Biochemical and structural basis for differential inhibitor sensitivity of EGFR with distinct exon 19 mutations

Tyrosine kinase inhibitors (TKIs) are used to treat non-small cell lung cancers (NSCLC) driven by epidermal growth factor receptor (EGFR) mutations in the tyrosine kinase domain (TKD). TKI responses vary across tumors driven by the heterogeneous group of exon 19 deletions and mutations, but the molecular basis for these differences is not understood. Using purified TKDs, we compared kinetic properties of several exon 19 variants. Although unaltered for the second generation TKI afatinib, sensitivity varied significantly for both the first and third generation TKIs erlotinib and osimertinib. The most sensitive variants showed reduced ATP-binding affinity, whereas those associated with primary resistance retained wild type ATP-binding characteristics (and low \( k_{M,\text{ATP}} \)). Through crystallographic and hydrogen-deuterium exchange mass spectrometry (HDX-MS) studies, we identify possible origins for the altered ATP-binding affinity underlying TKI sensitivity and resistance, and propose a basis for classifying uncommon exon 19 variants that may have predictive clinical value.

Almost two decades since the discovery of activating mutations in the tyrosine kinase domain (TKD) of the epidermal growth factor receptor (EGFR) that drive non-small cell lung cancer (NSCLC), tyrosine kinase inhibitors (TKIs) are now used routinely as the first-line treatment to target EGFR with significant clinical benefit. The most frequently observed EGFR mutations in NSCLC patients, together accounting for over 90% of oncogenic EGFR aberrations in this cancer\(^1\)–\(^4\), are in exons 19 (45%) and 21 (46%). The most common exon 19 mutation, which accounts for -33% of EGFR mutations in NSCLC, causes a 5-residue deletion (\( \Delta E746-A750 \)) from the loop connecting the third \( \beta \)-strand of the EGFR TKD (\( \beta_3 \)) to the key regulatory \( \alpha C \) helix (Fig. 1a). The most common exon 21 mutation is L858R in the TKD activation loop, which accounts for 44% of EGFR mutations in this cancer\(^1\). Five TKIs (erlotinib, gefitinib, afatinib, dacomitinib, and osimertinib) have been approved by the FDA for the treatment of EGFR-mutated NSCLC with exon 19 deletions or an L858R mutation, with first generation TKIs such as erlotinib previously recommended as optimal first-line treatment\(^6\). Unfortunately, development of resistance to these TKIs is very common, most frequently through emergence of the T790M mutation in exon 20. Efforts to improve patient outcomes led current National Comprehensive Cancer Network clinical practice guidelines to suggest the third generation irreversible TKI osimertinib for first-line treatment of EGFR-mutated NSCLC\(^5\).

1Department of Pharmacology, Yale University School of Medicine, New Haven, CT 06520, USA. 2Yale Cancer Biology Institute, Yale University West Campus, West Haven, CT 06516, USA. 3Yale Cancer Center, Yale University School of Medicine, New Haven, CT 06520, USA. 4Department of Medicine (Medical Oncology), Yale School of Medicine, New Haven, CT 06520, USA. 5Yale Center for Analytical Sciences, Yale School of Public Health, New Haven, CT 06520, USA. 6Department of Pathology, Yale University School of Medicine, New Haven, CT 06520, USA. 7These authors contributed equally: Iris K. van Alderwerelt van Rosenburgh, David M. Lu. e-mail: mark.lemmon@yale.edu; kumar.ashtekar@yale.edu; yuko.tsutsui@yale.edu
The Catalogue of Somatic Mutations in Cancer (COSMIC) lists over 100 different exon 19 mutations (including many singletons) in lung cancer, but the predominance of the 

exon 19 mutations seen in NSCLC, with a focus on their TKI sensitivities. These and other clinical studies have reported differential sensitivity of individual EGFR exon 19 mutations to EGFR TKIs in clinical settings, however—

The Catalogue of Somatic Mutations in Cancer (COSMIC) lists over 100 different exon 19 mutations (including many singletons) in lung cancer, but the predominance of the \Delta E746-A750 variant (Fig. 1b) —which accounts for 75% in this database—has led current clinical recommendations to consider them all as a single group. Numerous studies have reported differential sensitivity of individual EGFR exon 19 variants to EGFR TKIs in clinical settings, however\(^\text{19}^{-23}\). For example, in a cohort of 202 patients with tumors harboring various exon 19 deletions, uncommon deletion variants were associated with significantly worse survival than the common \Delta E746-A750 mutation with the first generation TKI gefitinib\(^\text{24}\). In addition, our group previously reported significantly shorter progression-free and overall survival in erlotinib-treated patients with tumors carrying the \Delta L747-A750InsP mutation compared to those with tumors driven by the common \Delta E746-A750 mutation\(^\text{24}\). Moreover, the L747P exon 19 mutation has been reported to be associated with primary resistance to first and second generation TKIs\(^\text{24,25}\), but to show a significantly better response and progression-free survival to the second generation TKI afatinib\(^\text{22}\). These and other clinical findings argue for the importance of understanding the molecular basis for the differential TKI sensitivity of exon 19-mutated variants.

Here, we describe biochemical and structural studies of EGFR exon 19 variants seen in NSCLC, with a focus on their TKI sensitivities. We previously described differential sensitivity of the common \Delta E746-A750 variant and the uncommon \Delta L747-A750InsP variant\(^\text{24}\). We recapitulate this observation biochemically and find that it reflects differences in ATP-binding affinity (and \(K_{M, ATP}\)) of the two variants, without measurable alterations in the inhibition constant (\(K_i\)) for erlotinib. By
AFatinib was only ~1.3-fold and was not statistically significant.

The increase in IC50 was greater for erlotinib (~7.5-fold) than for osimertinib (~2-fold increase in IC50), whereas the increase in IC50 for afatinib was only ~1.3-fold and was not statistically significant (P = 0.24).

As in our previously published study3, similar differences were also observed when assessing inhibition of intact EGFR variants expressed in CHO cells (right panels in Fig. 1e–g) – although the present studies did indicate slightly reduced afatinib sensitivity for ΔL747-Δ750nInsP in CHO cells, unlike our previous work. IC50 values cannot be directly compared across inhibitors in these studies, since erlotinib is a reversible inhibitor whereas osimertinib and afatinib are both irreversible covalent inhibitors.

### Results

**Isolated kinase domains recapitulate exon 19 variant TKI sensitivities**

In an effort to understand the biochemical basis for the differential TKI sensitivities of EGFR exon 19 variants, we first asked whether purified EGFR tyrosine kinase domains (TKDs) recapitulate the reduced erlotinib sensitivity seen in cells and patients for the ΔL747-Δ750nInsP variant2. We expressed and purified isolated EGFR TKDs harboring either the common exon 19 deletion (ΔE746-Δ750) or the uncommon ΔL747-Δ750nInsP mutation using a baculovirus expression system (numbers here correspond to UniProt entry P00533, and are divided by 10).

### Table 1 | Biochemical properties of EGFR TKDs with different mutations

| Mutation | IC50 erlotinib (nM) | IC50 afatinib (nM) | IC50 osimertinib (nM) | Kᵢ erlotinib (nM) | Kᵢ ATP (μM) | Kᵢ pept (μM) | apparent k<sub>cat</sub> (s⁻¹) | corrected k<sub>cat</sub> (s⁻¹) |
|----------|---------------------|-------------------|-----------------------|------------------|-------------|-------------|-----------------|-----------------|
| None (wild type) | n.d. | n.d. | n.d. | 12 ± 3 | 406 | 0.042 | 0.042 |
| **Profile 1** | | | | | | | |
| ΔL747-Δ750nInsP | 232 ± 48 | 36 ± 15 | 65 ± 21 | 5.3 ± 1.5 | 23 ± 6 | 15 | 0.91 | 1.26 |
| L747P | 141 ± 17 | 37 ± 4 | 159 ± 89 | 6.2 ± 4.0 | 21 ± 6 | 52 | 0.84 | 1.14 |
| ΔL747-Δ749 | 108 ± 17 | 42 ± 4⁻ | >1000 | n.d. | 13 ± 8 | n.d. | -0.044 | 0.052 |
| **Profile 2** | | | | | | | |
| ΔE746-Δ750 | 31 ± 8 | 27 ± 9 | 37 ± 17 | n.d. | 158 ± 24 | 261 | 0.60 | 1.1 |
| ΔE746-Δ751nInsA | 20 ± 3 | 8 ± 1 | 11 ± 2 | n.d. | 240 ± 59 | 156 | 0.12 | 0.75 |
| ΔL747-Δ751nInsP | 14 ± 5 | 6 ± 2 | 7 ± 2 | n.d. | 164 ± 48 | 38 | 0.16 | 1.33 |
| ΔL747-Δ753nInsS | 37 ± 10 | 23 ± 3 | 26 ± 16 | 4.2 ± 1.7 | 101 ± 22 | 292 | 0.49 | 1.07 |
| ΔS752-Δ759 | 25 ± 5 | 20 ± 7 | 22 ± 7 | n.d. | 91 ± 25 | 80 | 1.2 | 3.0 |
| **Other** | | | | | | | |
| L858R | 75 ± 12 | 23 ± 2 | 52 ± 2 | 4.7 ± 2.3 | 74 ± 14 | 218 | 1.4 | 3.04 |
| L858R/Δ750nInsM | -10,000 | 92 ± 23 | 101 ± 24 | 41 ± 14 | 45 ± 6 | 99 | 6.6 | - |

*IC₅₀ determinations for ΔL747-Δ749 were performed at 1 μM TKD, whereas experiments with all other variants were performed at 100 nM TKD. IC₅₀ values for ΔL747-Δ749 have therefore been divided by 10.*

*Titration to maximum peptide concentration was only performed once, because of the large amounts of peptide required, so Kᵢ pept and apparent k<sub>cat</sub> values are estimates with no S.D. quoted (n = 1).*  

*As afatinib is highly potent and its IC₅₀ is frequently less than 50% of the TKD concentration, we assumed an endpoint titration to estimate the active TKD concentration (equal to afatinib IC₅₀, which we assume is 50% of active [TKD]), and have corrected the apparent k<sub>cat</sub> based on this number. L858R/Δ750nInsM was excluded from this analysis.*

