Deep mutational scanning reveals EGFR mutations conferring resistance to the 4th-generation EGFR tyrosine kinase inhibitor BLU-945

Check for updates

Yueyang Wang[1], Yuan Hao[2], Michela Ranieri[2], Tigran M. Abramyan[3], Lev Tsidilkovski[3], Michelle Hollenberg[1], Alfonso Lopez[2], Xavier T. R. Moore[1], Fiona Sherman[2], Jiehui Deng[2], Hayk Saribekyan[3], Garegin Papoian[3], Kwok-Kin Wong[2], Elaine Shum[2] & John T. Poirier[2] ✉

Fourth-generation EGFR tyrosine kinase are in development to overcome common resistance mutations. We performed deep mutational scanning (DMS) of the EGFR kinase domain in the context of L858R by introducing a saturation library of ~17,000 variants into Ba/F3 cells. DMS library-expressing cells were exposed to osimertinib or BLU-945 to identify escape mutations. L718X mutations were enriched across all conditions. BLU-945 specific mutations included K714R, K716T, L718V, T725M, K728E, K754E/N, N771S/T, T783I, Q791L/K, G863S, S895N, K929I, and M971L. A secondary DMS screen combining osimertinib and BLU-945, exclusively enriched for L718X mutations. Clinically, L718X mutations emerged in two patients treated with BLU-945. One patient with baseline EGFR L858R and L718Q mutations experienced early progression. Another with baseline *EGFR* L858R, T790M, and C797S acquired an L718V mutation at progression. This study demonstrate how comprehensive resistance profiling of targeted therapies can predict clinically relevant mutations and guide rational combinations to delay or prevent resistance.

Lung adenocarcinoma (LUAD) accounts for ~40% of primary lung cancer[1], with most patients presenting with locally advanced or metastatic disease[2,3]. Molecular profiling of LUAD has identified driver mutations in over 80% of cases[3], with mutations in epidermal growth factor receptor (EGFR) among the most common[4]. Incidence in patients with Asian ancestry is 40–55%, notably higher than the 5–15% frequency observed in Caucasians[5]. The majority of EGFR mutations occur in the kinase domain and manifest as short 3–7 amino acid deletions in exon 19 (ex19del) or as an L858R (p.Leu858Arg) missense mutation encoded in exon 21[6,7]. Less common mutations such as G719X (p.Gly719X) and small insertions in exon 20 account for the remainder[6]. These mutations cause constitutive activation of EGFR, leading to unchecked cell proliferation, resistance to apoptosis and the promotion of metastasis[8].

Targeting mutant EGFR with small molecule tyrosine kinase inhibitors (TKIs) has proven to be an effective therapeutic strategy in EGFR mutant LUAD, leading to multiple FDA approvals[9–12]. The first- generation TKIs gefitinib and erlotinib target the ATP-binding pocket of the kinase domain,

effectively inhibiting downstream signaling[13]. Both gefitinib and erlotinib demonstrated significant improvement in progression-free survival (PFS) compared to platinum doublet chemotherapy in patients with tumors harboring EGFR L858R and ex19del alterations[11]. Despite initial clinical responses, acquired resistance frequently developed within one year[14]. A dominant mechanism of resistance to first-generation EGFR TKIs is the T790M (p.Thr790Met) Gatekeeper mutation, which abolishes TKI binding while retaining kinase activity[15,16].

One current standard of care for treatment-naïve EGFR mutant LUAD is osimertinib, a third-generation TKI, that was developed to overcome the T790M Gatekeeper mutation[17]. Osimertinib covalently target C797, inhibiting EGFR ex19del and L858R irrespective of the presence of T790M while sparing wild type EGFR[18]. Despite initial efficacy, 7–26% of patients develop resistance to osimertinib by acquiring additional second-site mutations that are enriched in an allele specific manner[19,20]. EGFR L858R typically escapes osimertinib through L718Q/V (p.Leu718Gln/Val) mutations while ex19del tends

[1]Vilcek Institute of Graduate Biomedical Sciences, NYU Grossman School of Medicine, New York, NY, USA. [2]Laura and Isaac Perlmutter Cancer Center, NYU Langone Health, New York, NY, USA. [3]Deep Origin, South San Francisco, CA, USA. ✉e-mail: john.poirier@nyulangone.org

THE HORMEL INSTITUTE
UNIVERSITY OF MINNESOTA

to escape *via* C797S (p.Cys797Ser), which destroys the cystine nucleophile required for covalent binding of osimertinib[21].

A diverse range of resistance mechanisms not attributable to second site mutations have also been described. These include activation of bypass signaling pathways, such as MET amplification-mediated resistance, which occurs in 7–15% following first-line osimertinib therapy[22], as well as histologic transformation to small cell lung cancer (SCLC)[23,24]. Other identified mechanisms include EGFR amplification and mutations in downstream effectors such as *KRAS, BRAF*, and *PIK3CA*[19]. Notably, these aberrations can co-occur within the same tumor and co-exist with on target EGFR mutations, underlying the complexity and heterogeneity of tumor evolution and the emergence of resistance to EGFR-TKI therapy[19].

For C797S -trans T790M mutations, where C797S and T790M are encoded on different alleles, combinations of first- and third-generation TKIs have shown some efficacy in preclinical models[18]. Conversely, in C797S -cis T790M mutations, where C797S and T790M are on the same allele, combination of first-generation TKIs with osimertinib are ineffective[18]. These data suggest that suppression of allelic diversity within cells or tumor cell populations may be an important consideration in the development of novel EGFR TKIs.

Fourth-generation ATP competitive EGFR TKIs are designed to retain efficacy in the context of C797S[25–27] mutations; however, the landscape of potential second-site mutation liabilities associated with these inhibitors remains unknown[28,29]. BLU-945 is an investigational fourth-generation EGFR TKI designed to target EGFR T790M and EGFR T790M/C797S mutations that confer resistance to first-, second- and third-generation TKIs[30]. Unlike osimertinib which binds covalently to the C797 residue, the scaffold of BLU-945 allows it to adopt a conformation that fits within the ATP binding pocket of EGFR in the presence of the C797S mutation while sparing wild-type EGFR[30]. BLU-945 has been evaluated clinically as a monotherapy and in combination with osimertinib in previously treated patients with EGFR driven LUAD[31].

In this study, we aimed to identify potential mutational resistance mechanisms to BLU-945 using deep mutational scanning to assess the functional impact of thousands of EGFR mutations in parallel. We screened a DMS library of mutant cDNAs covering the EGFR L858R kinase domain to saturation in the context of a panel of EGFR TKIs to systematically profile all possible amino acid substitutions and their downstream effects.

Results

Deep mutational scanning of the EGFR kinase domain recovers known osimertinib resistance mutations

We generated a DMS library encompassing the EGFR L858R kinase domain encoded by a lentiviral transfer vector that co-expresses green fluorescent protein (GFP) *via* internal ribosomal entry site (IRES) (Fig. 1A). The library was generated using a mutagenic PCR approach, which relies on oligonucleotide primer pools encoding degenerate codons[32–34]. Each codon in the 804 bp target region can be mutated to one of 63 possible codons, including stop codons, resulting in a library of 16,884 potential variants (268 amino acids × 63 possible codons). The library consists of over 1×10^7 primary clones, which corresponds to >500× representation over the theoretical number of variants. Sanger sequencing of individual clones indicated an average of 1.14 mutations per cDNA (Fig. 1B). The plasmid library was sequenced using barcoded error correction and showed negligible mutational bias in transitions and transversions (Fig. 1C). 16,882 unique variants were recovered, corresponding to >99.9% saturation with >99% of variants observed >10 times (Fig. 1D). The abundance of codon variants was approximately normally distributed with a mean of 168 unique observations (Fig. 1E). The EGFR L858R DMS library was packaged in VSV-g pseudo-typed lentivirus at large scale and concentrated by centrifugation. The titer of the lentiviral library was 2×10^7 transduction units (TU) as determined by detection of GFP expression by flow cytometry. The total TU corresponded to ~1185× representation.

We conducted all screens in Ba/F3 cells, a murine pro-B cell line dependent on extracellular interleukin-3 (IL-3) for survival. Upon transduction with potent oncogenes, these cells become dependent on the introduced oncogene for mitogenic signaling[35]. Ba/F3 cells were transduced at a low multiplicity of infection (MOI ~ 0.1) to bias toward single integrations while achieving >1000-fold representation of the library[33,36]. After allowing 7 days for provirus integration and EGFR L858R mutant protein expression, IL-3 was withdrawn from the culture media. This established a selective pressure under which Ba/F3 cell survival depended on the signaling activity of transduced EGFR mutants[37]. Cells expressing viable EGFR variants were expected to survive without IL-3, while those harboring non-functional or deleterious mutations, or that were not transduced, were outcompeted and depleted from the population. After 14 days of IL-3 withdrawal, over 95% of the surviving Ba/F3 cells were GFP positive, consistent with positive selection of the transduced pool (Fig. 1F)[36,38,39]. The resulting EGFR L858R DMS library-expressing Ba/F3 cells were subsequently challenged with various EGFR inhibitors or DMSO for 14 days, with regular replenishment of compounds to maintain consistent drug pressure. At the end of the treatment period, gDNA was harvested and subjected to deep sequencing with error correction as previously described[36] to identify outgrowing clones. These enriched variants were interpreted as candidate resistance mutations (Fig. 1G).

