SUPPLEMENTARY METHODS

Genomic ancestry prediction in Foundation Medicine cohort

To generate a genomic ancestry prediction for each patient in cohort #2, single-nucleotide polymorphisms (SNPs) targeted by CGP were superimposed with Phase 3 1000 Genomes data and projected down to the top five principal components and used to train a random forest classifier to recognize five general ancestries: African, admixed American, East Asian, European, and South Asian [1].

Programmed death ligand 1-tumor proportion score in NSCLC of cohort #1 and cohort #2

The PD-L1 TPS was determined for NSCLCs from DFCI, MSKCC and FMI by immunohistochemistry using validated anti-PD-L1 antibodies including E1L3N (Cell Signaling Technology, Danvers, MA), 22C3 (Dako North America Inc, Carpinteria, CA), and 28-8 (Epitomics Inc, Burlingame, CA), according to local institutional practice.

Tumor mutational burden assessment in cohort #1 and cohort #2

For cases from DFCI and MSKCC, TMB, defined as the number of somatic, coding, base substitution, and indel mutations per megabase (mut/Mb) of genome examined, was determined using OncoPanel and MSK-IMPACT,as previously described [2,3]. Regarding MSKCC NSCLC cases, comprehensive TMB data for harmonization from MSK MetTropism (MSK, Cell 2021) dataset were downloaded in cBioPortal. For NSCLC cases from DFCI and MSKCC, TMB distributions were

harmonized across the two sequencing platforms by applying a normal transformation followed by standardization to z-scores, as previously described [4]. TMB of cases from TCGA studies and China Pan-cancer according to each platform. For cohort #2, TMB was calculated by counting the number of non-driver synonymous and non-synonymous mutations across a 0.8–1.2 megabase (Mb) region, with computational germline status filtering, and reporting as mutations/Mb. This method has been previously validated for accuracy against whole exome sequencing [5].

Caris Life Sciences non-small cell lung cancer cohort

Next generation sequencing

NSCLC tumors molecularly profiled at Caris Life Sciences (Phoenix, AZ) between 2017-2022 were examined for *MET* TKD mutations. Next generation sequencing (NGS) was performed on genomic DNA isolated from formalin-fixed paraffin-embedded (FFPE) tumor samples using the NextSeq or NovaSeq 6000 platform (Illumina, Inc., San Diego, CA). For the Nextseq sequenced tumors, a custom-designed SureSelect XT assay was used to enrich 592 whole-gene targets (Agilent Technologies, Santa Clara, CA). Further, for the NovaSeq sequenced tumors, 719 clinically relevant genes were enriched at a high coverage and read-depth, along with another panel designed to enrich for >20,000 genes at lower read-depth. All variants were detected with >99% confidence based on allele frequency and amplicon coverage, with an average sequencing depth of coverage of >500 and an analytic sensitivity to identify variants with a variant allele frequency of ≥5%. Genetic variants identified were interpreted by board-certified molecular geneticists and categorized as ‘pathogenic (P),’ ‘likely pathogenic (LP),’ ‘variant of unknown significance,’

47 ‘likely benign,’ or ‘benign,’ according to the American College of Medical Genetics and
48 Genomics (ACMG) standards. When assessing mutation frequencies of individual genes,
49 ’pathogenic’ and ‘presumed pathogenic’ were counted as mutations. The copy number
50 alteration (CNA) of each exon was determined by calculating the average depth of the
51 sample along with the sequencing depth of each exon and comparing this calculated
52 result to a pre-caliberated value.
53
54 *Tumor Mutational Burdern Estimation*
55 After excluding all variants previously described as germline alterations according
56 to dbSNP151, Genome Aggregation Database (gnomAD) databases and those
57 characterized as benign variants by Caris geneticists, tumor mutational burden (TMB)
58 was calculated by counting non-synonmous, missense, nonsense, inframe
59 insertion/deletion and frameshift mutations present in each tumor. A cutoff point of ≥10
60 mutations per megabase was used to classify TMB-high (TMB-H) tumors, based on
61 findings from the KEYNOTE-158 study [6].
62
63 *Immunohistochemistry (IHC) for PD-L1 assessment*
64 IHC was performed on FFPE sections. Slides were stained using automated
65 staining techniques in accordance with the manufacturer’s protocol and were further
66 optimized and validated as per CLIA/CAO and ISO requirements. A board-certified
67 pathologist evaluated all IHC results independently. The primary PD-L1 antibody clone
68 was 22c3 (Dako). Tumor Proportion Score (TPS) was measured as the percentage of
69 viable tumor cells showing partial or complete membrane staining at any intensity. A

tumor was considered PD-L1 positive if its TPS was ≥1% and PD-L1 high if its TPS was ≥50%