In vitro quantitation of inhibitor sensitivity of other exon 19 variants

Having recapitulated the previously-observed differences between ΔE746-Δ750 and ΔL747-Δ750nInsP EGFR variants with in vitro biochemical studies, our next goal was to extend this analysis to other unstudied exon 19 variants. We generated expression constructs for an additional 13 different exon 19 variants (see Supplementary Fig. 1a), focusing on those that occur most frequently in lung cancer11,12. Unfortunately, about half of these variants expressed poorly or aggregated during purification; this likely explains why no crystal structure has yet been reported for EGFR’s TKD with an exon 19 deletion. Nonetheless, we could generate high quality TKD protein for six further exon 19 variants (Supplementary Fig. 1b, c), including TKD containing the L747P mutation—reported to be associated with clinical resistance to erlotinib and gefitinib13,14. Figure 1b lists the mutations in the β3/αc loop (Fig. 1a, b) that we could study biochemically, as well as their frequencies among exon 19 variants in the COSMIC database1. Kinetic studies showed that most of the activated exon 19 TKD variants had apparent k<sub>cat</sub> values ranging from 0.5 to 1.2 s⁻¹ (Table 1) compared with <0.05 s⁻¹ for wild type and 1.4 s⁻¹ for the L858R-mutated TKD, in agreement with previous reports15,31. Both ΔE746-Δ751nInsA and ΔL747-Δ751nInsP had lower apparent k<sub>cat</sub> values (0.12 and 0.16 s⁻¹ respectively). This presumably reflects smaller proportions of active protein in these preparations, given the very low values for afatinib IC₅₀ determined in these cases. As one approach to correct for this effect, we assumed that IC₅₀ for afatinib (see below) under these conditions is equal to 50% of the active TKD concentration, and corrected k<sub>cat</sub> accordingly (right-most column in Table 1). The resulting corrected k<sub>cat</sub> values range from 0.75 to 1.33 s⁻¹ with two outliers. ∆S752-Δ759 appears approximately two-fold more active, and the ΔL747-Δ749 variant did not show elevated kinase activity (Table 1), which was unexpected.

Comparing a series of different exon 19-mutated variants, we show that they fall into two categories. The common (ΔE746-Δ750) variant epitomizes a group of variants with reduced ATP-binding affinity that are sensitized to first and third generation inhibitors. The uncommon ΔL747-Δ750nInsP variant exemplifies a distinct group in which ATP-binding affinity (and Kᵢ ATP) remains at wild type levels, reducing sensitivity to erlotinib and osimertinib. Importantly, we find that afatinib sensitivity is relatively unchanged between the two profiles, which has clinical implications. Determining the first crystal structure of an exon 19-mutated EGFR TKD and analyzing structural dynamics using hydrogen-deuterium exchange mass spectrometry (HDX-MS) further suggested explanations for the different ATP-binding characteristics of exon 19 variants. This in turn allowed us to predict that exon 19 variants with β3/αc loop deletions of ≤3 residues will show primary resistance to first and third generation TKIs. Analysis of outcomes in erlotinib-treated patients was consistent with this prediction, which may be valuable for guiding future clinical decisions.
We next analyzed inhibition of the exon 19 variants (at 50–100 nM protein) by erlotinib, osimertinib, and afatinib (at 1 mM ATP). IC_{50} values corresponded to ~50% of active TKD concentration in most cases (~20–40 nM), suggesting very high afinity inhibitor binding—although the ΔL747-A750InsP and ΔE746-A750InsA variants gave lower values (indicating that these preparations have reduced levels of active kinase). Whereas this high sensitivity to afatinib was maintained across all exon 19 variants studied (green bars in Fig. 2a), a subset of variants showed significantly reduced sensitivity to erlotinib and osimertinib (Fig. 2a; see also Table 1 and Supplementary Fig. 2). Thus, two distinct sensitivity classes or profiles could be defined. In the first (profile 1: Fig. 2a), which comprises ΔL747-A750InsP and L747P, sensitivity to erlotinib and osimertinib was substantially reduced (IC_{50} increased) compared with ΔE746-A750. In profile 2 variants, by contrast, sensitivity to erlotinib and osimertinib was more similar to that seen for the common ΔE746-A750 variant (Fig. 2a). The L858R control appears to be intermediate between the two profiles. L858R is slightly less sensitive to erlotinib than ΔE746-A750 (by ~2.5-fold)—consistent with previous reports. It is important to note that we cannot directly compare sensitivity across different inhibitors without time-dependent inhibition studies.

The reduced erlotinib sensitivity of profile 1 variants was also discernible for mutated full-length receptors expressed in CHO cells (Fig. 2b)–with L747P showing the effect most clearly, consistent with...
clinical reports. We also investigated another rare exon 19 deletion variant, ΔL747-E749, which has 33 entries in COSMIC. Unexpectedly, we were unable to express the intact receptor with this mutation in CHO cells, and the purified TKD was much less active in vitro than TKDs with the other mutations listed in Fig. 1b. When corrected for the fact that experiments with this variant required 10-fold more TKD, estimated IC50 values for ΔL747-E749 were 108 nM for erlotinib, 42 nM for afatinib, and ~1000 nM for osimertinib (Supplementary Fig. 2f, Table 1) — placing it in profile 1.

These data demonstrate that sensitivity to a given inhibitor can differ significantly between individual EGFR exon 19 variants at the level of the kinase domain itself. Most notably, and consistent with recent clinical data, we find that relative resistance to erlotinib is seen with L747P and ΔL747-A750InsP (in profile 1)—but remarkably not with ΔL747-E749 (in profile 2), which differs solely in having just one more amino acid deleted from its β3/αC loop.

Profile 1 and profile 2 variants have similar drug-binding affinities

We next wanted to understand the origin of different inhibitor sensitivities of EGFR exon 19 deletion variants. We undertook steady state kinetic measurements of TKD activities to measure their Michaelis constants for ATP (Km, ATP) and, where possible, inhibition constants (KI) for erlotinib using the Cheng–Prusoff equation. We were able to obtain reliable KI estimates for the ΔL747-A750InsP, L747P, and ΔL747-P753InsS variants (Fig. 2c–e), which all gave values between 4.2 nM (ΔL747-P753InsS) and 6.2 nM (ΔL747-A750InsP), similar to previously reported values. Importantly, the erlotinib KI values for the profile 1 ΔL747-A750InsP and L747P variants were both within 1.3–1.5 fold of KI for the profile 2 variant ΔL747-P753InsS, despite their erlotinib IC50 values being 4–6 fold higher (Fig. 2a). The small differences observed in KI when compared with ΔL747-P753InsS are statistically significant (P = 0.02 for L747P and P = 0.01 for ΔL747-A750InsP). However, they are not sufficient to explain the reduced erlotinib sensitivity of these two profile 1 variants, arguing that reduced drug-binding affinity cannot fully explain resistance to erlotinib—contrary to the suggestions of our previous computational studies. Equivalent experiments with L858R- and L858R/T790M-mutated TKDs (Supplementary Fig. 3a, b) gave KI values for 4.7 ± 2.3 and 41 ± 14 nM respectively, indicating that some of the erlotinib resistance caused by the T790M mutation must arise from reduced drug-binding affinity (increasing KI)—consistent with previous studies. Unexpectedly, Cheng–Prusoff equation plots for other profile 2 variants showed a lack of dependence of IC50 on ATP concentration over a substantial range (Supplementary Fig. 3c, d), suggesting either non-competitive inhibition effects that are difficult to understand and/or non-specific binding of inhibitor to inactive protein in these preparations. As a result, we could not estimate KI values for the ΔL747-A750, ΔE746-T751InsA, ΔL747-T751InsP or ΔS752-I759 variants.

TKI sensitivity is determined by Km, ATP

Previous studies of acquired TKI resistance in EGFR have shown that alterations in ATP-binding affinity play a determining role. Measurements of Km, ATP for all the EGFR variants listed in Fig. 1b revealed that this is also true for exon 19 mutations in EGFR. Remarkably, the profile 1 and profile 2 variants also cluster in the same groups based on their Km, ATP values (Fig. 3a, b). Notably, Km, ATP for ΔL747-A750InsP is ~seven-fold lower (23 ± 6 nM) than the value of 158 ± 24 nM measured for ΔE746-A750 (P < 0.0001), protein preparations, with three independent experiments for each. c Overlay of the 2.96 Å crystal structure of the ΔL747-E749 exon 19 variant (PDB ID: 7TVD – see Table 2), shown in black, with the wild type EGFR TKD from PDB entry 4M16, shown in grey. Both TKDs are in the active conformation, but the deletion in the β3/αC loop displaces αC slightly towards the ATP-binding site in ΔL747-E749 (see green arrow). The residues deleted in ΔL747-E749 (L747, R748, and E749) are shown in orange in the wild type structure. d Close-up of the β3/αC region in the overlay of ΔL747-E749 on the wild type EGFR TKD, showing that β3 is slightly truncated at its C-terminus in ΔL747-E749 and αC loses a helical turn at its N-terminus—allowing the shortened β3/αC loop to link the two elements. Source data are provided as a Source data file.
This difference in $K_m\text{_{ATP}}$ can almost fully account for the 7.5-fold difference in erlotinib $IC_{50}$ between the two variants. Similarly, $K_m\text{_{ATP}}$ for the L747P variant (7.5-fold smaller than for $\Delta E746-A750$; $P < 0.0001$) can more than account for its five-fold reduced sensitivity to erlotinib. All the exon 19 variants in profile 2 have higher $K_m\text{_{ATP}}$ values—ranging from 91 to 240 $\mu$M (Table 1)—all of which are greater than the $K_m\text{_{ATP}}$ value that we measured for L858R-mutated TKD (74 $\mu$M).