Deep mutational scanning reveals EGFR mutations conferring resistance to BLU-945

To address whether known resistance mutations could be recovered from the previously introduced Ba/F3 DMS pool, we exposed the EGFR L858R DMS library-expressing Ba/F3 cells to osimertinib at 100 nM (>EC$_{90}$ of osimertinib) over a 14-day period with regular replenishment of drug to maintain consistent pressure Independent biological replicates demonstrated high concordance in the enriched variant profiles, supporting the reproducibility of the results. Among the osimertinib-treated samples, L718X (p. Leu718X) mutations were enriched compared to control (Supplementary 1A, B). As L718X is the dominant resistance mutation to osimertinib in the context of EGFR L858R oncogenic driver mutation, these results validate the validity of our screening approach.

We next employed the same screening workflow in the context of BLU-945 at 100 nM (>EC$_{90}$) for 14 days with regular drug replenishment. Endpoint samples were harvested, gDNA was extracted, and deep sequencing was performed as above. In contrast to the osimertinib screen, we identified numerous potential resistance mutations for BLU-945 (Fig. 2A). Candidate resistance mutations clustered in the hinge region amino acids Q791 and M793, which form hydrogen bonds with the adenosine ring of ATP[40]. Q791L/K and M793X mutations are supported as likely resistance mutations by structural evidence of a previously published co-crystal structure of the BLU-945 analog 'compound 24' and EGFR kinase domain (PDB:8D76)[30]. The amino-naphthyridine NH group of BLU-945 forms a hydrogen bond to the backbone carbonyl of Q791, while the naphthyridine N2 forms a hydrogen bond with the backbone NH of M793, stabilizing the inhibitor-kinase complex. We conclude that the results of the screen are consistent with the binding mode of BLU-945.

DMS screens explore all potential mutations at each amino acid position that are compatible with sustained EGFR L858R signaling, including complex codon changes that are less likely to occur in patients. We focused on amino acid changes that could arise from a single nucleotide substitution (Hamming distance = 1) from the wild-type codon. After filtering, Q791L/K mutations remain while M793X are removed, suggesting that mutations to M793 would be a less likely escape mechanism (Fig. 2B). We identified the following mutations as resistance mutations that can arise through single base mutations: K714R (p.Lys714Arg), K716T (p.Lys716Thr), L718V (p.Leu718Val), T725M (p.Thr725Met), K728E (p.Lys728Glu), K754E/N (p.Lys754Glu/Asn), N771S/T (p.Asn771Ser/Thr), T783I (p.Thr783Ile), Q791L/K (p.Gln791Leu/Lys), G863S (p.Gly863Ser), S895N (p.Ser895Asn), K929I (p.Lys929Ile), and M971L (p.Met971Leu). Several of these mutations have been observed clinically, such as K714R[41], L718V[42], T725M[43], K728E[44], T751P[45,46], K754E/N[47,48], and N771S[49].
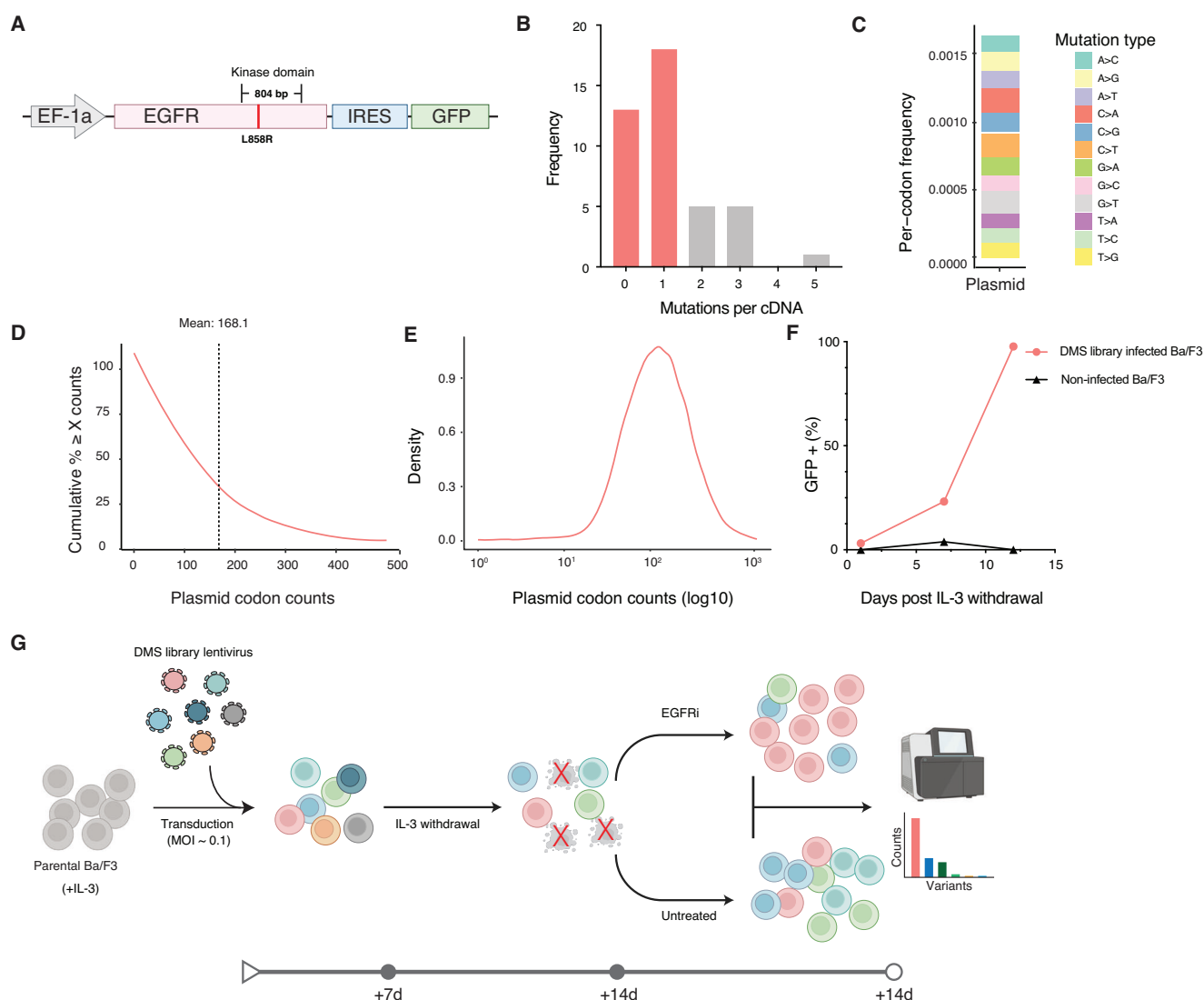
Fig. 1 | Deep mutational scanning of the EGFR L858R kinase domain in Ba/F3 cells. A Map of pHAGE-EGFR-L858R with kinase domain annotated. **B** Frequency of mutations count per cDNA. **C** Per-codon frequency of the indicated mutation types. **D** Cumulative distribution of read counts for all expected variants in the library. **E** Density plot of plasmid library codon counts. **F** Enrichment of library transduced Ba/F3 cells after IL-3 withdrawal. **G** Positive selection screening scheme.

When viewing BLU-945 mutations in the context of kinase domain structure, several patterns become apparent (Fig. 2C). First, mutations tend to be clustered in the N lobe and hinge regions of the kinase domain, which play a comparatively larger role in the catalytic activity of EGFR than the C lobe. Second, we observed frequent mutations in lysine residues in the β1-3 strands of the N lobe. Third, we see a cluster of mutations in the region between the αC-helix and β4-5 strands involving N771 and P772, two residues involved in ex20ins variants[50]. We suspect that these point mutations in this region might distort P-loop and αC-helix conformation in a manner that indirectly affects BLU-945 binding.

Many of the mutations highlighted here involve contact residues that are affinity determinants of BLU-945 (Fig. 2D). Notably, Q791 and M793 hinge contact points were present in compound 7, an early lead compound in the BLU-945 discovery campaign[30], while other mutations such as K716T and K728E involve residues engaged by the sulfone group characteristic of compound 24 and eventually BLU-945, developed late in the medicinal chemistry program. The large number of potential mutational escape routes for BLU-945 relative to osimertinib may be due to a greater number of required binding affinity determinants. Importantly, residues most critical for BLU945 binding, and therefore most vulnerable to resistance mutations, tend to cluster around the ATP binding pocket occupied by the inhibitor. These include contact residues such as K716, K728, L718, and Q791.

Not observed in the screen were mutations to the back pocket residue K745, which forms a hydrogen bond with the piperidinol ring of BLU-945[30]. Molecular interactions with K745 are a common and distinct feature of reversible, fourth-generation ATP-competitive EGFR kinase inhibitors[40]. We suspect that K745 mutations are not observed since this residue is the conserved "catalytic lysine" which is necessary to form a salt bridge with E762 in the kinase active state that is essential for phosphoryl transfer activity[51].