Oncogenic driver mutations in NSCLC

Tumors positive for *MET*ex14 alterations were classified as those having likely LP/P mutations in exon 14 of the *MET* gene or that were positive for *MET*ex14 alterations (determined by WTS). Oncogenic driver positive NSCLC harbored at least one of the following alterations: LP/P mutations in *KRAS*, *EGFR*, *BRAF*, *ERBB2*; copy number amplification of *ERBB2* (≥6 copies), *MET*ex14 alterations, LP/P rearrangements in *ALK*, *ROS1*, *RET*, *NTRK 1-3*, or *NRG1*.

Preclinical models

Antibodies and compounds

Antibodies against pMET (Tyr-1234; sc-101736; 1:1000) and MET (sc-161; 1:1000) were purchased from Santa Cruz Biotechnology; phospho-ERK1/2 (#4370; 1:1000); ERK1/2 (#9102;1:1000), phospho-AKT (#4060; 1:1000), AKT (#9272; 1:1000), HSP 90 (#4877; 1:1000), phospho-EGFR (#3777; 1:1000); EGFR (#4267; 1:1000); anti-rabbit IgG HRP-linked antibody (#7074; 1:5000 or 1:10000), and anti-mouse IgG HRP-linked antibody (#7076; 1:5000 or 1:10000) from Cell Signaling Technology. Crizotinib, capmatinib, tepotinib, cabozantinib, and osimertinib were purchased from Selleckchem. Elzovantinib (TPX-0022) was provided by Turning Point Therapeutics.

Transfection and Luciferase assay

Transfection experiments were performed using FuGENE HD Transfection Reagent (Promega) as per the manufacturer's protocol. For luciferase assays, the pGL4.33[luc2P/SRE/Hygro] and pRL-TK vectors (Promega) were co-transfected with MET expression vectors. At 72 hours after transfection, the cells were lysed and subjected to either western blotting or luciferase assay. Luciferase activity was measured using the Dual-Luciferase Reporter Assay System (Promega). All experimental conditions were set up in 3 wells.

Structural analysis

Generation of the MET Models

The cocrystal structure of MET with ATP (PDBID:3DKC) was used to build the simulation system. Initial modeling was done in the Schrodinger 2022-3. The protein preparation wizard was used to assign bond orders, add hydrogens, create disulfide bonds, and fill in missing side chains using the Prime module. Ionization and tautomeric states for heteroatom groups were generated using the Epik module at neutral pH (pH 7.0 ± 2.0), and protonated states of titratable residues were determined by pKa calculations at physiological conditions (pH 7.0 ± 2.0) using the PROPKA module. A restrained minimization was also performed on all atoms using the OPLS4 force field [7]. The minimization was considered converged once heavy atom displacement was below 0.3 angstroms. Orientations of asparagine and glutamine were sampled at this step to optimize the hydrogen bonding network using the built in workflow. Waters within 5.0 angstroms of the ATP and Mg+2 were retained [8,9]. This initial structure contained 3 engineered residues at positions 1194, 1234, and 1235. As such, we used this

structure as a template for the generation of active state ATP bound MET kinase homology models for mutants. To model the wild type structure, the canonical sequence (Primary accession number: P08581) was used. Appropriate substitutions in the canonical sequence was done to generate the H1094Y, F1200I, and R1170Q models [10]. These three sequences were aligned using ClustalW to the structure of crystal structure. ATP, Mg+2 ion, and the one water molecule that chelates the Mg+2 ion were used as co-factors when building the homology models.

ATP Parameterization

The initial ATP bound in wild type MET crystal was used for parametrization. In this model, the triphosphate tail of ATP is fully deprotonated resulting in a -4 charge. We generated a force field based optimized complex for the initial geometry for the quantum mechanical calculations. The initial complex underwent a global unrestrained minimization using the OPLS4 forced field using the Prime minimization panel. An implicit (variable-dielectric generalized Born model) solvent model for water was used during the minimization. From this minimized complex, the ATP, Mg+2 ion, chelating water, residues 1209 and 1222 (residues with side chains that interact with the Mg+2 ion), and residues 1159 and 1160 (residues that hydrogen bond with the ATP adenosine group) were extracted for quantum mechanical treatment. The three peptide fragments (residues 1159 and 1160, 1209, and 1222), were capped using acetyl group or N-methyl amide on the N and C terminus, respectively. These capping groups were orientated to retain the backbone angles of neighboring residues. This complex served as the initial coordinates for quantum mechanical calculations for charges.