These data suggest that—as with L858R—most activating exon 19 mutations reduce ATP-binding affinity (higher $K_m\text{_{ATP}}$), resulting in enhanced sensitivity to ATP-competitive inhibitors like erlotinib. The increase in $K_m\text{_{ATP}}$ value appears to be greater for these exon 19 mutations than for L858R, suggesting that they may promote higher sensitivity to TKIs than L858R—possibly explaining differences in outcome between these two groups first described in early clinical studies. By contrast, the profile 1 exon 19 variants (L747P and $\Delta L747-A750$) activate EGFR without increasing $K_m\text{_{ATP}}$ (Table 1), and retain the ATP-binding characteristics of the wild type kinase and the $\Delta L747-E749$ variant (Fig. 3b). This is associated with lower TKI sensitivity (and clinical resistance). As previously reported by Eck and colleagues, the T790M mutation reverses the $K_m\text{_{ATP}}$ increase caused by the L858R mutation to reduce TKI sensitivity. However, T790M does not fully bring $K_m\text{_{ATP}}$ down to wild type levels under our assay conditions (Fig. 3b and Supplementary Fig. 3e), and the $K_m\text{_{ATP}}$ change is not sufficient to account for the >130-fold increase in $IC_{50}$ for erlotinib (Supplementary Fig. 2g, h) seen for the L858R/T790M double mutation. Taken together, our findings with exon 19 variants suggest that, just as suggested for T790M, alterations in ATP-binding affinity are the primary (if not sole) cause of reduced sensitivity to erlotinib and osimertinib.

**Crystal structure of $\Delta L747-E749$ EGFR TKD**

Despite significant efforts, we were unable to crystallize any of the activated exon 19 EGFR TKD variants listed in Fig. 1b. We did succeed in determining a 3 Å crystal structure of the $\Delta L747-E749$ variant, however (PDB ID 7TVD, see Supplementary Fig. 4 and Table 2). This variant was indistinguishable in kinetic assays from wild type (Table 1), unexpectedly failing to show elevated kinase activity. The $\Delta L747-E749$ variant crystallized in the same space group as the wild type EGFR TKD, and formed the same symmetric dimer reported to stabilize the active kinase conformation in crystals of other EGFR TKDs. As shown in Fig. 3c, d, the only discernible differences between the wild type TKD and $\Delta L747-E749$ are a shortening of strand $\beta 3$ (from its C-terminal end) and loss of the first turn of helix $\alpha C$. These changes result from deletion of the three residues (L747, R748, and E749: orange in Fig.3d) from the beginning of the $\beta 3/\alpha C$ loop. Truncation of both $\beta 3$ and $\alpha C$ is necessary to allow the truncated $\alpha C$ to the ATP-binding site compared with its position in dimers of wild type, L858R or other variants with activating mutations (green arrow in Fig. 3c). The predicted salt bridge between E762 in $\alpha C$ and K745 in strand $\beta 3$, required to stabilize ATP interactions of the latter in the active TKD, is precisely maintained in $\Delta L747-E749$ (Supplementary Fig. 4). The $\Delta L747-E749$ TKD structure also shows that a 3-residue deletion from the $\beta 3/\alpha C$ loop can be tolerated without major disruption. AlphaFold$^{12}$—based predictions performed using ColabFold$^{9}$ suggested that longer deletions in profile 2 variants (notably $\Delta S752-I759$) further truncate $\alpha C$ from its amino-terminus to allow the loop still to reach between the $\beta 3$ and $\alpha C$ secondary structure elements (Supplementary Fig. 5). This likely leads to more profound distortion and/or alterations in $\alpha C$ position and interactions—consistent with our failure to crystallize such exon 19 variants. The profile 1 variants have less truncated $\alpha C$ helices.

---

**Table 2 | Crystallization conditions, data collection, and refinement statistics**

| Amino acid residue boundaries | 696–1022 |
|-----------------------------|---------|
| PDB ID                      | 7TVD    |
| Crystallization conditions   | Protein in 20 mM Tris pH 8.0, 150 mM NaCl and 2% glycerol; Reservoir solution 28% PEG 400, 0.1 M HEPES pH 7.5, 0.2 M CaCl$_2$. Ratio = 1:1, 16 °C |
| Date of collection           | March 19, 2019 |
| Wavelength (Å)               | 0.97918 |
| Space group                  | I 2 3   |
| Cell dimensions              | a, b, c (Å) 149.42, 149.42, 149.42 |
| α, β, γ (°)                  | 90.00, 90.00, 90.00 |
| Resolution (Å)               | 105.66–2.96 |
| Completeness                 | 99.9 (99.7) |
| Redundancy                   | 36.1 (33.5) |
| R$_{sym}$ (%)                | 44 (263) |
| l/σ                          | 15.4 (1.5) |
| CC1/2 (%)                    | 0.999 (0.44) |
| Number of reflections        | 11,749 (1880) |
| Refinement                   | R$_{work}$/R$_{free}$ (%) 24.0/27.7 |

| Number of atoms              |
| Protein                     | 2338 |
| Ions                        | 0 |
| Ligands                     | 0 |
| Water                       | 0 |

| Average B factor (Å)         |
| Protein                     | 104.0 |
| Ions                        | - |
| Ligands                     | - |
| Water                       | - |

| Geometry (Ramachandran)      |
| Favor($\%$)                  | 94.06 |
| Allowed($\%$)                | 5.24 |
| Outliers($\%$)               | 0.70 |
| RMSD (Å)                     | 0.004 |
| Bond length                  | 0.004 |
| Bond angle                   | 0.77 |

*Numbers in parentheses denote highest resolution shell.

It is not clear from this analysis why the $\Delta L747-E749$ TKD is not activated whereas $\Delta L747-A750$nsP and L747P are both among the most active of the variants (Table 1). These two activated variants have slight amino-terminal truncations of $\alpha C$ compared with $\Delta L747-E749$ in AlphaFold models (Supplementary Fig. 5a), but share the introduction of a proline at position 747 (two residues after the $\beta 3$ lysine). This could alter local dynamics in a way that mimics a larger deletion to activate the kinase.

**HDMS shows variant-specific differences in structural dynamics**

To gain insight into structural differences between profile 1 and profile 2 exon 19 variants, we used hydrogen–deuterium exchange mass spectrometry (HDMS), which probes region-specific structural dynamics by measuring rates of backbone amide hydrogen atom exchange$^{11,14}$. Because amide hydrogens in structurally flexible

---

*Notes:*

- $K_m\text{_{ATP}}$: Michaelis constant of ATP
- $IC_{50}$: Half-maximal inhibitory concentration
- $\alpha C$: Alpha C helix
- $\beta 3$: Beta 3 strand
- $\Delta E746-A750$: Deletion from amino acid 746 to 750
- $\Delta L747-E749$: Deletion from amino acid 747 to 749
- $\Delta S752-I759$: Deletion from amino acid 752 to 759
- $\Delta L747-A750$nsP: Deletion from amino acid 747 to 750 with an added proline
- $\Delta L747-E749$nsP: Deletion from amino acid 747 to 749 with an added proline
- HDMS: Hydrogen–deuterium exchange mass spectrometry
or solvent-accessible regions exchange with deuterium (present as D2O in the solvent) more rapidly than in structurally rigid regions, time-dependent deuterium uptake can be used to assess local backbone fl exibility. We compared wild type EGFR TKD with two file 1 variants (ΔL747-A750InsP and L747P) and two file 2 variants (ΔE746-A750 and ΔL747-P753InsS). In the absence of erlotinib, all mutated variants showed a generally increased level of backbone amide hydrogen exchange (Fig. 4a, b and Supplementary Figs. 6a, 7) across most regions of the protein. This effect was most notable in the region between helix αC and strand β7 (residues 750–850), which surrounds the binding site for ATP and ATP-competitive TKIs (Fig. 4b45) — notably in the αC/β4 l o o p and αE/HRD.

Fig. 4 | Native state dynamics of exon 19 variants differ for profile 1 and profile 2. a Backbone amide hydrogen exchange from HDX-MS experiments for all analyzable peptides at the 1 min labeling timepoint for wild type (grey circles), ΔL747-A750InsP (blue circles), L747P (blue diamonds), ΔE746-A750 (red circles), and ΔL747-P753InsS (red triangles) variants without ATP or inhibitor. Mean percent exchange (± SD) for three independent experiments across each of three separate protein preparations (two for ΔL747-A750InsP) is plotted against the median residue number of the peptide in wild type EGFR numbering. Secondary structure element positions are denoted at the top of the figure, with functionally important regions (β5/αD hinge, HRD motif, and activation loop) indicated. Wild type data are represented by two dashed grey lines corresponding to the range (± SD) for each point. The ΔE746-A750 and ΔL747-P753InsS variants both showed EX1 kinetics48 in the HRD and DFG motif regions, so these data are shown separately (see Fig. 6). Positions of peptides i through vi (see c) are denoted at the bottom of the graph. b 1 min percent exchange data for the indicated EGFR variants mapped onto a crystal structure of wild type EGFR kinase domain in its active conformation (PDB ID: 7KXZ), with ATP (sticks) placed based on its position in PDB ID: 3VJO45. The percent exchange of each amino acid was assigned using the DynamX software package (Waters) as described in Methods, and is color coded as shown on a scale from orange (100% exchange) to dark grey (no exchange). White regions in the binding pocket and elsewhere represent regions with EX1 exchange kinetics or with no coverage. Corresponding representations of 10 s and 10 min data are provided in Supplementary Fig. 7. c Mean percent exchange data (± SD) for key individual peptides i through vi, as marked in (a), as a function of the logarithm of deuterium labeling time are shown. See Supplementary Fig. 8 for additional peptides. Three independent HDX-MS experiments were performed on each of three separate protein preparations (two for ΔL747-A750InsP), with experimental parameters listed in Supplementary Table 1. Source data are provided as a Source data file.
motif region. Importantly, profile 2 variants (red in Fig. 4a and Supplementary Fig. 6a) generally showed higher degrees of exchange in these regions than profile 1 variants (blue in Fig. 4a and Supplementary Fig. 6a). Peptides around αC (residue 920) also showed some differences—particularly in ΔL747-P753InsS—unexpectedly given their significant distance from the ATP-binding site. Interestingly, parallel HDX-MS analysis of the less active ΔL747-E749F variant showed that it does not show the enhanced exchange seen for other exon 19 deletion variants (Supplementary Fig. 6c). This finding suggests that increased structural dynamics (as manifest in HDX-MS studies) may play an important role in activation of EGFR by exon 19 deletions.