In addition to mutations of key contact point residues, DMS identified L718V as a key resistance mutation to BLU-945, like the osimertinib screen. Structurally, L718 is positioned in the first β-strand (β1) of the N-lobe near the ATP-binding site. Given the higher clinical prevalence of L718Q over L718V in patients with L858R-mutant EGFR, we sought to further characterize the structural impact of the L718Q variant. To this end, we conducted all-atom molecular dynamics (MD) simulations and absolute binding free energy perturbation (ABFEP) calculations, using structures modeled on PDB: 8D76. We hypothesized that isoform-specific dynamics
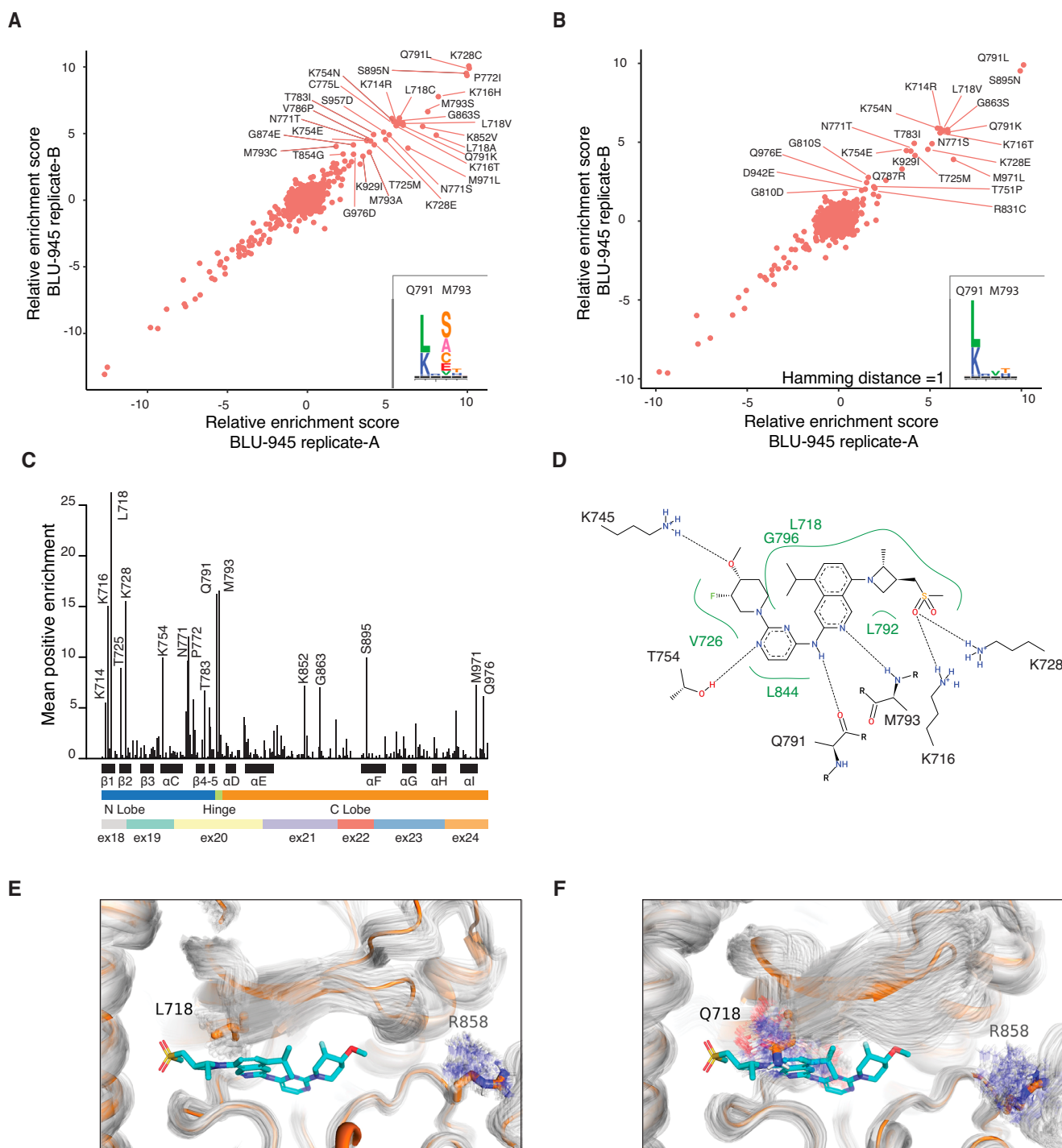
Fig. 2 | Identification of resistance mutations for BLU-945. A Scatter plot of enrichment scores after BLU-945 selection. Inset: hinge mutations involved in APT-binding. **B** Scatter plot of enrichment scores after BLU-945 selection, Hamming distance of 1. Inset: hinge mutations present after filtering. **C** Enrichment of BLU-945 resistance mutations relative to the EGFR kinase domain structure. **D** Poseview diagram of BLU-945, illustrating key contact residues. **E** L858R single mutant MD showed no steric clashes with BLU-945 overlayed from PDB: 8D76 (cyan sticks, edited from 8D76 ligand). Transparent gray shows all MD frames; orange highlights the cluster centroid. **F** L858R/L718Q double mutant MD frames sterically clash with BLU-945.

in the apo state modulate the architecture of the binding pocket and influence BLU-945 affinity.

Our 100 ns MD simulations showed that the L718Q mutation induces the greatest conformational perturbation in the binding site, particularly within the P-loop, as quantified by root-mean-square deviation (RMSD) analyses (Supplementary 1C, D). Overlay of the most populated MD conformers with the ligand-bound structure revealed a steric clash between the

Q718 side chain and BLU-945, suggesting impaired binding compatibility (Fig. 2E, F).

Functional validation of individual BLU-945 resistance mutations in multiple allelic contexts

To test whether any resistance mutations identified in the primary screen were sufficient to confer resistance to BLU-945, we chose to focus on the

Q791L hinge mutation given its importance in the primary screen for BLU-945 and the relative ease of mutation at that site as well as L718Q, since this mutation is most frequently identified as a resistance mutation in patients with L858R mutations[30,52]. We generated individual Ba/F3 cell lines stably expressing candidate resistance mutations in the context of L858R and exon 19 deletion (Del19) mutations (delE746-A750), which represent the most common EGFR-activating mutations in LUAD. Additionally, we introduced co-mutations of T790M and C797S to model clinically relevant resistance scenarios. Although T790M clones may be suppressed when osimertinib is used in the first-line setting, not all fourth-generation EGFR TKIs are designed to against this mutation (e.g., BDTX-1535). Therefore, we included T790M to ensure a more comprehensive evaluation of resistance mechanisms. All mutant constructs conferred growth factor independence in Ba/F3 cells apart from del19/L718Q, consistent with published reports in patients and mouse model systems suggesting that this co-mutation is non-functional[53,54].

We observed that Ba/F3 cells in L858R or del19 context, in conjunction with T790M/C797S mutations, displayed increased resistance to BLU-945 when either L718Q or Q791L mutations were present. The EC$_{50}$ values for L858R/T790M/C797S and del19/T790M/C797S were 1.71 nM (1.26 nM–2.40 nM 95%CI) and 7.81 nM (6.09 nM–9.96 nM 95%CI) respectively; however, the EC$_{50}$ was >100 nM with the addition of L718Q or Q791L resistance mutations. (Fig. 3A–D).

Western blot analysis revealed that the phosphorylation levels of EGFR in L858R/T790M/C797S and del19/T790M/C797S mutant cells were markedly reduced following treatment with 100 nM BLU-945. However, no significant reduction in phosphorylation was observed in cells harboring L718Q or Q791L cDNAs, indicating that these second-site mutations confer resistance to BLU-945 by escaping EGFR inhibition (Fig. 3E).

To further dissect the biophysical mechanisms underlying resistance, we performed ABFEP calculations using representative structures from MD simulations. Starting from the BLU-945 bound crystal structure (PDB: 8D76), we aligned and equilibrated ligand–protein complexes and calculated ΔG values for various EGFR isoforms. The predicted binding affinities closely mirrored experimental trends: L858R, L858R/T790M, and L858R/T790M/C797S all showed strong binding to BLU-945, consistent with low EC$_{50}$ values. L858R/C797S also retained favorable binding, though with a modest reduction.

In contrast, the L858R/T790M/Q791L mutant exhibited weaker binding affinity, with a ΔG shift consistent with the observed >2.5 kcal/mol drop in binding affinity. This change likely stems from a subtle entropic gain, as the hydrophobic leucine substitution at Q791 exposed to solvent and not directly involved in hydrogen bonding which may increase local flexibility relative to the more constrained wild-type hinge. The L858R/L718Q variant showed the most substantial loss in binding, with a calculated ΔG nearly 10 kcal/mol less favorable than L858R/T790M, reinforcing its role as a potent resistance mutation (Supplementary Fig. 1E; Table 1).

Osimertinib co-treatment suppresses BLU-945 kinase domain resistance mutants apart from L718X

In the first-in-human Phase I/II SYMPHONY trial, patients were treated with BLU-945 in combination with osimertinib[31]. To explore the potential of overcoming resistance to BLU-945 through combination therapy, we tested whether the resistance profiles of individual compounds, in this case BLU-945 and osimertinib, could predict the resistance profile for the combination treatment. Pairwise comparison of single agent resistance profiles suggested that osimertinib could effectively inhibit most resistance mutations observed with BLU-945 monotherapy, with the notable exception of mutations at position L718 (Fig. 4A).

To test our hypothesis, we performed a DMS experiment using BLU-945 and osimertinib in combination. The enrichment plot generated from these co-treatment endpoint samples indicated that resistance to both agents exclusively occurs at the L718 position where only proline (L718P) is enriched after filtering for Hamming distance >1 (Fig. 4B), illustrating a common resistance liability at this position.

While osimertinib and BLU-945 share a common resistance liability at position L718, we were encouraged that all other second-site mutations from the BLU-945 screen were effectively suppressed. We tested the resistance profiles of 9 distinct EGFR inhibitors to determine to what extent they shared resistance mechanisms with BLU-945 (Fig. 4C). Notably, BLU-945 suppressed T790M/C797S in both L858R and Del19 contexts. Mutations conferring resistance to BLU-945 such as Q791L were sensitive to all drugs tested in the context of either L858R or ex19del. However, rescue by osimertinib, sunvozertinib, lazertinib, or zipalertinib was all abolished by the C797S co-mutation. Examination of cross-resistance patterns of these drugs suggests that the combination of BLU-945 with one of the several EGFR TKIs with activity against Q791X or L718X might be an attractive strategy in contexts where the two mutations do not occur *in cis*.