This complex then underwent geometry optimization at the B3LYP-D3/6-311G(d,p)++. Diffuse and polarizable functions were added to the basis set. Default parameters were chosen for Jaguar optimization [11,12]. The initial Hessian for the geometry optimization was calculated from the basis set. Backbones of the three peptide fragments (residues 1159 and 1160, 1209, and 1222) including capping groups were constrained in the X, Y, and Z axis keeping their position fixed. The side chains of residues 1159 and 1160 were also held fixed. Due to the side chains of residues 1209 and 1222 chelating the Mg+2, no constraints were applied to these side chains. Output of the geometry optimization was used as the initial coordinates for single point energy calculations for electrostatic potential charge (ESP) generation utilizing the B3LYP-D3 with the 6-311G(d,p)++ basis set [13,14]. This was used in building the simulation system.

Model System Generation

Using CHARMM-GUI enhanced sampler, we generated input files for the 4 ATP containing MET, the active wild type, H1094Y, F1200I, and R1170Q systems [15,16]. A rectangular water box was fitted to the protein structure. The edge distance was specified as 12.5 angstroms. In addition to neutralizing ions, a KCl concentration of 150 mM was used. This accounts for approximately 130 ions in each system. TIP3P was used as the water model for these systems. Parameters for ions were derived from the TIP3P leaprc file and included a 12-6 non-bonded correction. The FF19SB force field was used to parameterize protein residues [17]. Hydrogen mass repartitioning was also

used in these simulations to enable a 4-fs time step by redistributing heavy atom mass to the attached hydrogens.

All-Atom MD Simulations

All molecular dynamics simulations were performed in Amber 22 using reference GeForce RTX 3060 ti or 3080 ti GPUs. The PMEMD.CUDA program in AMBER22 was used to conduct all simulations [17]. The equilibration protocol was consisted of a two-stage minimization protocol with the steepest descent (5,000 steps) and conjugate gradient (5,000 steps) for each model system. A restraint potential of 1.0 kcal/mol·Å2 was applied to protein residues during the initial minimization to remove any potential steric clashes from K+ or Cl− ions, and water molecules. Equilibrium simulations were carried out after initial minimization calculations. These equilibration simulations were carried out for a total of 0.25 ns using an integration time step of 1 fs. Langevin dynamics was used to control the temperature of the system. Initial velocities were generated for a target temperature of 303.15K. Long-range electrostatics were calculated using the particle mesh Ewald method with a 9.0 angstrom cutoff. The SHAKE algorithm was used to treat the solvent molecules**.** Production molecular dynamics simulations were carried out for 1000ns with snapshots taken every 0.1 ns.

Analysis of MD Simulations

All analyses were performed on the last 500ns of the production trajectories. After removal of water and charged ions except for the Mg+2 ion, production trajectories were outputted in dcd and netcdf formats for follow-up analysis. The production

trajectories additionally had been saved with 1ns interval between each frame. The distance measurements were taken using CPPTRAJ over the last 500 ns of the simulations. RMSF and RMSD values were also calculated using CPPTRAJ [18]. The MMPBSA.py utility in AmberTools was used to calculate the ΔG of binding based on a single trajectory MM/GBSA protocol. The last 500 frames of the reduced trajectories representing snap shots from 500 ns to 1000 ns at an interval of 1ns were used. ATP was used as the ligand in calculations. The Mg+2 ion was treated as part of the receptor. The ante-MMPBSA.py utility was used to generate the initial input files based on the parameterized complex [19]. These calculations were carried out using a modified GB mode} with the reference salt concentration of 0.15 M [20]. Statistical analysis was performed using GraphPad. A Mann–Whitney–Wilcoxon test was used to test interaction distances.