In Fig. 4c, HDX data for six individual peptides (i–vi: locations also noted in Fig. 4a) are compared for ΔE746-A750 (red), ΔL747-A750InsP (blue), and wild type (grey). These results illustrate the greater level of exchange for ΔE746-A750 compared with ΔL747-A750InsP in peptides from the β1 strand (ii), αC/β4 loop (iii), αD (iv), αE (v), and the activation loop (vi) regions. The observed differences suggest significantly increased structural flexibility or disorder in the ATP/inhibitor binding site for the most common (ΔE746-A750) exon 19 deletion (profile 2) compared with wild type or the ΔL747-A750InsP (profile 1) variant. Interestingly, the most dramatic difference in exchange was seen for peptides encompassing residues 821–827 (v in Fig. 4c). This corresponds to the C-terminal half of αC in the EGFR TKD (see also Supplementary Fig. 8e, f). Helix αE makes direct contacts with the C-terminal part of helix αC in both active and inactive EGFR TKD structures (Fig. 5a), so would be expected to be altered when the stability or dynamics of αC are changed. As shown in Fig. 5a, the Y827 side chain (in helix αE) forms a predicted hydrogen bond with the backbone amide of D770 (close to the C-terminus of αC, and the site of activating exon 20 insertions). The R831 side chain (at the very end of αE) also interacts with the S768 main chain in αC (Fig. 5a). Helix αC changes that result from profile 2 exon 19 deletions between β3 and αC are presumably communicated to αE through these interactions, which interestingly were also recently implicated in differential gefitinib sensitivity in molecular dynamics studies8. The altered αE peptide also includes M825, the side chain of which contacts H835 in the conserved HRD motif at the beginning of the TKD’s catalytic loop—propagating αC changes to the catalytic loop. Changes in αE dynamics also appear to be propagated to αD, which plays a key role in stabilizing the ATP-binding site (Fig. 5a). In addition, the Q820 side chain adjacent to this peptide appears to help stabilize the β7/β8 hairpin that contributes to ATP binding. Other peptides that show the most elevated exchange in ΔE746-A750 also lie adjacent to the ATP-binding site (Fig. 5a), suggesting that their structural fluctuations seen by HDX-MS reflect disorder in this binding site caused by the deletion between β3 and αC. The profile 2 variant ΔL747-P753InsS also shares these dynamic properties with ΔE746-A750, whereas the L747P variant appears intermediate. Exchange plots are compared for selected peptides captured for all (or most) variants studied in Supplementary Fig. 8, and also illustrate the generally increased exchange seen in different regions of the TKD for profile 2 compared with profile 1 exon 19 variants.

When HDX-MS experiments were repeated in the presence of erlotinib, the differences outlined above were much less evident (Fig. 5b and Supplementary Fig. 6b, d), and the dynamics of the erlotinib-bound exon 19 variants were similar to those of erlotinib-bound wild type TKD. As expected, therefore, erlotinib binding causes structural rigidification of exon 19 mutants, and the conformational space sampled by the erlotinib-bound variants is narrowed so that it is no longer significantly different from that of the wild type TKD. One exception is ΔL747-P753InsS, which shows more exchange than the other erlotinib-bound variants in several regions (Fig. 5b), notably around αG.

**ATP-binding pockets of profile 2 variants unfold on minute timescales**

One of the most notable features of the HDX-MS data without erlotinib was a bimodal distribution in the exchange kinetics of a few key peptides from profile 2 variants (Fig. 6 and Supplementary Fig. 9). In typical HDX-MS studies of native proteins, the centroid mass of most peptides shows a progressive shift to higher m/Z values over time, in the so-called EX2 limit47–49. This reflects frequent but transient/very short-lived local unfolding or structural opening events that refold/close at a higher rate than that of hydrogen/deuterium exchange. By contrast, bimodal mass envelopes are seen in the EX1 limit50, where the locally unfolded state is much longer lived—reflecting a high energy barrier between the unfolded (exchange-competent) and folded (exchange-incompetent) states, indicating structural
heterogeneity. In this EX1 limit, all backbone hydrogens in the peptide either become exchanged simultaneously (high mass envelope) or remain unexchanged (low mass envelope). Bimodal mass envelopes of this sort were seen in both profile 2 variants (Fig. 6 and Supplementary Fig. 9), specifically in peptides from the catalytic loop (which includes the HRD motif) and from the activation loop (which includes the DFG motif). This phenomenon was either not seen, or was much less prominent, in the profile 1 variants studied. These data argue for significant local unfolding of regions around the ATP-binding pocket on the minute timescale – but only for profile 2 and not profile 1 exon 19 variants. The fused-bimodal envelopes seen for the DFG-motif region in ΔE746-A750 (Fig. 6b) and ΔL747-A750InsP (Supplementary Fig. 9b) indicate that this region does not fully display EX1 behavior, but again argues for significant disorder in the ATP/drug-binding pocket. Different regions within the same protein or assembly have been reported to display EX1 and EX2 kinetics in several other cases. Recent proposals based on where the exon 19 deletion begins, also supported by recent clinical data, argue that the proline insertion per se is not a relevant determinant factor for whether activated variants fall into profile 1 counterparts.

Exon 19 variant features that predict clinical response to erlotinib

The low frequencies of uncommon EGFR exon 19 variants make them challenging to study in statistically-powered clinical trials. It is therefore important to devise predictive methods to guide clinical decision making and trial design—by identifying which uncharacterized variants fall into profile 1 versus profile 2, for example. It is particularly notable that the (profile 1) ΔL747-A750lnSP variant shows erlotinib and osimertinib resistance, whereas the (profile 2) ΔL747-T751lnSP variant—which just one more β3/αC residue deleted—does not (Fig. 2a). The different TKI sensitivities of these two variants are also supported by recent clinical data, and argue that the proline insertion per se is not a relevant defining factor for whether activated variants are in profile 1 or profile 2. Recently suggested classifications based on where the exon 19 deletion begins are also not robust,
since we find variants with deletions starting at L747 in both profiles 1 and 2, for example.

The defining characteristic of EGFR exon 19 variants with wild-type-like $K_m$ values ($\Delta$L747-A750InsP, L747P, $\Delta$L747-E749) and wild type is that they all have just three or fewer residues missing from the $\beta$3/c loop (Fig. 1b). Our $\Delta$L747-E749 crystal structure (Fig. 3) shows how such a 3-residue deletion can be tolerated without major disruption (consistent with HDX-MS studies of this variant showing a rigid wild-type-like loop). Introducing a proline at position 747 (as in $\Delta$L747-A750InsP and $\Delta$L747P) seems to be sufficient to activate the kinase (Table 1), with some enhanced structural dynamics (Fig. 4 and Supplementary Fig. 6). Longer deletions—with or without introduction of a proline—require significant structural distortion. Consistent with this, profile 2 variants with longer deletions from the $\beta$3/c loop show increased amide hydrogen exchange in the C-terminal part of helix $\alpha$C (near the $\alpha$C/$\beta$4 loop—as shown in Supplementary Fig. 8a–c). This disruption is propagated to helices $\alpha$D, $\alpha$E (Supplementary Fig. 8d–f) and beyond—to cause substantial disorder in the ATP-binding site as well as TKD activation.

Defining profile 1 and profile 2 exon 19 variants as having $\beta$3/c loop deletions of ≤3 or ≥4 residues respectively, we analyzed deidentified outcome data from the GENIE+ database for patients with a diagnosis of non-small cell lung cancer whose tumors harbored an EGFR exon 19 mutation and were treated with erlotinib. Although the data available for this analysis were very limited, Kaplan–Meier survival analyses (Fig. 7a) showed that a profile 1 mutation is associated with a significantly shorter median progression-free survival (2.3 months) on erlotinib than a profile 2 mutation (8.5 months). Importantly, the association still holds (Fig. 7b) when we reanalyze the data excluding the common $\Delta$E746-A750 deletion (which we know is erlotinib sensitive), with median progression-free survival of 11.2 months ($P < 0.0001$). Predicted profile 1 mutations in the GENIE+ dataset were restricted to: $\Delta$L747-A750InsP (4 patients) plus $\Delta$E746-$\Delta$L747InsP (1 patient). Predicted profile 2 mutations were restricted to: $\Delta$E746-A750 (43), $\Delta$L747-T751 (3), $\Delta$E746-S752InsV (3), $\Delta$S752InsF (2), $\Delta$L747-S752 (2), $\Delta$E746-S752InsI (1), $\Delta$E746-T751InsA (1), $\Delta$L747-T751InsP (1), and $\Delta$E746-T751InsP (1). Similarly, among patients treated with erlotinib at the Yale Cancer Center, a profile 1 exon 19 mutation predicts a significantly shorter median median progression-free survival (2.6 months) compared to a profile 2 mutation (9.8 months) with $P = 0.0002$. After adjusting for baseline covariates of age, sex, race, and smoking history using the Cox proportional hazards model, the hazard for progression was nine times greater for patients with tumors with profile 1 compared to profile 2 mutations (95% confidence interval [2.6–31.4], $P = 0.0005$). This holds true for overall survival (OS) on erlotinib as well, with median OS of 24.5 months for profile 1 and 41.3 for profile 2 ($P = 0.03$). The adjusted hazard ratio for death was 3.0 (95% confidence interval [1.2–7.7], $P = 0.02$) for patients with tumors harboring profile 1 mutations compared to profile 2 mutations. Predicted profile 1 mutations in the Yale dataset were restricted to: $\Delta$L747-A750InsP (6 patients). Predicted profile 2 mutations were restricted to: $\Delta$E746-A750 (55), $\Delta$L747-T751InsS (5), $\Delta$L747-T751 (3), $\Delta$L747-S752InsV (1), $\Delta$L747-S752InsA (1), $\Delta$L747-T751InsP (1), $\Delta$L747-T751InsP (1), and $\Delta$E746-T751InsP (1). e Similarly, among patients treated with erlotinib at the Yale Cancer Center, a profile 1 exon 19 mutation predicts a significantly shorter median progression-free survival (2.3 months) on erlotinib than a profile 2 mutation (8.5 months). Importantly, the association still holds (Fig. 7b) when we reanalyze the data excluding the common $\Delta$E746-A750 deletion (which we know is erlotinib sensitive), with median progression-free survival of 11.2 months for profile 2. We also performed a similar analysis for a cohort of patients from Yale Cancer Center (Fig. 7c). From March 2005 to December 2021, six patients with tumors with profile 1 EGFR exon 19 mutations and seventy with profile 2 mutations were treated with erlotinib as the first TKI. Baseline characteristics between patients with tumors with profile 1 vs. profile 2 mutations were similar with respect to age (median 60.5 years [range, 50.3–67.8] for profile 1 compared to 62.8 years [36.9–94.5] for profile 2, $P = 0.70$), sex (83.3% female for profile 1 compared to 68.5% for profile 2, $P = 0.66$), and smoking status (66.7% with no prior smoking history for profile 1 compared to 42.8% in profile 2, $P = 0.39$). As with analysis of the GENIE+ data, a profile 1 deletion is associated with a significantly shorter PFS (1.8 months) on erlotinib than a profile 2 deletion (9.8 months). After adjusting for baseline covariates of age, sex, race, and smoking status, the risk of progression was higher for patients with tumors with profile 1 mutations compared to profile 2 (adjusted PFS hazard ratio 9.1 [95% confidence interval, 2.6–31.4], $P = 0.0005$). Profile 1 mutations are also associated with shorter overall survival on erlotinib (24.5 months) compared to profile 2 mutations (41.3 months), as well as a significantly increased risk of death, with an adjusted overall survival hazard ratio of 3.0 [95% confidence interval, 1.2–7.7], $P = 0.02$ (Fig. 7d).
In the absence of crystal structures, much of the available structural information about exon 19 variants has been inferred from molecular dynamics (MD) simulations. Significant discrepancies exist between the results of the different studies, however. In one such study, the ΔE746-A750 TKD was reported to occupy a predominantly αC-conformation and to bind ATP more tightly than the wild type EGFR—contrasting with our experimental findings. In another study, the ΔE746-A750 TKD was reported instead to have an αC-out conformation in the deepest energy minimum. Our own initial computational studies also suggested impaired erlotinib binding, again not supported by the experimental data presented here. MD simulation results rely heavily on accurate initial protein structures, and these previous analyses will not have taken into account the different degrees of structural disorder that the individual variants show in their native states. Thus, our HDX-MS results provide an important dynamic perspective for considering ATP and TKI binding.