BLU-945 and osimertinib cross-resistant L718X mutations identified in two patients with primary resistance and on-treatment disease progression

Our institution participated in the SYMPHONY clinical trial evaluating BLU-945, both as monotherapy and in combination with osimertinib, for patients with metastatic EGFR-mutant LUAD. We identified two patients harboring L718X EGFR mutations.

Patient A (Fig. 5A) harbored an EGFR L858R mutation and was previously treated with osimertinib. The patient had a liquid biopsy obtained on day –659 which identified the presence of *EGFR* L858R and L718Q mutations. Baseline imaging on day 0 revealed two nodules in the right lung, measuring 1.2 × 0.7 cm (middle lobe) and 2.7 × 2.2 cm (lower lobe). The patient started BLU-945 on day 12. Although initial follow-up imaging demonstrated stable disease, evaluations on day 37 and day 40 revealed disease progression with new bone metastases and spinal cord compression, necessitating surgical intervention. Post-surgery imaging on day 68 showed stable lung lesions, but further scans on day 121 demonstrated continued disease progression (+20% from nadir) along with new brain metastases. Additional progression in the spine was confirmed on day 177 and day 181.

Patient B (Fig. 5B) presented with *EGFR* L858R and developed on-treatment resistance mutations (*EGFR* T790M and C797S) following multiple prior treatments including osimertinib. A pre-treatment liquid biopsy performed on day-34 revealed *EGFR* L858R (34.7% VAF), T790M (13.7% VAF), and C797S (14.2% VAF). Baseline imaging on day 0 showed a right middle lobe lung nodule measuring 1.0 × 0.8 cm (TL1) as well as a liver lesion measuring 1.5 × 1.1 cm (TL2). BLU-945 therapy commenced on day 6. Follow-up imaging on day 33 demonstrated stable disease, with the sum of the target lesions decreasing by −1.7% overall from baseline. However, by day 61, the sum of target lesions increased by 3.3%. On day 117, imaging demonstrated progression due to new liver lesions, prompting a BLU-945 dose escalation to 400 mg.

On day 173, CT imaging indicated an overall growth of the target lesions by > 41%. Although the patient initially exhibited stable disease, therapy was ultimately discontinued due to rapid disease progression.

Four liquid biopsies were obtained from Patient A at day −659, −413, −51, and 180 relative to day 0 (Fig. 5C). During this interval, the *EGFR* L858R mutation rose from 15.8% to 80.0% VAF. Two distinct L718Q variants were also detected: L718Q (2151_2153GCT > ACA) was first measurable at 0.62% on day -51 and increased to 2.29% by day 180. L718Q (2153 T > A) was detected at 2.4% on day -659, fluctuating thereafter and ultimately reaching 12.11% on day 180.

Two on-treatment liquid biopsies were obtained from Patient B on days 119 and 147. The first confirmed *EGFR* L718V (0.15% VAF) in addition to L858R (39.69% VAF), T790M (17.19% VAF), and C797S (17.01% VAF) (Fig. 5D). Liquid biopsy on day 147 coincided with an increase in L718V to 0.2% VAF, with reduced levels of L858R (13.9% VAF), T790M (6% VAF), and C797S (6% VAF).
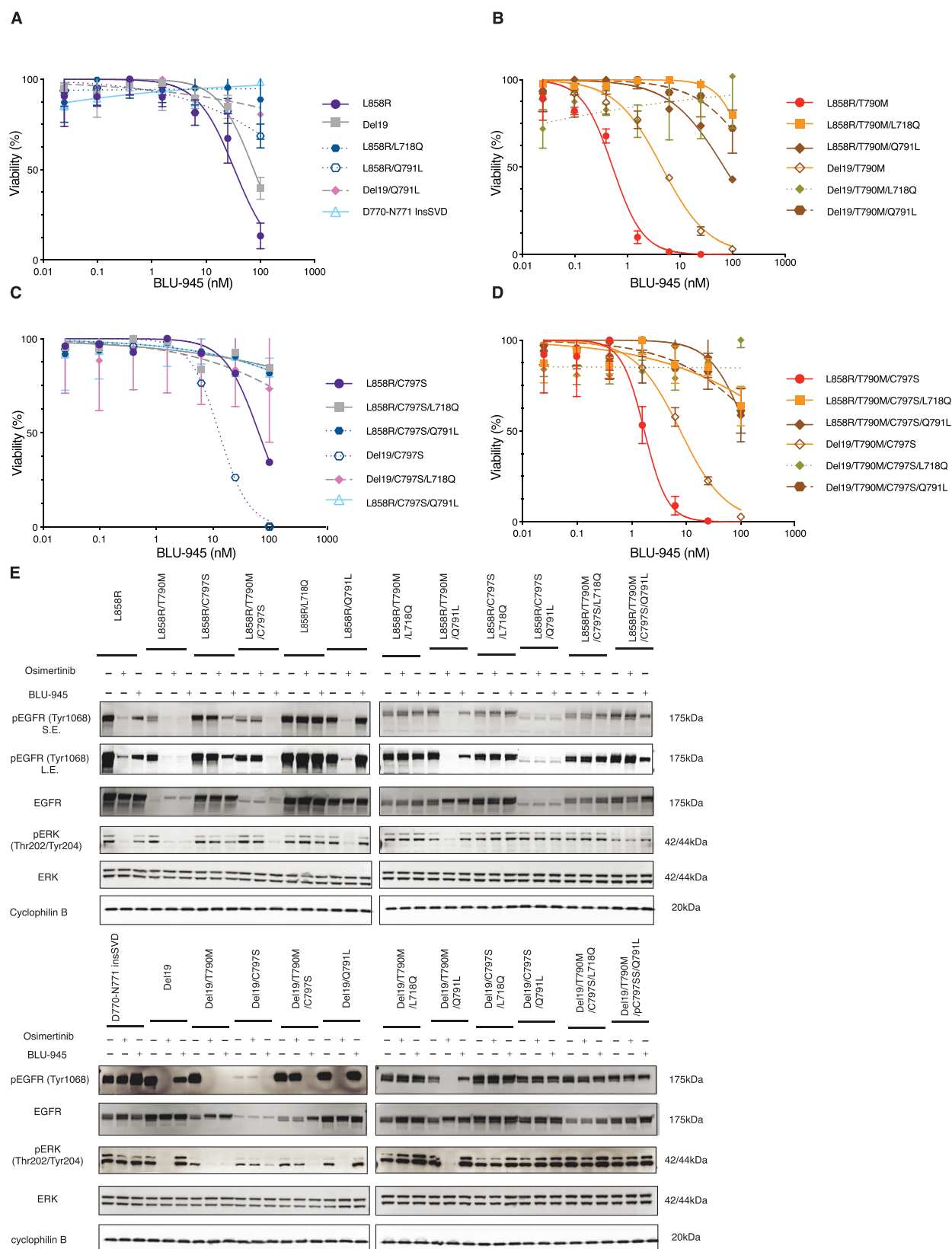
Fig. 3 | Testing of individual resistance mutations in the context of L858R, Del19, T790M, and C797S. A Viability dose-response curves showing the sensitivity of Ba/F3 cells expressing EGFR L858R or Del19 alone, as well as those harboring candidate resistance mutations (L718Q, Q791L), (D770-N771 InsSVD as a EGFR ex20ins control). **B** Viability dose-response curves of Ba/F3 cells expressing EGFR L858R/T790M or Del19/T790M, with or without co-occurring L718Q or Q791L mutations. **C** Viability dose-response curves of Ba/F3 cells expressing EGFR L858R/C797S or Del19/C797S, with or without co-occurring L718Q or Q791L mutations. **D** Viability dose-response curves of Ba/F3 cells harboring the triplet mutation combinations L858R/T790M/C797S or Del19/T790M/C797S, with or without L718Q or Q791L. **E** Western blot analysis of Ba/F3 cells under the indicated conditions. Both short exposure (S.E.) and long exposure (L.E.) images are shown to capture signal intensity differences.
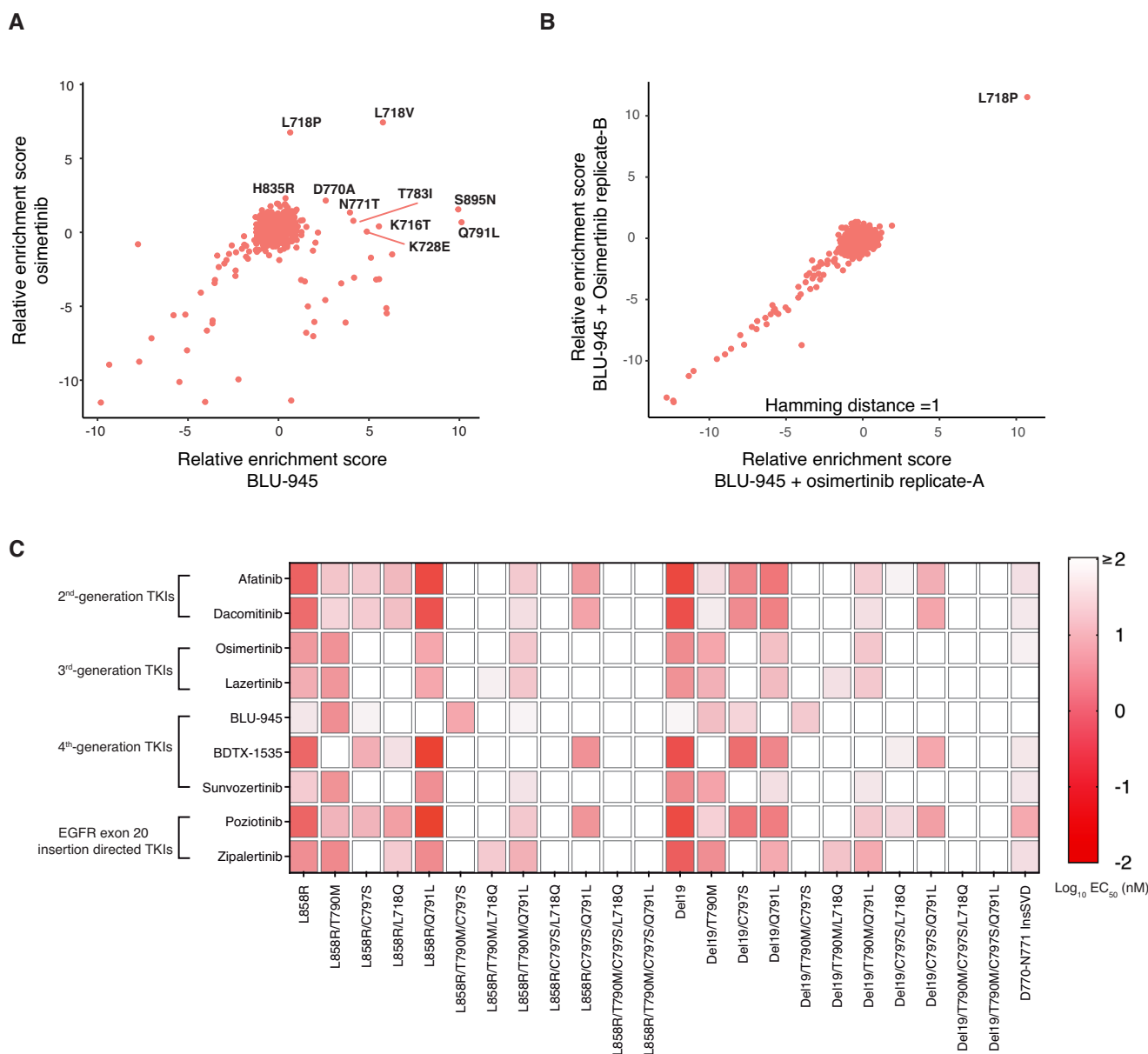
Fig. 4 | Prediction and experimental validation of EGFRi cross-resistance.
A Pairwise comparison of osimertinib and BLU-945 resistance profiles. **B** Scatter plot of enrichment scores after combination osimertinib and BLU-945 selection.