Supplementary References

1. Connelly CF, Carrot-Zhang J, Stephens PJ, Frampton GM. Abstract 1227: Somatic genome alterations in cancer as compared to inferred patient ancestry. Cancer Res. 2018; 78: 1227.
2. Garcia EP, Minkovsky A, Jia Y, Ducar MD, Shivdasani P, Gong X, et al. Validation of OncoPanel: A Targeted Next-Generation Sequencing Assay for the Detection of Somatic Variants in Cancer. Arch Pathol Lab Med. 2017;141:751–8.
3. Cheng DT, Mitchell TN, Zehir A, Shah RH, Benayed R, Syed A, et al. Memorial Sloan Kettering-Integrated Mutation Profiling of Actionable Cancer Targets (MSK-IMPACT): A Hybridization Capture-Based Next-Generation Sequencing Clinical Assay for Solid Tumor Molecular Oncology. J Mol Diagn. 2015;17:251–64.
4. Vokes NI, Liu D, Ricciuti B, Jimenez-Aguilar E, Rizvi H, Dietlein F, et al. Harmonization of Tumor Mutational Burden Quantification and Association With Response to Immune Checkpoint Blockade in Non–Small-Cell Lung Cancer. JCO Precis Oncol. 2019;3:1–12.
5. Lee JK, Sivakumar S, Schrock AB, Madison R, Fabrizio D, Gjoerup O, et al. Comprehensive pan-cancer genomic landscape of KRAS altered cancers and real-world outcomes in solid tumors. NPJ Precis Oncol. 2022; 6:91.

6. Marabelle A, Fakih M, Lopez J, Shah M, Shapira-Frommer R, Nakagawa K, et al. Association of tumour mutational burden with outcomes in patients with advanced solid tumours treated with pembrolizumab: prospective biomarker analysis of the multicohort, open-label, phase 2 KEYNOTE-158 study. Lancet Oncol. 2020; 21:1353–65.
7. Lu C, Wu C, Ghoreishi D, Chen W, Wang L, Damm W, et al. OPLS4: Improving Force Field Accuracy on Challenging Regimes of Chemical Space. J Chem Theory Comput. 2021; 17:4291–4300.
8. Shelley JC, Cholleti A, Frye LL, Greenwood JR, Timlin MR, Uchimaya M. Epik: A software program for pKa prediction and protonation state generation for drug-like molecules. J Comput Aided Mol Des. 2007; 21:681–691.
9. Madhavi Sastry G, Adzhigirey M, Day T, Annabhimoju R, Sherman W. Protein and ligand preparation: Parameters, protocols, and influence on virtual screening enrichments. J Comput Aided Mol Des. 2013; 27:221–234.
10. Larkin MA, Blackshields G, Brown NP, Chenna R, Mcgettigan PA, McWilliam H, et al. Clustal W and Clustal X version 2.0. Bioinformatics. 2007; 23:2947–2948.
11. Bochevarov AD, Watson MA, Greenwood JR, Philipp DM. Multiconformation, Density Functional Theory-Based pKa Prediction in Application to Large, Flexible Organic Molecules with Diverse Functional Groups. J Chem Theory Comput. 2016;12:6001-6019.
12. Yu HS, Watson MA, Bochevarov AD. Weighted Averaging Scheme and Local Atomic Descriptor for pKa Prediction Based on Density Functional Theory. J Chem Inf Model. 2018; 58:271–286.
13. Ditchfield R, Hehre WJ, Pople JA. Self-Consistent Molecular-Orbital Methods. IX. An Extended Gaussian-Type Basis for Molecular-Orbital Studies of Organic Molecules. J Chem Phys. 1971; 54:724–728.
14. Grimme S, Antony J, Ehrlich S, Krieg H. A consistent and accurate ab initio parametrization of density functional dispersion correction (DFT-D) for the 94 elements H-Pu. J Chem Phys. 2010; 132:154104.
15. Jo S, Cheng X, Lee J, Kim S, Park SJ, Patel DS, et al. CHARMM-GUI 10 years for biomolecular modeling and simulation. J Comput Chem. 2017;38:1114–1124.
16. Jo S, Kim T, Iyer VG, Im W. CHARMM-GUI: a web-based graphical user interface for CHARMM. J Comput Chem. 2008;29:1859-1865.
17. Alabbas AB. Identification of promising methionine aminopeptidase enzyme inhibitors: A combine study of comprehensive virtual screening and dynamics simulation study. Saudi Pharm J. 2023;31:101745.
18. Roe DR, Cheatham TE. PTRAJ and CPPTRAJ: Software for Processing and Analysis of Molecular Dynamics Trajectory Data. J Chem Theory Comput. 2013;9:3084-3095.
19. Miller BR, McGee TD, Swails JM, Homeyer N, Gohlke H, Roitberg AE. MMPBSA.py: An Efficient Program for End-State Free Energy Calculations. J Chem Theory Comput. 2012;8:3314-3321.
20. Onufriev A, Bashford D, Case DA. Exploring protein native states and large-scale conformational changes with a modified generalized born model. Proteins. 2004;55:383-394.