From a clinical perspective, another key observation from our study is that profile 1 and profile 2 exon variants differ in their sensitivity to erlotinib and osimertinib, but not to afatinib (Figs. 1, 2a). This suggests that tumors driven by ΔE746-A750InsP, L747P or possibly other profile 1 exon 19 mutations, should respond better to afatinib than to first or third generation EGFR TKIs. Indeed, our previous cellular data suggest this for ΔE746-A750InsP, and published in vitro and clinical data support this finding for L747P ΔE746-A750. Why might profile 1 exon variants show diminished inhibition by erlotinib and osimertinib but not by afatinib? For erlotinib, the increased ATP-binding affinity of profile 1 variants simply demands that more inhibitor (~10-fold more) is required to achieve the same level of inhibition seen with a profile 2 variant like ΔE746-A750 that binds ATP more weakly. For osimertinib, a covalent inhibitor, profile 1 mutations might weaken initial (reversible) binding either directly and/or through increased ATP competition. This would reduce the lifetime of the initial encounter complex and thus reduce the efficiency with which osimertinib reacts with C797 in the active site. Alternatively (or in addition), the profile 1 mutations could alter the pose of osimertinib in the TKD’s ATP-binding site—as suggested for other EGFR mutations—so that the reaction between the drug and C797 is impaired. These are the same mechanisms proposed to allow osimertinib to spare wild type EGFR in favor of the mutated receptor. Reversible binding of afatinib to EGFR is substantially stronger than seen with osimertinib, which should allow it to react more efficiently with profile 1 variants—and thus retain more complete inhibition—regardless of the specific exon 19 mutation. Although further detailed mechanistic study would be necessary to ascribe the retained afatinib sensitivity of profile 1 variants to enhanced affinity or reactivity, our data argue strongly that afatinib may be a more effective option for tumors with this class of uncommon exon 19 mutations.

In summary, our biochemical and structural analyses suggest that a simple classification of exon 19 variants based on β3/αC deletion length (and proline introduction) can provide a much-needed rationale to inform clinical decision-making for patients with tumors driven by uncommon exon 19 deletions, which is supported by analysis of patient outcome data.

Methods

Plasmid construction

For intact EGFR expression in CHO cells, cDNA for full-length EGFR (residues 1–1210 in precursor protein numbering—including the 24-residue signal sequence) in pcDNA3.1− was used (NCBI NP_005219.2, UniProt Id: P00533). EGFR mutations (L858R, T790M/L858R, ΔE746-A750, ΔL747-P753InsS, ΔL747-A750InsP, ΔE746-S752insL, ΔL747-S752, ΔE746-T751insA, ΔL747-T751insP, ΔE746-T751, ΔL747-P753, ΔL747-E749, ΔS752-T759, L747P, ΔL747-T751insS, ΔL747-T751insQ, or ΔE746-S752insL) were introduced into the wild type vector using QuikChange® site-directed mutagenesis (Agilent Technologies). For
expressing the corresponding kinase domains, TKD sequences (spanning residues 696–1022 in precursor protein numbering or 672–998 in mature protein numbering) were amplified from the relevant EGFR cDNA in pcDNA3.1(−) using primers that introduced an N-terminal hexa-histidine (6x His) tag and Spe I/Xho I restriction sites for subcloning into pFastBac1. Recombinant baculoviruses were then generated using the Bac-to-Bac system (Invitrogen) for protein expression in Spodoptera frugiperda Sf9 cells (Expression Systems).

Protein expression and purification
Sf9 cells at 1.6 × 10^7 cells/ml were infected with recombinant baculovirus and cultured for 60–65 h at 27 °C. Cells were harvested by centrifugation at 2,200 g, and cell pellets resuspended in cold (4 °C) lysis buffer (20 mM Tris/His, pH 8.0, containing 500 mM NaCl, 5 mM 2-mercaptoethanol, and 10% w/v glycerol—supplemented with Roche protease inhibitor cocktail). Cells were then lysed using a microfluidizer (Microfluidics M-110P). The cell lysate was centrifuged at 75,000 × g for 1 h to remove cell debris and insoluble aggregates, and the resulting supernatant was filtered through a 0.45 μm filter before loading onto a 5 ml HiTrap Chelating Ni2+ affinity column (Cytiva). The column was equilibrated in buffer A (20 mM Tris/His, pH 7.8, containing 500 mM NaCl), and EGFR TKD was eluted with buffer B (buffer A supplemented with 200 mM imidazole) using a step gradient over 20 column volumes. Pooled fractions were passed through a 0.22 μm syringe filter and further purified at 4 °C by gel filtration, using a HiLoad 16/60 Superdex 200 column (Cytiva) equilibrated in buffer C (20 mM Tris/His, pH 7.5, containing 250 mM NaCl, and 250 mM KCl). The purity and molecular weights were assessed by SDS-PAGE (Supplementary Fig. 1) imaged using a Bio-Rad GelDoc-EZ imager, running Image Lab Version 5.2.1, and intact mass spectrometry—with typical yields of purified EGFR TKDs ranging from 1.5 to 3 mg per liter of Sf9 culture.

Steady-state kinase assays
All steady-state parameters for kinase activity and inhibition were determined using a quantitative fluorometric peptide assay that employed AssayQuant PhosphoSenS® peptide substrate #0001 (AQpeptide), containing a sulfonamido-oxine (Sox) fluorophore that shows chelation-enhanced fluorescence upon peptide phosphorylation28,29. The assay reaction (20 μl) contained 50 mM HEPES (pH 7.5), 0.01% Brij-35, 10 mM MgCl2, 0.5% w/v glycerol, 0.1 mg/ml BSA, 1 mM DTT, and varying concentrations of AQpeptide accumulation were monitored using a BioTek Synergy microplate reader attached to EGFR with ATP. Progress curves of phosphorylated peptide were initiated by adding 10 μM AQPeptide, and fluorescence intensity counts corresponding to phosphorylated AQPeptide were monitored using the BioTek Synergy 2 microplate reader. Initial velocities were determined as described above. IC50 values for each inhibitor were then determined by fitting the measured velocity (normalized to 100% in the absence of inhibitor) to the simple equation: Rate = 100/[1 + (I/Ki)] using GraphPad Prism.

Inhibition constants (Ki) for EGFR TKD variants were measured only for the reversible inhibitor erlotinib, and were determined by monitoring progress curves at different concentrations of inhibitor in the presence of 1 mM ATP. Purified TKD at an estimated final concentration of 50–100 nM (10 times greater for ΔL747-T751) was mixed with 0 to 40 μM inhibitor (erlotinib, afatinib, or osimertinib) at 1 mM ATP (and 10 mM MgCl2) for 30 min. Reactions were initiated by adding 10 μM AQPeptide, and fluorescence intensity counts corresponding to phosphorylated AQPeptide were monitored using the BioTek Synergy 2 microplate reader. Initial enzyme velocities were determined as described above. Osimertinib IC50 values were then determined for each [ATP], and Ki was estimated by fitting these IC50 values to the Cheng-Prusoff Equation16:

\[ IC_{50} = \frac{K_i}{1 + [ATP]/K_{M,ATP}} \]

EGFR TKD concentrations for the different variants were: 5 nM (L858R/T790M), 10 nM (ΔL747-T750-P), 50 nM (L858R), 100 nM (ΔL747-P, ΔL747-T753InsS, ΔL747-T751InsP, ΔL747-E749), 200 nM (ΔE746-T751InsA), 500 nM (ΔE746-A750), 1 μM (ΔL747-E749), or blank (no EGFR).

Mammalian cell lines and culture conditions
For routine cell culture, CHO cells (ATCC #CCL-61) were cultured in 100 mm culture dishes (Corning Falcon #353003) at 37 °C, in a humidified incubator with 5% CO2 in F-12 medium (Kaighn’s modification of Ham’s F-12 medium, ATCC #30-2004) supplemented with 10% fetal bovine serum ( Gibco #16140071) and penicillin (100 U/ml) and streptomycin (100 μg/ml) (Gibco #15140022). Cells were obtained from ATCC, and routinely tested for mycoplasma contamination (Lonza #LT-07-I18).