C Heatmap of Log_{10} EC_{50} values for 9 distinct inhibitors calculated from 7-point dose response curves.

Discussion

In this study, we employed deep mutational scanning to systematically map amino acid substitutions in the EGFR L858R kinase domain that confer resistance to BLU-945, a fourth-generation EGFR inhibitor. We validated Q791L and L718Q mutations, which emerged as crucial resistance determinants in L858R and ex19del contexts, irrespective of T790M or C797S co-mutations. L718X mutations are notable as a singular shared liability between BLU-945 and osimertinib, highlighting the challenge of covering all potential mutational escape routes even in the case of structurally distinct inhibitors.

DMS has the potential to provide comprehensive foreknowledge of on-target resistance liabilities not previously possible by methods such as long-term drug treatment or chemical mutagenesis. This approach can be integrated into early drug development pipelines to refine chemical scaffolds against known escape mutations. We show that forward genetics can complement structural information by highlighting constrained contact

points such as M793 and K745, while minimizing the vulnerability at sites like Q791 and L718. These positions in the EGFR kinase domain may be more "mutationally pliable" than others, enabling multiple potential substitutions that preserve kinase function and reduce inhibitor efficacy. Defining these mutationally permissive hotspots could support the development of inhibitors with increased robustness to mutational mechanisms of acquired resistance.

Anticipating drug resistance mutations early in drug development could also inform rational design of combination or sequential regimens in the clinic. For example, if preclinical data predict that L718X and Q791L/K will arise under selective pressure from BLU-945, treatment strategies could include pairing BLU-945 with a complementary inhibitor with orthogonal resistance liabilities to diminish the potential for outgrowth of resistant clones. From the perspective of clinical trial design, comprehensive resistance profiles could be used to further stratify patients to enhance response rates, increasing the probability of successful trial. This approach would also
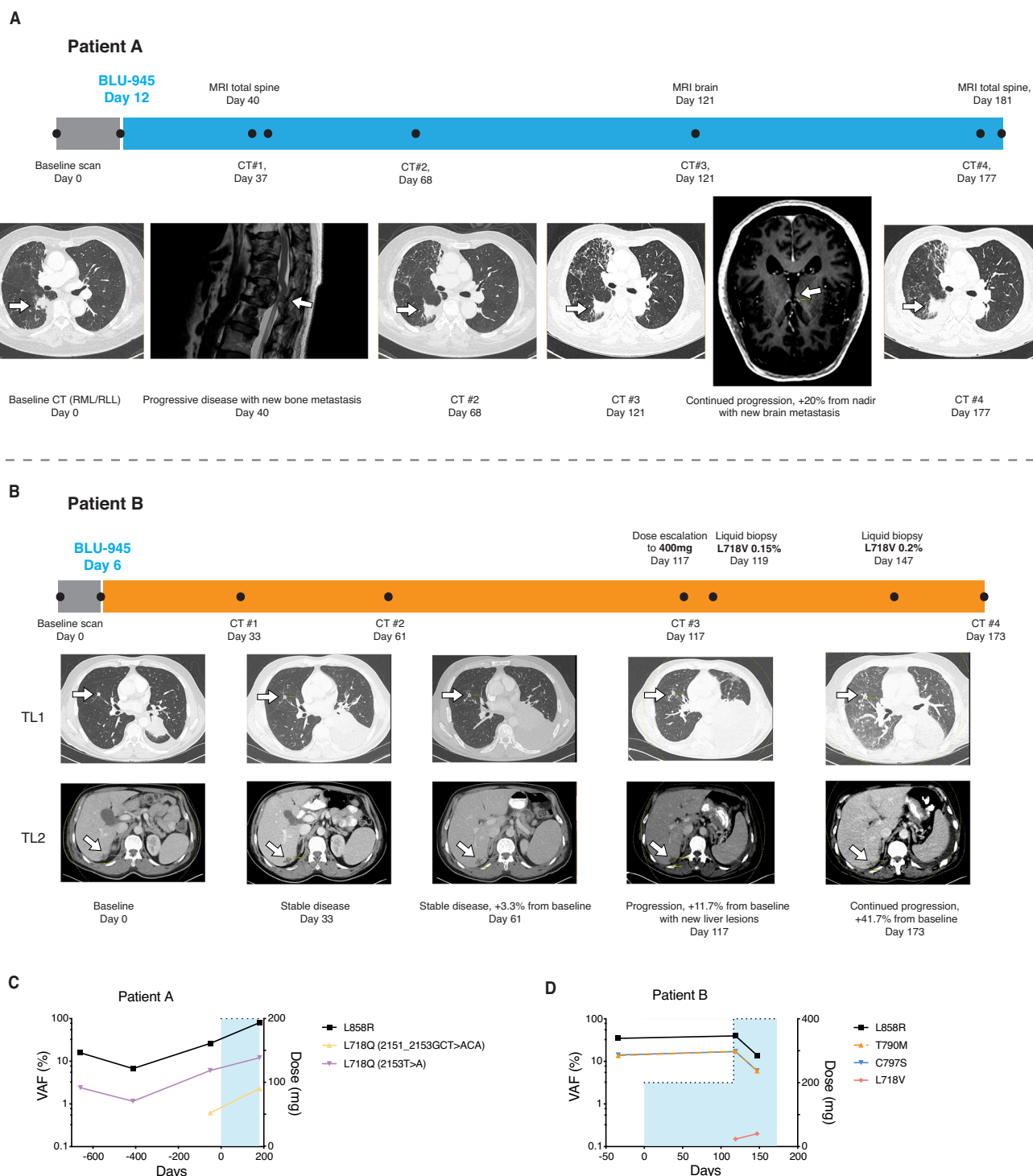
Fig. 5 | Clinical courses of two patients with L718X mutation. A Clinical course of patient A. B Clinical course of patient B. C ctDNA sequencing results for patient A. Left axis is $\log_{10}$ variant allele frequency, right axis is BLU-945 dose. D ctDNA sequencing results for patient B. Left axis is $\log_{10}$ variant allele frequency, right axis is BLU-945 dose.

respect ethical considerations by providing the necessary information to direct patients away from therapies that have reduced potential for efficacy. It is worth noting that while BLU-945 is no longer in active clinical development, this example highlights the value of such preclinical knowledge for guiding strategy of future drug candidates.

Our study is not without limitations. Our screen focuses entirely on the EGFR kinase domain in the context of a single L858R mutant allele. Mutations outside the kinase domain or that require different allelic context would be missed. On target mutations represent only one facet of clinical resistance, whereas real-world tumors can escape *via* other genetic processes

such as gene amplification, bypass signaling, or histological transformations[55]. The Ba/F3 system, while well-established for oncogene screening, may not fully capture the complexities of human lung adenocarcinoma, and in some cases, there was variability in total EGFR protein expression level between cell lines, limiting interpretation of pharmacodynamic experiments to comparisons that can be made within cell lines. Screening at a single high-dose condition (> EC_{90}) could potentially miss subtle resistance phenotypes that might emerge with lower drug concentrations. Finally, while our analysis tested candidate mutations in ex19del and L858R, our primary screen focused on the L858R context,

whereas exon ex19del, ex20ins, or comparatively less common point mutations such as G719X or L861Q might yield distinct allele-specific resistance mutations.

