–298 (1988).

255. Jiang, J. *et al.* Epidermal Growth Factor–Independent Transformation of Ba/F3 Cells with Cancer-Derived Epidermal Growth Factor Receptor Mutants Induces Gefitinib-Sensitive Cell Cycle Progression. *Cancer Res* **65**, 8968–8974 (2005).

256. Greulich, H. *et al.* Oncogenic Transformation by Inhibitor-Sensitive and -Resistant EGFR Mutants. *PLoS Med* **2**, e313 (2005).

257. Shah, P., Andrew, & & Westwell, D. D. The role of fluorine in medicinal chemistry. *J Enzyme Inhib Med Chem* **22**, 527–540 (2007).

258. Zhai, X., Ward, R. A., Doig, P. & Argyrou, A. Insight into the Therapeutic Selectivity of the Irreversible EGFR Tyrosine Kinase Inhibitor Osimertinib through Enzyme Kinetic Studies. *Biochemistry* **59**, 1428–1441 (2020).

259. Copeland, R. A. Irreversible Enzyme Inactivators. *Evaluation of Enzyme Inhibitors in Drug Discovery* 345–382 (2013) doi:10.1002/9781118540398.CH9.

260. Strelow, J. M. A Perspective on the Kinetics of Covalent and Irreversible Inhibition. *SLAS Discovery* **22**, 3–20 (2017).

261. Krippendorff, B. F., Neuhaus, R., Lienau, P., Reichel, A. & Huisinga, W. Mechanism-Based Inhibition: Deriving KI and kinact Directly from Time-Dependent IC50 Values. *SLAS Discovery* **14**, 913–923 (2009).

262. Cha, S. Tight-binding inhibitors—I: Kinetic behavior. *Biochem Pharmacol* **24**, 2177–2185 (1975).

263. Copeland, R. A. Drug–Target Residence Time. *Evaluation of Enzyme Inhibitors in Drug Discovery* 287–344 (2013) doi:10.1002/9781118540398.CH8.

264. Szedlacsek, S. E. & Duggleby, R. G. Kinetics of Slow and Tight-Binding Inhibitors. *Methods Enzymol* **249**, (1995).

265. Williams, J. W. & Morrison, J. F. The Kinetics of Reversible Tight-Binding Inhibition. *Methods Enzymol* **63**, (1979).

266. Stone, S. R. & Morrison, J. F. Mechanism of inhibition of dihydrofolate reductases from bacterial and vertebrate sources by various classes of folate analogues. *Biochimica et Biophysica Acta (BBA)/Protein Structure and Molecular* **869**, (1986).

267. Morin, A. *et al.* Collaboration gets the most out of software. *Elife* **2**, (2013).

268. Winter, G. *et al.* DIALS: implementation and evaluation of a new integration package. *Acta Crystallogr D Struct Biol* **74**, 85–97 (2018).

269. Winter, G. Xia2: An expert system for macromolecular crystallography data reduction. *J Appl Crystallogr* **43**, 186–190 (2010).

270. Adams, P. D. *et al.* PHENIX: a comprehensive Python-based system for macromolecular structure solution. *Acta Crystallogr D Biol Crystallogr* **66**, 213–221 (2010).

271. Emsley, P., Lohkamp, B., Scott, W. G. & Cowtan, K. Features and development of Coot. *Acta Crystallogr D Biol Crystallogr* **66**, 486–501 (2010).

272. Moriarty, N. W., Grosse-Kunstleve, R. W. & Adams, P. D. electronic Ligand Builder and Optimization Workbench (eLBOW): a tool for ligand coordinate and restraint generation. *Acta Crystallogr D Biol Crystallogr* **65**, 1074 (2009).

273. Mohammadi, M. *et al.* Structures of the tyrosine kinase domain of fibroblast growth factor receptor in complex with inhibitors. *Science (1979)* **276**, 955–960 (1997).

274. Schindler, T. *et al.* Structural mechanism for STI-571 inhibition of Abelson tyrosine kinase. *Science (1979)* **289**, 1938–1942 (2000).

275. Yosaatmadja, Y. *et al.* Binding mode of the breakthrough inhibitor AZD9291 to epidermal growth factor receptor revealed. *J Struct Biol* **192**, 539–544 (2015).

276. Lategahn, J. *et al.* Insight into Targeting Exon20 Insertion Mutations of the Epidermal Growth Factor Receptor with Wild Type-Sparing Inhibitors. *J Med Chem* (2021) doi:10.1021/ACS.JMEDCHEM.1C02080/SUPPL_FILE/JM1C02080_SI_006.PDF.

277. Gunaydin, H. & Bartberger, M. D. Stacking with No Planarity? *ACS Medicinal Chemistry Letters* vol. 7 Preprint at https://doi.org/10.1021/acsmedchemlett.6b00099 (2016).

278. Molcanov, K. & Kojić-Prodić, B. Towards understanding π-stacking interactions between non-Aromatic rings. *IUCrJ* vol. 6 Preprint at https://doi.org/10.1107/S2052252519000186 (2019).

279. Cheng, H. *et al.* Discovery of 1-{(3R,4R)-3-[({5-Chloro-2-[(1-methyl-1H-pyrazol-4-yl)amino]-7H-pyrrolo[2,3-d]pyrimidin-4-yl}oxy)methyl]-4-methoxypyrrolidin-1-yl}prop-2-en-1-one (PF-06459988), a Potent, WT Sparing, Irreversible Inhibitor of T790M-Containing EGFR Mutants. *J Med Chem* **59**, 2005–2024 (2016).

280. Gajiwala, K. S. *et al.* Insights into the aberrant activity of mutant EGFR kinase domain and drug recognition. *Structure* **21**, 209–219 (2013).

281. Olson, C. M. *et al.* Development of a Selective CDK7 Covalent Inhibitor Reveals Predominant Cell-Cycle Phenotype. *Cell Chem Biol* **26**, (2019).

282. Hassan, M., Brown, R. D., Varma-O'brien, S. & Rogers, D. Cheminformatics analysis and learning in a data pipelining environment. *Mol Divers* **10**, 283–299 (2006).

283. Rogers, D. & Hahn, M. Extended-connectivity fingerprints. *J Chem Inf Model* **50**, 742–754 (2010).

284. Zimmerman, J. J., von Saint André-von Arnim, A. & McLaughlin, J. Chapter 74 – Cellular Respiration. in *Pediatric Clinical Care* (2011).