Transient transfection and inhibitor treatment of CHO cells
For transfection, CHO cells were plated at a density of 2.0 × 10^4 in 6-well plates (Corning Falcon #353046) in F-12K medium supplemented with 3% FBS and no antibiotics. The following day, FuGENE™transfection reagent was added to each well as provided by the manufacturer. Cells were then incubated for 3 days before harvest.

To convert relative fluorescence intensity readings to molar concentrations of phosphorylated AQpeptide, we fully phosphorylated AQPeptide at a range of different concentrations up to 200 μM by incubating for 1 h with 5 nM purified T790M/L858R variant (plus 1 mM ATP and 10 mM MgCl2) and plotted the fluorescence intensity counts of fully phosphorylated AQPeptide plotted against the known peptide concentration.

Michaelis-Menten constants (Km) of each EGFR fragment for ATP (Km,ATP) were determined using 20 μM AQPeptide and a range of different ATP concentrations (0.98–3 mM), with purified EGFR TKDs at the following concentrations: 5 nM (L858R/T790M), 10 nM (ΔL747-T750InsP), 50 nM (L858R), 100 nM (ΔL747-P, ΔL747-T753InsS, ΔL747-T751InsP, 200 nM (ΔE746-T751InsA), 500 nM (ΔE746-A750), 1 μM (ΔL747-E749), or blank (no EGFR).

\[ IC_{50} = \frac{K_i}{1 + [ATP]/K_{M,ATP}} \]

EGFR TKD concentrations for the different variants were: 5 nM (L858R/T790M), 10 nM (ΔL747-T750-P), 50 nM (L858R), 100 nM (ΔL747-P, ΔL747-T753InsS, ΔL747-T751InsP, ΔL747-E749), 200 nM (ΔE746-T751InsA), 500 nM (ΔE746-A750), 1 μM (ΔL747-E749), or blank (no EGFR).

Mammalian cell lines and culture conditions
For routine cell culture, CHO cells (ATCC #CCL-61) were cultured in 100 mm culture dishes (Corning Falcon #353003) at 37 °C, in a humidified incubator with 5% CO2 in F-12K medium (Kaighn’s modification of Ham’s F-12 medium, ATCC #30-2004) supplemented with 10% fetal bovine serum ( Gibco #16140071) and penicillin (100 U/ml) and streptomycin (100 μg/ml) (Gibco #15140022). Cells were obtained from ATCC, and routinely tested for mycoplasma contamination (Lonza #LT-07-I18).

Transient transfection and inhibitor treatment of CHO cells
For transfection, CHO cells were plated at a density of 2.0 × 10^4 in 6-well plates (Corning Falcon #353046) in F-12K medium supplemented with 3% FBS and no antibiotics. The following day, FuGENE™transfection reagent was added to each well as provided by the manufacturer. Cells were then incubated for 3 days before harvest.

\[ IC_{50} = \frac{K_i}{1 + [ATP]/K_{M,ATP}} \]
HD transfection reagent (Promega #PRE2311) was used to transfect each well with 0.7 μg of pcDNA3.1(−) containing the relevant EGFR variant. After 48 h, cells were starved using F-12K medium with no supplementation and treated after 72 h with erlotinib (SelleckChem #S7786), afatinib (SelleckChem #S1011) or osimertinib (SelleckChem #S7297) for 1 h at the noted concentrations. Inhibitors were made as 100× stocks in DMSO, and diluted to the final concentrations in starvation F-12K medium. After treatment, cells were placed on ice, medium aspirated, and 150 μl of ice-cold lysis buffer added per well. Lysis buffer was made from 10× stock (CST #9803) and supplemented with protease and phosphatase inhibitors (ThermoFisher #78440). Using a cell scraper (ThermoFisher #179693), lysates were harvested, vortexed, and centrifuged for 10 min at 10,000 × g at 4 °C. Protein concentrations in the clarified lysates were quantitated using a detergent-compatible protein assay (BioRad #5000111). Clarified cell lysates were combined with 4× Laemmli sample buffer (BioRad #1610747) containing a final concentration of 50 mM DTT, and heated for 4 min at 95 °C.

Quantitative immunoblotting
Samples corresponding to 20 μg of total protein in clarified lysates were subjected to SDS-PAGE using 4–20% polyacrylamide stain-free gels (Biolyte #4568096). Proteins were transferred onto 0.2 μm nitrocellulose membranes (BioRad #1620112). Membranes were blocked for 1 h using Intercept TBS blocking buffer (LICOR #972-60001), washed extensively with TBS-T (10 mM Tris-HCl pH 8.0, 150 mM NaCl, 0.1% w/v Tween20), and incubated with primary antibodies diluted in Intercept T20 antibody diluent (LICOR #972-65001) overnight at 4 °C on an orbital shaker. The following primary antibodies were used (at 1:2000 dilution): phospho-EGFR (pE749) pY1173 (Rabbit, CST #4407), total EGFR (Mouse, ThermoFisher #MS-665-P0), Grb2 (Rabbit, CST #3972). Note that the EGFR pY1173 antibody recognizes phosphorylated Y1179 in the PO0533 Uniprot sequence. After extensive washing with TBS-T, blots were incubated with goat anti-mouse IRDye® 680RD (LICOR #926-68070) and/or goat anti-rabbit IRDye® 800CW (LICOR #9926-32211) secondary antibodies at 1:20,000 dilutions for 1 h at room temperature. After final extensive washes with TBS-T, the 680RD total EGFR and 800CW pEGFR bands were simultaneously detected using a LICOR Odyssey® DLx imaging system. As in previous work15,26,29,30, Grb2 was used as a loading control because it is less abundant than housekeeping proteins such as actin, yielding more accurate normalization. For quantitation and plotting of bands, bands were simultaneously detected using a LICOR Odyssey® DLx imaging system. As in previous work68,69, Grb2 was used as a loading control because it is less abundant than housekeeping proteins such as actin, yielding more accurate normalization. For quantitation and plotting of bands, bands were simultaneously detected using a LICOR Odyssey® DLx imaging system.

Hydrogen–deuterium exchange mass spectrometry (HDX-MS)
Freshly purified EGFR exon 19 variant TKDs at 7.8–15.6 μM in 20 mM HEPES (pH 7.4), 100 mM NaCl, 130 mM NaCl, in the presence of 8 M urea-d4 (Cambridge Isotope Laboratories, Inc.). Quenched samples were flash-frozen in liquid N2, and stored at −80 °C until mass spectrometry analysis. Frozen samples were quickly thawed and manually injected onto an Enzyme BEH pepsi column (Waters) at 2 °C, and the deuterium-labeled sample was digested for 3 min (at a flow rate of 100 μl/min/water/0.1% formic acid). Peptic peptides were trapped and separated using an Acquity UPLC BEH C18, 130 Å, column (1.0 × 100 mm, 1.7 μm, Waters catalogue number 186003975) and Acquity UPLC BEH C18, 130 Å, column (1.0 × 100 mm, 1.7 μm, Waters catalogue number 186002346), respectively, using a linear gradient from 5% to 40% acetonitrile over 7 min, on an Acquity UPLC M-Class w/HDX-2 automation (Waters), running at 40 μl/min. UPLC solvents were water containing 0.1% (v/v) formic acid (A) and acetonitrile containing 0.1% (v/v) formic acid (B). MS’ data were acquired using a Synapt G2-Si mass spectrometer (Waters) using 0.4 s scan time in sensitivity TOF mode. A ramp collision energy of 5 to 10 V was used for low energy acquisition, and IS to 40 V for high energy acquisition with continuous lock mass (Leu-Enk) for mass accuracy correction. Other instrument parameters were: capillary voltage, 3 kV; cone voltage, 10 V; source offset, 80 V; source temperature, 90 °C; desolvation temperature, 150 °C; cone gas flow, 0 l/h; desolvation gas flow, 600 l/h; nebulizer gas, 6 bar; MS scan range, 300–2000 m/z.

HDX-MS data analysis
Peptides were sequenced using the ProteinLynx Global Server 3.03 (PLGS, Waters), and the deuterium uptake of each peptic peptide was determined using DynamX 3.0 (Waters), with minimum intensity set to 5000, minimum product per amino acid to 0.3, and maximum MH+ error set to 10 ppm. The deuterium uptake of all analyzed peptides presented in this study is the average uptake of three biological samples (two for ΔL747-E749 and ΔL747-A750insP) with each biological repeat performed in technical triplicate (each a separate labeling reaction). The percent exchange of each peptic peptide (%D) was calculated by the following equation:

\[ \%D = \frac{m - m_0}{m_f - m_0} \times 100 \]

where \( m \) is the centroid mass of a peptic peptide at time, \( t \); \( m_0 \) is the centroid mass of a peptic peptide without deuterium labeling; and \( m_f \) is the centroid mass of a peptic peptide for the fully-deuterated standard sample. For plotting exchange data against residue number in Fig. 4a, the percent exchange for each peptide was plotted against the median residue number of the peptide. Mass spectra of peptides with EX1 exchange kinetics were analyzed, and fit to a Gaussian distribution using Origin (OriginLab). All data were collected and analyzed according to consensus HDX-MS guidelines20, and are detailed the Source data file and Supplementary Table 1.

The DynamX software package (Waters) was used to map percent exchange data onto the structure of the EGFR TKD in Fig. 4b and Supplementary Fig. 7. In brief, for a residue within a given region, the percent exchange determined for the shortest peptide that includes it was assigned. For residues found in multiple overlapping peptides, the percent exchange value for a peptide that retains the overlapping region at its C-terminus is used.