While the clinical development of BLU-945 has since been discontinued, our results overall illustrate the power of systematic mutational scanning to identify resistance liabilities for emerging EGFR TKIs. Future studies could build upon these findings by exploring combination strategies, building prospective surveillance of likely resistance mutations into early phase clinical trials, and designing novel EGFR TKIs that target common resistance escape routes.

Methods

Cell culture

HEK293T cells were cultured in high-glucose DMEM (Gibco #11995073) containing L-glutamine and sodium pyruvate, supplemented with 1% Penicillin-Streptomycin (Gibco #15140-122) and 10% Foundation™ fetal bovine serum (Gemini Bio-Products). Parental Ba/F3 cells were cultured in RPMI 1640 medium (Gibco #11875085) supplemented with 1% Penicillin-Streptomycin (Gibco #15140-122), 5% Foundation™ fetal bovine serum (Gemini Bio-Products), and 10 ng/mL recombinant mouse interleukin-3 (BioLegend #575504). Both HEK293T and Ba/F3 parental cell lines were obtained from the American Typed Culture Collection. All cell lines were maintained at 37 °C in a humidified incubator with 5% CO_2.

Plasmids

The plasmid pHAGE-EGFR-L858R was a gift from Gordon Mills & Kenneth Scott (Addgene plasmid #116276; http://n2t.net/addgene:116276; RRID:Addgene_116276). pHAGE-EGFR-L858R was used as the backbone for subsequent deep mutational scanning (DMS) experiments. Plasmids were transformed and propagated in NEB Stable Competent *E.coli* (New England Biolabs #C3040H). To prepare the plasmid template for DMS library construction, the plasmid was linearized using the restriction enzyme BSU36I (New England Biolabs #R0524S), which recognizes distinct sequences flanking the region of interest. The digestion reaction was prepared in a final volume of 50 μL, consisting of 1 μg of plasmid DNA, 5 μL of 10× rCutSmart Buffer (New England Biolabs #B6004S), 1 μL of BSU36I, and nuclease-free water (New England Biolabs #B1500S) to bring the reaction to final volume. The mixture was incubated at 37 °C for 2 h to ensure complete digestion. Following digestion, the reaction products were resolved on a 0.7% agarose gel to confirm successful cleavage. The linear fragment encompassing the EGFR kinase domain was excised and purified using a gel extraction kit.

Construction of EGFR L858R DMS library

To introduce mutations into the EGFR L858R kinase domain, we employed separate forward and reverse mutagenesis PCRs. Forward and reverse mutagenic primer were designed using the CodonTilingPrimers python script with minimum length set to 32 bp and synthesized as two separate oligonucleotide pools (Integrated DNA Technologies)[33]. In the forward reaction, 50 ng of linearized EGFR L858R template was combined with Q5 Master Mix (2×, New England Biolabs #M0491S), a 1 μL of a 10 μM forward oligo primer pool, and 1 μL of a 10 μM fixed reverse primer (5'-GCTCCAATAAATTCACTGCTTTG-3') in a total volume of 50 μL. The PCR protocol was carried out for 7 cycles using standard Q5 cycling conditions, and the resultant product was purified with a PCR cleanup kit (Qiagen #28506). The reverse mutagenesis PCR was set up similarly using a fixed forward primer (5'-GCATACAGTGCCACCCAGAG-3'). The PCR protocol was also carried out for 7 cycles using standard Q5 cycling conditions, and the resultant product was purified with the PCR cleanup kit (Qiagen 28506). Each of the purified forward and reverse products was then treated with Exonuclease I (NEB M0293S) to remove any residual mutagenic primers. DNA concentrations were adjusted to approximately 140 ng/μL, and each product was further diluted by adding 60 μL of nuclease-free water to reach a final volume of 90 μL.

The forward and reverse products were joined to produce the full-length, mutated insert. Each 30 μL joining PCR reaction consisted of 15 μL of Q5 Master Mix (2×), 4 μL of the diluted forward PCR product, 4 μL of the diluted reverse PCR product, 1 μL of a 10 μM fixed forward primer, 1 μL of 10 μM fixed reverse primer, and 5 μL of nuclease-free water. Six parallel joining PCRs were carried out to ensure sufficient product yield, using the same Q5 cycling conditions for 20 cycles. The products were then purified by PCR cleanup.

The mutagenized kinase domain was cloned into pHAGE-EGFR-L858R using restriction enzyme digestion and ligation. 1.6 μg of pHAGE-EGFR-L858R and 1 μg of the insert were digested with BSU36I and cleaned up by gel extraction. The digestion products were then ligated with T4 DNA ligase (New England Biolabs #M0202S) in a 20 μL reaction containing 166 ng of vector DNA, 100 ng of insert DNA, and 1 μL of T4 DNA ligase. The reaction was incubated at 16 °C overnight. A portion of the ligation mixture was tested by chemical transformation to confirm successful ligation prior to proceeding with large-scale transformation.

Electrotransformation was performed following the Endura™ Competent Cells (COMCEL-002) protocol. Briefly, 25 μL of electrocompetent cells were mixed with up to 5 μL of ligated DNA (final DNA concentration ~55 ng/μL) in each electroporation cuvette. The SOC buffer was warmed to 37 °C, and 5 mL culture tubes were prepared in advance. After electroporation, 970 μL of prewarmed SOC buffer was immediately added to each cuvette, and the cells were transferred to 5 mL tubes and recovered at 37 °C for 1 h with shaking at 250 rpm.

For library expansion, 1 mL of the recovered culture was plated on each Square BioAssay LB agar plate. To determine transformation efficiency by titration, ten-fold serial dilutions were prepared. Plates were incubated overnight at 37 °C, and the resulting colony counts were used to calculate transformation efficiencies and ensure sufficient library complexity.

After overnight incubation at 37 °C, the colonies on the large LB plate were harvested by gently scraping the plate surface with an appropriate volume of LB medium. The collected cells were transferred to a fresh tube, and plasmid DNA was extracted using a standard midi-prep protocol. This pooled plasmid DNA represented the final EGFR L858R DMS library. Individual colonies from the titration plates were selected for Sanger sequencing to determine the presence and frequency of intended mutations.

Lentivirus production

On Day 0, HEK293T cells were seeded in a 15 cm dish to reach ~70% confluence on Day 1. On Day 1, cells were transfected with 6 μg total DNA (EGFR L858R DMS library plasmid 3 μg, pSPAX 2 μg, pMD2.G 1 μg) in 200 μL Opti-MEM (Gibco #31985062), using a 2:1 ratio of linear polyethyleneimine (PEI) (Polysciences #23966) to total DNA. The DNA–PEI mixture was incubated at room temperature for 10 min before adding dropwise to the cells. On Day 2, the medium was refreshed. On Day 4, viral supernatants were collected, spun at 300 rcf for 5 min, and filtered through a 0.45 μm filter.

Genomic DNA extraction

Genomic DNA (gDNA) was extracted from ~1.5×10^7 cells using salt precipitation[56]. Cells were resuspended in 200 μL of DPBS in a 15 mL tube. Lysis was achieved by adding 3 mL of lysis buffer (10 mM Tris, 100 mM EDTA, 2% SDS, pH 8.0) containing RNase A (10 mg/mL), followed by vortexing and incubation at 37 °C for 5 min, then cooling on ice for 3 min. Protein contaminants were removed by adding 1 mL of protein precipitation solution (7.5 M ammonium acetate), vortexing, and centrifuging at $2000 \times g$ for 10 min.

The supernatant was transferred to a new tube containing 3 mL of isopropanol to precipitate DNA, and the mixture was gently inverted 50 times. DNA was pelleted by centrifugation at $2000 \times g$ for 3 min, washed with 3 mL of 70% ethanol, and air-dried for 5–10 min. The pellet was resuspended in 400 μL of hydration buffer (10 mM Tris, pH 8.0), incubated at 65 °C for 1 h, and hydrated overnight at room temperature with shaking at 600 rpm.

Barcoded Subamplicon Sequencing

cDNAs were sequenced using a two-step, barcoded subamplicon PCR strategy based on the method described as previously described[33,36]. In the first step (PCR1), the EGFR kinase domain was divided into two equal subamplicons, and each half was amplified in a separate reaction. Amplicon 1 used primers 5'-CTCTTGAGGATCTTGAAGGAAACTGAA-3' and 5'-CTTGACATGCTGCGGTGTTTT-3'. Amplicon 2 used primers 5'-CTTGACATGCTGCGGTGTTTT-3' and 5'-ATGCATTCTTTCATCCCCCT-GAATGAC-3'. Approximately 5 µg of the DMS library DNA was split between these two reactions, each containing primers encoding 8 degenerate nucleotides to introduce unique molecular identifiers (UMIs). Following amplification, PCR1 products were purified by isopropanol precipitation to remove residual primers and short fragments, and the concentration of each purified product was determined using the PicoGreen method (Invitrogen #P11496).

For the second PCR step (PCR2), 500 fg of the purified PCR1 product was used as the template for each subamplicon. This PCR step incorporated Illumina sequencing adapters and sample-specific indexes for multiplexed sequencing. The resulting indexed libraries were then purified using SPRI beads.

Libraries were pooled at equimolar concentrations and sequenced on an Illumina NovaSeq SP (500 cycles) platform using paired end reads configured to capture the UMI, barcode, and subamplicon regions.

System preparation and molecular dynamics simulations

All systems were prepared using the EGFR T790M_L858R_V948R mutant crystal structure (PDB: 8D76 chain A). Respective reverse or forward mutations were performed to design each isoform for simulations using open source PyMOL version 3.0.0. The bound ligand in the structure was edited to BLU-945. For holo isoform simulations, the ligand was removed. The resulting complex was solvated in a rectangular periodic box with a 1.2 nm buffer of explicit water molecules on all sides.

Simulations were carried out using the Deep Origin Simulation Suite. The protein was parameterized with the AMBER FF14SB force field[57], while the ligand was assigned parameters using the GAFF2 small molecule force field, with atomic charges derived via the AM1-BCC method[58]. Solvation employed the TIP3P water model[59] and a physiological ionic strength of 0.15 M was maintained by converting water molecules to Na^+ and Cl^- ions to achieve both electroneutrality and the desired salt concentration.

To enhance simulation efficiency, hydrogen mass repartitioning (HMR)[60] was applied by increasing the hydrogen mass to 2 Da, enabling the use of extended integration timesteps. All simulations were conducted in the NPT ensemble at 298.15 K and 1 bar, using a BAOAB Langevin integrator and a Monte Carlo barostat[61,62].

To assess structural dynamics and system equilibration, 100 ns molecular dynamics (MD) simulations were performed in duplicate for both the holo (ligand-bound) and apo (ligand-free) systems. The resulting trajectories were subjected to cluster analysis, and the two most populated conformations were selected for subsequent free energy perturbation (FEP) calculations. Prior to FEP, each of these cluster-derived structures was equilibrated with 10 ns of MD using a 4 fs timestep to enable the derivation of Boresch-style restraint parameters[63].

In-house Python scripts leveraging MDTraj v1.10.3[64] were used to analyze MD trajectories and compute root-mean-square deviations (RMSD). For each isoform, RMSD was calculated for the binding site backbone atoms, defined as all residues within 4 Å of BLU-945 in the reference structure (PDB: 8D76 chain A). The selected residues included: K716, L718, F723, V726, K728, A743, K745, E762, M766, C775, M790, Q791, L792, M793, P794, G796, L844, T854, D855, and D1006.

Free energy calculations

Absolute binding free energy perturbation (ABFEP) simulations were conducted using a standard ligand annihilation approach. For the binding leg of the calculations, 37 intermediate λ windows were employed, while 31 windows were used for the solvation leg. Each simulation was independently run in triplicate to ensure reproducibility and to estimate statistical uncertainty.

For the binding simulations, each window included 3 ns of equilibration followed by 5 ns of production dynamics, using a 4 fs timestep. In the solvation simulations, each window consisted of 0.2 ns of equilibration and 2.5 ns of production, run with a 5 fs timestep.

Free energy differences were estimated using the multistate Bennett acceptance ratio (MBAR) method. Convergence was assessed based on time-series stability of the free energy estimates, the standard deviation across the triplicate runs, and overlap matrices to confirm sufficient phase space sampling between adjacent windows.

Western blotting

Cells were lysed in RIPA buffer (Thermo Fisher #89900) supplemented with 1× Halt protease inhibitor cocktail (Thermo Fisher #78437) and phosphatase inhibitor cocktail (Thermo Fisher #78420). Total protein concentrations were determined using the BCA assay (Pierce). For each sample, 25 µg of total protein was mixed with 6× Laemmli SDS sample buffer (VWR, AAJ61337-AC) and heated at 95 °C for 5 min. Samples were then separated on 4–12% gradient Bis-Tris gel and transferred onto PVDF membranes.

Membranes were blocked in 5% nonfat milk or 3% BSA (for phospho-specific detection) in TBS-T (20 mM Tris-HCl, 150 mM NaCl, 0.1% Tween-20) for 1 h at room temperature. The blots were then incubated overnight at 4 °C with primary antibodies against total EGFR (Cell Signaling Technology #2232S), phospho-EGFR (CST #3777S), total ERK (CST #9102S), phospho-ERK (CST #9101S), and cyclophilin B (CST #43603S), following the manufacturers' recommended dilutions. After three washes in TBS-T, membranes were incubated with the HRP-conjugated secondary antibodies (Invitrogen #31460) at 1:10,000 for 1 h at room temperature. Following three additional TBS-T washes, protein bands were visualized using the chemiluminescence substrate (Thermo Fisher #34577) and detected by the Invitrogen iBright imaging system.

Cell viability assays

Ba/F3 cells were counted, adjusted to 20,000 cells/mL, and 25 µL (500 cells) was seeded into each well of a half-area 96-well plate (Corning #3688). Drug stocks were prepared at 2× concentrations using a 4-fold serial dilution (100 nM down to 0.0244 nM) in medium containing DMSO. Each drug concentration (including the DMSO-only control) was tested in triplicate by dispensing 25 µL of the 2× dilution into separate wells, yielding a final well volume of 50 µL at the 1× drug concentration. Plates were incubated at 37 °C in 5% CO_2 for 72 h, after which 25 µL of CellTiter-Glo 2.0 reagent (Promega #G9242) was added per well. Following a 10-minute incubation at room temperature with gentle shaking, luminescence endpoint was measured on a microplate reader. Viability values were normalized to the DMSO control and used to generate dose–response curves for IC_{50} determination using GraphPad Prism (Dotmatics).

Circulating tumor DNA sequencing

Peripheral blood was drawn at the NYU Perlmutter Cancer Center with the approval of the institutional review board (NYU IRB#S18-00980) and in accordance with the Declaration of Helsinki. Blood was collected in Streck Cell-Free DNA blood collection tubes, and samples were mixed gently per manufacturer's instructions, transported at 4 °C, processed to aliquot serum, and subsequently stored at −80 °C. Circulating tumor DNA sequencing was performed using the FoundationOne® Monitor (F1M) assay (Foundation Medicine, Inc.), a tissue-naïve method designed to monitor therapy response by quantifying ctDNA tumor fraction and variant allele frequency (VAF).

Data availability

Sequencing data that support the findings of this study is available at the Gene Expression Omnibus (GEO), a public functional genomics data repository (GSE305057).

References

1. Midha, A., Dearden, S. & McCormack, R. Mutation incidence in non-small-cell lung cancer of adenocarcinoma histology: a systematic review and global map by ethnicity (mutMapII). *Am. J. Cancer Res.* **5**, 2892–+ (2015).
2. Crino, L., Weder, W., van Meerbeeck, J. & Felip, E. Group EGW. Early stage and locally advanced (non-metastatic) non-small-cell lung cancer: ESMO Clinical Practice Guidelines for diagnosis, treatment and follow-up. *Ann. Oncol.* **21**, v103–v115 (2010).
3. Dearden, S., Stevens, J., Wu, Y. L. & Blowers, D. Mutation incidence and coincidence in non small-cell lung cancer: meta-analyses by ethnicity and histology (mutMap). *Ann. Oncol.* **24**, 2371–2376 (2013).
4. Smits, A. J. J. et al. EGFR and KRAS mutations in lung carcinomas in the Dutch population: increased mutation frequency in malignant pleural effusion of lung adenocarcinoma. *Cell Oncol.* **35**, 189–196 (2012).
5. Li, K. D., Yang, M. J., Liang, N. X. & Li, S. Q. Determining EGFR-TKI sensitivity of G719X and other uncommon EGFR mutations in non-small cell lung cancer: perplexity and solution (Review). *Oncol. Rep.* **37**, 1347–1358 (2017).
6. 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**, 167–179 (2020).
7. 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).
8. Uribe M. L., Marrocco I., Yarden Y. EGFR in cancer: signaling mechanisms, drugs, and acquired resistance. *Cancers*. 13. Epub 20210601. https://doi.org/10.3390/cancers13112748 (2021).
9. 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).
10. Paez, J. G. et al. EGFR mutations in lung cancer: correlation with clinical response to gefitinib therapy. *Science* **304**, 1497–1500 (2004).
11. Mok, T. S. et al. Gefitinib or carboplatin-paclitaxel in pulmonary adenocarcinoma. *N. Engl. J. Med.* **361**, 947–957 (2009).
12. Soria, J. C. et al. Investigators F. Osimertinib in untreated EGFR-mutated advanced non-small-cell lung cancer. *N. Engl. J. Med.* **378**, 113–125 (2018).
13. El-Rayes, B. F. & LoRusso, P. M. Targeting the epidermal growth factor receptor. *Br. J. Cancer* **91**, 418–424 (2004).
14. Wu S. G., Shih J. Y. Management of acquired resistance to EGFR TKI-targeted therapy in advanced non-small cell lung cancer. *Mol. Cancer*. **17**. https://doi.org/10.1186/s12943-018-0777-1 (2018).
15. 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**, 225–235 (2005).
16. Kobayashi, S. et al. EGFR mutation and resistance of non-small-cell lung cancer to gefitinib. *N. Engl. J. Med.* **352**, 786–792 (2005).
17. Mok, T. S. et al. Investigators A. Osimertinib or platinum-pemetrexed in EGFR T790M-positive lung cancer. *N. Engl. J. Med.* **376**, 629–640 (2017).
18. Fu, K., Xie, F. C., Wang, F. & Fu, L. W. Therapeutic strategies for EGFR-mutated non-small cell lung cancer patients with osimertinib resistance. *J. Hematol. Oncol.* **15**, 173 (2022).
19. Leonetti, A. et al. Resistance mechanisms to osimertinib in EGFR-mutated non-small cell lung cancer. *Br. J. Cancer* **121**, 725–737 (2019).
20. Enrico, D. et al. Overcoming resistance to osimertinib by T790M loss and C797S acquisition using gefitinib in a patient with EGFR-mutant NSCLC: a case report. *Jto Clin. Res Rep.* **4**, 100456 (2023).
21. Suzuki, M. et al. A macrocyclic kinase inhibitor overcomes triple resistant mutations in EGFR-positive lung cancer. *Npj Precis Oncol.* **8**, 46 (2024).
22. Qin K., Hong L., Zhang J., Le X. MET amplification as a resistance driver to TKI therapies in lung cancer: clinical challenges and opportunities. *Cancers*. 15. Epub 20230118. https://doi.org/10.3390/cancers15030612.
23. Leonetti, A. et al. Small cell lung cancer transformation as a resistance mechanism to osimertinib in epidermal growth factor receptor-mutated lung adenocarcinoma: case report and literature review. *Front. Oncol.* **11**, 642190 (2021).
24. Ferrer, L. et al. A brief report of transformation from NSCLC to SCLC: molecular and therapeutic characteristics. *J. Thorac. Oncol.* **14**, 130–134 (2019).
25. Gunther, M., Juchum, M., Kelter, G., Fiebig, H. & Laufer, S. Lung cancer: EGFR inhibitors with low nanomolar activity against a therapy-resistant L858R/T790M/C797S mutant. *Angew. Chem. Int. Ed. Engl.* **55**, 10890–10894 (2016).
26. Gunther, M. et al. Trisubstituted pyridinylimidazoles as potent inhibitors of the clinically resistant L858R/T790M/C797S EGFR mutant: targeting of both hydrophobic regions and the phosphate binding site. *J. Med. Chem.* **60**, 5613–5637 (2017).
27. Juchum, M. et al. Trisubstituted imidazoles with a rigidized hinge binding motif act as single digit nM inhibitors of clinically relevant EGFR L858R/T790M and L858R/T790M/C797S mutants: an example of target hopping. *J. Med. Chem.* **60**, 4636–4656 (2017).
28. Patel, H., Pawara, R., Ansari, A. & Surana, S. Recent updates on third generation EGFR inhibitors and emergence of fourth generation EGFR inhibitors to combat C797S resistance. *Eur. J. Med. Chem.* **142**, 32–47 (2017).
29. Fukuda, S. et al. Potential utility of a 4th-generation EGFR-TKI and exploration of resistance mechanisms-an in vitro study. *Biomedicines* **12**, 1412 (2024).
30. 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).
31. Elamin Y. Y. et al. BLU-945 monotherapy and in combination with osimertinib (OSI) in previously treated patients with advanced *EGFR*-mutant (*EGFRm*) NSCLC in the phase 1/2 SYMPHONY study. *J. Clin. Oncol*. **41**, https://doi.org/10.1200/JCO.2023.41.16_suppl.901146 (2023).
32. Doud, M. B., Hensley, S. E. & Bloom, J. D. Complete mapping of viral escape from neutralizing antibodies. *PLoS Pathog.* **13**, e1006271 (2017).
33. Dingens, A. S., Haddox, H. K., Overbaugh, J. & Bloom, J. D. Comprehensive mapping of HIV-1 escape from a broadly neutralizing antibody. *Cell Host Microbe* **21**, 777–+ (2017).
34. Doud, M. B., Lee, J. M. & Bloom, J. D. How single mutations affect viral escape from broad and narrow antibodies to H1 influenza hemagglutinin. *Nat. Commun.* **9**, 1386 (2018).
35. Palacios, R. & Steinmetz, M. Il-3-dependent mouse clones that express B-220 surface antigen, contain Ig genes in germ-line configuration, and generate B lymphocytes in vivo. *Cell* **41**, 727–734 (1985).
36. Doud, M. B. & Bloom, J. D. Accurate measurement of the effects of all amino-acid mutations on influenza hemagglutinin. *Viruses Basel.* **8**, 155 (2016).
37. Koga, T., Suda, K. & Mitsudomi, T. Utility of the Ba/F3 cell system for exploring on-target mechanisms of resistance to targeted therapies for lung cancer. *Cancer Sci.* **113**, 815–827 (2022).
38. Wu, N. C. et al. High-throughput profiling of influenza A virus hemagglutinin gene at single-nucleotide resolution. *Sci. Rep.* **4**, 4942 (2014).
39. Zhang, T. H., Wu, N. C. & Sun, R. A benchmark study on error-correction by read-pairing and tag-clustering in amplicon-based deep sequencing. *BMC Genomics.* **17**, 108 (2016).