Determination of ΔL747-E749 EGFR TKD crystal structure
Purified EGFR TKD (ΔL747-E749) was subjected to an additional anion exchange chromatography step21, concentrated to -6 mg/ml in 20 mM Tris pH 8.0, 150 mM NaCl and 2% glycerol. Crystals were obtained using the hanging-drop vapor diffusion method with a reservoir solution of 28% PEG 400, 0.1M HEPES pH 7.5, 0.2 M CaCl2 at 16 °C. Crystals appeared after two weeks and were cryo-protected in reservoir solution supplemented with 20% (w/v) glycerol and 5% (w/v) ethylene glycol prior to flash freezing in liquid nitrogen. X-ray
diffraction data were collected on the synchrotron beamline 24-ID-C of NE-CAT at the Advanced Photon Source. Data were processed using XDS\textsuperscript{2} Version 20200417, and scaled using SCALSA (Version 3.3.22) in the CCP4 program suite (Version 7.1)\textsuperscript{3}. The structure was solved by molecular replacement using Phaser\textsuperscript{4}. The active state EGFR TKD structure (PDB code 1M17\textsuperscript{5}) was used as the search model. Repeated cycles of manual building/rebuilding were performed using Coot\textsuperscript{6} and were alternated with refinements using Phenix (Version 1.18.2.3874)\textsuperscript{7} and the PDB-REDO web server\textsuperscript{8} (https://pdb-redo.eu). The final structure was validated with the MolProbity (version 4.02b-467) and wwPDB (version 2.26) servers. Structural figures were generated using PyMOL (http://www.pymol.org).

Clinical data analysis
De-identified patient data were obtained from the GENIE BPC non-small cell lung cancer v2.1 consortium database\textsuperscript{9}, for patients with a diagnosis of NSCLC whose tumors harbored an EGFR exon 19 deletion. Only patients who received erlotinib and had outcome data for progression free survival (PFS) on erlotinib were included. Data were also obtained retrospectively from a cohort of patients with metastatic NSCLC harboring an EGFR exon 19 deletion who were treated with erlotinib at the Yale Cancer Center. All patients were enrolled in Institutional Review Board-approved protocols at Yale University School of Medicine. Ethics oversight was provided by the Yale University Institutional Review Board, and consent was obtained for use of patient clinical data.

For these analyses, PFS was defined as the time from treatment initiation to time of clinically significant growth of existing lesions or new lesions on imaging, or death. Overall survival (OS) was defined as time from treatment initiation to death. Profile 1 variants were defined as having deletions from the β3/αC loop of three or fewer residues, whereas profile 2 variants were defined as having β3/αC loop deletions of four or more residues. Profile 1 and profile 2 variants were grouped and compared using Kaplan-Meier survival analyses, using the Mantel–Cox test. Censoring occurred at the time of last clinical follow-up (denoted by an upward tick on the Kaplan–Meier plot). The Cox proportional hazards model\textsuperscript{10} was used to generate adjusted hazard ratios with 95% confidence intervals for PFS and OS, adjusting for baseline covariates of age, sex, race, and smoking history.

Reporting summary
Further information on research design is available in the Nature Research Reporting Summary linked to this article.

Data availability
The refined coordinates for the ΔL747-E749 EGFR TKD crystal structure have been deposited into the Protein Data Bank, with accession code PDB 7TVD. PDB entries 1IM1, 7KXZ, and 3VJO were also used in generation of structural figures. HDX-MS mass spectrometry data have been deposited into the Protein Data Bank, with accession code PDB7TVD. PDB entries 1M17, 7KXZ, and 3VJO were used in generation of structural figures. HDX-MS mass spectrometry data have been deposited into the Protein Data Bank, with accession code PDB7TVD. PDB entries 1M17, 7KXZ, and 3VJO were used in generation of structural figures. HDX-MS mass spectrometry data have been deposited into the Protein Data Bank, with accession code PDB7TVD. PDB entries 1M17, 7KXZ, and 3VJO were used in generation of structural figures. HDX-MS mass spectrometry data have been deposited into the Protein Data Bank, with accession code PDB7TVD. PDB entries 1M17, 7KXZ, and 3VJO were used in generation of structural figures. HDX-MS mass spectrometry data have been deposited into the Protein Data Bank, with accession code PDB7TVD. PDB entries 1M17, 7KXZ, and 3VJO were used in generation of structural figures. HDX-MS mass spectrometry data have been deposited into the Protein Data Bank, with accession code PDB7TVD.

References
1. Lynch, T. J. et al. Activating mutations in the epidermal growth factor receptor underlying responsiveness of non-small-cell lung cancer to gefitinib. N. Engl. J. Med. 350, 2129–2139 (2004).
2. Sharma, S. V., Bell, D. W., Settleman, J. & Haber, D. A. Epidermal growth factor receptor mutations in lung cancer. Nat. Rev. Cancer 7, 169–181 (2007).
3. Tate, J. G. et al. COSMIC: The catalogue of somatic mutations in cancer. Nucleic Acids Res. 47, D941–D947 (2019).
4. Graham, R. P. et al. Worldwide frequency of commonly detected EGFR mutations: A large, international study. J. Thorac. Oncol. 14, 157–164 (2019).
5. Roskoski, R. Jr. Small molecule inhibitors targeting the EGFR/ErbB family of protein-tyrosine kinases in human cancers. Pharmacol. Res. 139, 395–411 (2019).
6. Tan, D. S. et al. The International Association for the Study of Lung Cancer consensus statement on optimizing management of EGFR mutation-positive non-small cell lung cancer. J. Thorac. Oncol. 11, 946–963 (2016).
7. 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).
8. Ramalingam, S. S. et al. Overall survival with osimertinib in untreated, EGFR-mutated advanced NSCLC. N. Engl. J. Med. 382, 41–50 (2020).
9. Chung, K. P. et al. Clinical outcomes in non-small cell lung cancers harboring different exon 19 deletions in EGFR. Clin. Cancer Res. 18, 3470–3477 (2012).
10. Lee, V. H. et al. Association of exon 19 and 21 EGFR mutation patterns with treatment outcome after first-line tyrosine kinase inhibitor in metastatic non-small-cell lung cancer. J. Thorac. Oncol. 8, 1148–1155 (2013).
11. Wang, Y. et al. Patients harboring uncommon EGFR exon 19 deletion-insertion mutations respond well to first-generation EGFR inhibitors and osimertinib upon acquisition of T790M. BMJ Cancer Res. 21, 1215 (2021).
12. Truini, A. et al. The EGFR exon 19 mutant L747–A750P exhibits distinct sensitivity to tyrosine kinase inhibitors in lung adenocarcinoma. Clin. Cancer Res. 25, 6382–6391 (2019).
13. Tian, Y. et al. Different subtypes of EGFR exon19 mutation can affect prognosis of patients with non-small cell lung adenocarcinoma. PLoS One 13, e0201682 (2018).
14. Su, J. et al. Molecular characteristics and clinical outcomes of EGFR exon 19 indel subtypes to EGFR TKIs in NSCLC patients. Oncotarget 8, 11246–11257 (2017).
15. Improma, G. et al. Uncommon frame-shift exon 19 EGFR mutations are sensitive to EGFR tyrosine kinase inhibitors in non-small cell lung carcinoma. Med. Oncol. 35, 28 (2018).
16. Zhao, C. et al. The impact of EGFR exon 19 deletion subtypes on clinical outcomes in non-small cell lung cancer. Transl. Lung Cancer Res. 9, 1149–1158 (2020).
17. Chen, Y. et al. A multicenter-retrospective study of non-small-cell lung carcinoma harboring uncommon epidermal growth factor receptor (EGFR) mutations: different subtypes of EGFR exon 19 deletion-insertions exhibit the clinical characteristics and prognosis of non-small cell lung carcinoma. Transl. Lung Cancer Res. 11, 238–249 (2022).
18. 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).
19. Peng, X. et al. Clinical impact of uncommon epidermal growth factor receptor exon 19 insertion-deletion variants on epidermal growth factor receptor-tyrosine kinase inhibitor efficacy in non-small-cell lung cancer. Eur. J. Cancer 141, 199–208 (2020).
20. Rossi, S. et al. Impact of exon 19 deletion subtypes in EGFR-mutant metastatic non-small-cell lung cancer treated with first-line tyr- osine kinase inhibitors. Clin. Lung Cancer 20, 82–87 (2019).
21. Wang, Y. T., Ning, W. W., Li, J. & Huang, J. A. Exon 19 L747P mutation presented as a primary resistance to EGFR-TKI: a case report. J. Thorac. Dis. 8, E542–E546 (2016).
22. Xu, C.-W. et al. Molecular characteristics and clinical outcomes of EGFR exon 19 C-helix deletion in non-small cell lung cancer and response to EGFR TKIs. Transl. Oncol. 13, 100791 (2020).

23. Xu, H. et al. Heterogeneous response to first-generation tyrosine kinase inhibitors in non-small-cell lung cancers with different EGFR exons 19 mutations. Target. Oncol. 15, 357–364 (2020).

24. Sutiman, N. et al. EGFR mutation subtypes influence survival outcomes following first-line gefitinib therapy in advanced Asian NSCLC patients. J. Thorac. Oncol. 12, 529–538 (2017).

25. Huang, J., Wang, Y., Zhai, Y. & Wang, J. Non-small cell lung cancer harboring a rare EGFR L747P mutation showing intrinsic resistance to both gefitinib and osimertinib (AZD9291): A case report. Thorac. Cancer 9, 745–749 (2018).

26. Yu, G. et al. EGFR mutation L747P led to gefitinib resistance and accelerated liver metastases in a Chinese patient with lung adenocarcinoma. Int. J. Clin. Exp. Pathol. 8, 8603–8606 (2015).

27. Liang, S. K., Ko, J. C., Yang, J. C. & Shih, J. Y. Afatinib is effective in the treatment of lung adenocarcinoma with uncommon EGFR p.L747P and p.L747S mutations. Lung Cancer 153, 103–109 (2019).

28. Luković, E., González-Vera, J. A. & Imperiali, B. Recognition-domain focused chemosensors: versatile and efficient reporters of protein kinase activity. J. Am. Chem. Soc. 130, 12821–12827 (2008).

29. Shults, M. D., Carrico-Moniz, D. & Imperiali, B. Optimal Sox-based fluorescent chemosensor design for serine/threonine protein kinases. Anal. Biochem. 352, 198–207 (2006).

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

31. 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).

32. 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).

33. Eck, M. J. & Yun, C. H. Structural and mechanistic underpinnings of the differential drug sensitivity of EGFR mutations in non-small cell lung cancer. Biochim. Biophys. Acta 1804, 559–566 (2010).