40. Damghani, T. et al. Structural elements that enable specificity for mutant EGFR kinase domains with next-generation small-molecule inhibitors. *Methods Enzymol.* **685**, 171–198 (2023).

41. Vu, H. A. et al. Spectrum of EGFR gene mutations in Vietnamese patients with non-small cell lung cancer. *Asia Pac. J. Clin. Oncol.* **12**, 86–90 (2016).

42. Liu, Y. et al. Acquired EGFR L718V mutation mediates resistance to osimertinib in non-small cell lung cancer but retains sensitivity to afatinib. *Lung Cancer* **118**, 1–5 (2018).

43. Sharma, P. et al. A Novel EGFR germline mutation in lung adenocarcinoma: case report and literature review. *Clin. Lung Cancer* **25**, 479–482 (2024).

44. Mao, L. et al. Mutation spectrum of EGFR From 21,324 Chinese patients with non-small cell lung cancer (NSCLC) successfully tested by multiple methods in a CAP-accredited laboratory. *Pathol. Oncol. Res.* **27**, 602726 (2021).

45. Si, J. F., Gu, X. D., Wang, W. X., Ying, S. P. & Song, Z. B. Clinical outcomes of lung adenocarcinoma patients harboring uncommon epidermal growth factor receptor () mutations treated with EGFR-tyrosine kinase inhibitors (TKIs). *Ann. Palliat. Med.* **11**, 1624–1634 (2022).

46. Ruiz R. et al. Mutation pattern of EGFR gene in Peruvian patients with non-small cell lung cancer. *J. Clin. Oncol.* **42**, https://doi.org/10.1200/JCO.2024.42.16_suppl.e20615 (2024).

47. Ma, C. H. et al. Efficacy of afatinib for pulmonary adenocarcinoma with leptomeningeal metastases harboring an epidermal growth factor receptor complex mutation (exon 19del + K754E). A case report. *Medicine* **99**, e22851 (2020).

48. Miyamae, Y. et al. Mutation detection of epidermal growth factor receptor and genes using the smart amplification process version 2 from formalin-fixed, paraffin-embedded lung cancer tissue. *J. Mol. Diagn.* **12**, 257–264 (2010).

49. Tang, H. R. et al. EGFR mutations in small cell lung cancer (SCLC): genetic heterogeneity and prognostic impact. *J. Thorac. Oncol.* **12**, S710–S711 (2017).

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

51. Carrera, A. C., Alexandrov, K. & Roberts, T. M. The conserved lysine of the catalytic domain of protein kinases is actively involved in the phosphotransfer reaction and not required for anchoring ATP. *Proc. Natl. Acad. Sci. USA* **90**, 442–446 (1993).

52. Rotow, J. K. et al. Real-world genomic profile of EGFR second-site mutations and other osimertinib resistance mechanisms and clinical landscape of NSCLC post-osimertinib. *J. Thorac. Oncol.* **19**, 227–239 (2024).

53. Li, M. et al. L718Q/V mutation in exon 18 of EGFR mediates resistance to osimertinib: clinical features and treatment. *Discov. Oncol.* **13**, 72 (2022).

54. Starrett, J. H. et al. Drug sensitivity and allele specificity of first-line osimertinib resistance mutations. *Cancer Res.* **80**, 2017–2030 (2020).

55. Morgillo, F., Della Corte, C. M., Fasano, M. & Ciardiello, F. Mechanisms of resistance to EGFR-targeted drugs: lung cancer. *Esmo Open.* **1**, e000060 (2016).

56. Schneider, W. M. et al. Genome-scale identification of SARS-CoV-2 and pan-coronavirus host factor networks. *Cell* **184**, 120–32.e14 (2021).

57. Maier, J. A. et al. ff14SB: improving the accuracy of protein side chain and backbone parameters from ff99SB. *J. Chem. Theory Comput.* **11**, 3696–3713 (2015).

58. Jakalian, A., Jack, D. B. & Bayly, C. I. Fast, efficient generation of high-quality atomic charges. AM1-BCC model: II. Parameterization and validation. *J. Comput Chem.* **23**, 1623–1641 (2002).

59. Jorgensen, W. L., Chandrasekhar, J., Madura, J. D., Impey, R. W. & Klein, M. L. Comparison of simple potential functions for simulating liquid water. *J. Chem. Phys.* **79**, 926–935 (1983).

60. Hopkins, C. W., Le Grand, S., Walker, R. C. & Roitberg, A. E. Long-time-step molecular dynamics through hydrogen mass repartitioning. *J. Chem. Theory Comput.* **11**, 1864–1874 (2015).

61. Leimkuhler B., Matthews C. Rational Construction of Stochastic Numerical Methods for Molecular Sampling. *Appl. Math. Res. Expres*. 34-56. https://doi.org/10.1093/amrx/abs010. (2013).

62. Åqvist, J., Wennerström, P., Nervall, M., Bjelic, S. & Brandsdal, B. O. Molecular dynamics simulations of water and biomolecules with a Monte Carlo constant pressure algorithm. *Chem. Phys. Lett.* **384**, 288–294 (2004).

63. Boresch, S., Tettinger, F., Leitgeb, M. & Karplus, M. Absolute binding free energies: a quantitative approach for their calculation. *J. Phys. Chem. B.* **107**, 9535–9551 (2003).

64. McGibbon, R. T. et al. MDTraj: a modern open library for the analysis of molecular dynamics trajectories. *Biophys. J.* **109**, 1528–1532 (2015).

Competing interests
K.K.W. and J.T.P. have received research funding from Blueprint Medicines. The authors declare no non-financial interest.

Additional information
Supplementary information The online version contains supplementary material available at https://doi.org/10.1038/s41698-025-01086-2.

Correspondence and requests for materials should be addressed to John T. Poirier.

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.