34. Cheng, Y.-C. & Prusoff, W. H. Relationship between the inhibition constant (I₅₀) and the concentration of inhibitor which causes 50 percent inhibition (I₉₀) of an enzymatic reaction. Biochem. Pharmacol. 22, 3099–3108 (1973).

35. Hoemmerber, M., Pittsawong, W. & Kern, D. Cumulative mechanism of several major imatinib-resistant mutations in Ab1 kinase. Proc. Natl Acad. Sci. USA 117, 19221–19227 (2020).

36. Bresler, S. C. et al. ALK mutations confer differential oncogenic activation and sensitivity to ALK inhibition therapy in neuroblastoma. Cancer Cell 26, 682–694 (2014).

37. Jackman, D. M. et al. Exon 19 deletion mutations of epidermal growth factor receptor are associated with prolonged survival in non-small cell lung cancer patients treated with gefitinib or erlotinib. Clin. Cancer Res. 12, 3908–3914 (2006).

38. Riely, G. J. et al. Clinical course of patients with non-small cell lung cancer and epidermal growth factor receptor exon 19 and exon 21 mutations treated with gefitinib or erlotinib. Clin. Cancer Res. 12, 839–844 (2006).

39. 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 123, 1137–1149 (2006).

40. 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. J. Biol. Chem. 277, 46265–46272 (2002).

41. Tunyasuvun kok, K. et al. Highly accurate protein structure prediction for the human proteome. Nature 596, 590–596 (2021).

42. Mirdita, M. et al. ColabFold - Making protein folding accessible to all. Nat. Methods 19, 679–682 (2022).

43. Engen, J. R. & Wales, T. E. Analytical aspects of hydrogen exchange mass spectrometry. Annu. Rev. Anal. Chem. 8, 127–148 (2015).

44. Hofnung, A. N., Resing, K. A. & Ahn, N. G. Practical methods for deuterium exchange/mass spectrometry. Methods Mol. Biol. 250, 283–298 (2004).

45. Yoshikawa, S. et al. Structural basis for the altered drug sensitivities of non-small cell lung cancer-associated mutants of human epidermal growth factor receptor. Oncogene 32, 27–38 (2013).

46. Li, M., Li, M., Xie, Y. & Guo, J. Uncovering the molecular basis for the better gefitinib sensitivity of EGFR with complex mutations over single rare mutation: Insights from molecular simulations. Molecules 27, 3844 (2022).

47. Percy, A. J., Rey, M., Burns, K. M. & Schriemer, D. C. Probing protein interactions with hydrogen/deuterium exchange and mass spectrometry—a review. Anal. Chim. Acta 721, 7–21 (2012).

48. Ferraro, D. M., Lazo, N. & Robertson, A. D. EX1 hydrogen exchange and protein folding. Biochemistry 43, 587–594 (2004).

49. Hofnung, A. N., Resing, K. A. & Ahn, N. G. Protein analysis by hydrogen exchange/mass spectrometry. Annu. Rev. Biophys. Biomol. Struct. 32, 1–25 (2003).

50. Roberts, J. M. et al. Dynamics of the T4c-family tyrosine kinase SH3 domains. Protein Sci. 25, 852–864 (2016).

51. Beresznzak, J. Z., Watts, N. R., Wingfield, P. T., Steven, A. C. & Heck, A. J. Assessment of differences in the conformational flexibility of hepatitis B virus core-antigen and e-antigen by hydrogen deuterium exchange-mass spectrometry. Protein Sci. 23, 884–896 (2014).

52. Fang, J., Engen, J. R. & Beuning, P. J. Escherichia coli processivity clamp β from DNA polymerase III is dynamic in solution. Biochemistry 50, 5958–5968 (2011).

53. Morgan, C. R. et al. Conformational transitions in the membrane scaffold protein of phospholipid bilayer nanodiscs. Mol. Cell. Proteom. 10, M111.010876 (2011).

54. Consortium, A. P. G. AACR Project GENIE: Powering precision medicine through an international consortium. Cancer Discov. 7, 818–831 (2017).

55. Grisztina, V. et al. The significance of epidermal growth factor receptor uncommon mutations in non-small cell lung cancer: A systematic review and critical appraisal. Cancer Treat. Rev. 85, 101994 (2020).

56. Mulliy, R. et al. Epidermal growth factor receptor mutants from human lung cancers exhibit enhanced catalytic activity and increased sensitivity to gefitinib. Cancer Res. 67, 2323–2330 (2007).

57. Bresler, S. C. et al. Differential inhibitor sensitivity of anaplastic lymphoma kinase variants found in neuroblastoma. Sci. Transl. Med. 3, 108ra114 (2011).

58. Foster, J. H. et al. Activity of crizotinib in patients with ALK-aberrant relapsed/refractory neuroblastoma: A Children’s Oncology Group Study (ADVL0912). Clin. Cancer Res. 27, 3543–3548 (2021).

59. Foster, S. A. et al. Activation mechanism of oncogenic deletion mutations in BRAF, EGFR, and HER2. Cancer Cell 29, 477–493 (2016).

60. Freed, D. M., Park, J. H., Radhakrishnan, R. & Lemmon, M. A. Deletion mutations keep kinase inhibitors in the loop. Cancer Cell 29, 423–425 (2016).

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

62. Galdidas, I. et al. Structural basis of the effect of activating mutations on the EGFR receptor. Elife 10, e65824 (2021).

63. Amirat, M. Z., Koivu, M., Elenius, K. & Johnson, M. S. Structural characterization of EGFR exon 19 deletion mutation using molecular dynamics simulation. PLoS One 14, e0222814 (2019).
64. Kim, S. Y. et al. Modeling clinical responses to targeted therapies by patient-derived organoids of advanced lung adenocarcinoma. Clin. Cancer Res. 27, 4397–4409 (2021).

65. Yang, G. et al. The lifted veil of uncommon EGFR mutation p.L747P in non-small cell lung cancer: Molecular feature and targeting sensitivity to tyrosine kinase inhibitors. Front. Oncol. 12, 843299 (2022).

66. Schwartz, P. A. et al. Covalent EGFR inhibitor analysis reveals importance of reversible interactions to potency and mechanisms of drug resistance. Proc. Natl. Acad. Sci. USA 111, 173–178 (2014).

67. Yan, X. E. et al. Structural basis of AZD9291 selectivity for EGFR T790M. J. Med. Chem. 63, 8502–8511 (2020).

68. Fey, B., Aksamitiene, E., Kiyatkin, A. & Khodolenko, B. N. Modeling of receptor tyrosine kinase signaling: Computational and experimental protocols. Methods Mol. Biol. 1636, 417–453 (2017).

69. Aksamitiene, E., Hoek, J. B. & Kiyatkin, A. Multistrip Western blotting: A tool for comparative quantitative analysis of multiple proteins. Methods Mol. Biol. 1312, 197–226 (2015).

70. Masson, G. R. et al. Recommendations for performing, interpreting and reporting deuterium exchange mass spectrometry (HDEx-MS) experiments. Nat. Methods 16, 595–602 (2019).

71. 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).

72. Kabsch, W. XDS. Acta Crystallogr. D. Biol. Crystallogr. 66, 125–132 (2010).

73. CCP4. The CCP4 suite: Programs for crystallography. Acta Crystallogr. D. Biol. Crystallogr. 50, 760–763 (1994).

74. McCoy, A. J. et al. Phaser crystallographic software. J. Appl. Cryst. 40, 658–667 (2007).

75. Emsley, P. & Cowtan, K. Coot: Model-building tools for molecular graphics. Acta Crystallogr. D. Biol. Crystallogr. 60, 2126–2123 (2004).

76. Afonine, P. V. et al. Towards automated crystallographic structure refinement with pheno.x.refine. Acta Crystallogr. D. Biol. Crystallogr. 68, 352–367 (2012).

77. Joosten, R. P., Long, F., Murshudov, G. N. & Perrakis, A. The PDB_REDO server for macromolecular structure model optimization. IUCrJ 1, 213–220 (2014).

78. Cox, D. R. Regression models and life-tables. J. Roy. Stat. Soc. B 34, 187–220 (1972).

79. Perez-Riverol, Y. et al. The PRIDE database resources in 2022: A hub for mass spectrometry-based proteomics evidences. Nucleic Acids Res. 50, D543–D552 (2022).

Acknowledgements

We thank members of the Lemmon, Ferguson, Politi laboratories and Prof. Kathryn Ferguson for comments on data and on the manuscript, Dr. Jin H. Park for work at the beginning of the project, and Dr. Joshua B. Sheetz for assistance from S.E.S. I.K.VAV.R., M.J.G., Z.W., and S.B.G. analyzed and interpreted clinical data with statistical advice from M.P. M.A.L., K.P., and S.B.G. supervised the overall project. D.M.L., I.K.VAV.R., and Y.T., and M.A.L. wrote the manuscript and all authors commented on it.

Competing interests

S.B.G. has received research funding from AstraZeneca, Boehringer Ingelheim, and Mirati. S.B.G. is also a consulting/advisory board member for AstraZeneca, Boehringer Ingelheim, Bristol-Myers Squibb, Genentech, Amgen, Blueprint Medicine, Sanofi Genzyme, Daiichi-Sankyo, Regeneron, Takeda, and Janssen. K.P. has received research funding from AstraZeneca, D2G Oncology, Roche/Genentech and Boehringer Ingelheim; and honoraria from Halda, Janssen, and AstraZeneca. K.P. is co-inventor on a patent for EGFR T790M mutation testing (WO2006086777A3) issued, licensed, from Memorial Sloan Kettering Cancer Center, for which she receives royalties. All remaining authors report no competing interests.

Additional information

Supplementary information The online version contains supplementary material available at https://doi.org/10.1038/s41467-022-34398-z.

Correspondence and requests for materials should be addressed to Mark A. Lemmon, Kumar D. Ash特kar or Yuko Tsuttsui.

Peer review information Nature Communications thanks John Burke, David Heppner and the other anonymous, reviewer(s) for their contribution to the peer review of this work.

Reprints and permissions information is available at http://www.nature.com/reprints

Publisher’s note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article’s Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article’s Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article’s Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article’s Creative Commons license, and indicate if changes were made. If material is not included in the article’s Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit http://creativecommons.org/licenses/by/4.0/.

© The Author(s) 2022