Data-driven construction of anti-tumor agents with controlled polypharmacology

Chenxiao Da[†], Dehui Zhang[†], Michael Stashko[†], Eleana Vasileiadi[⊥], Rebecca E. Parker[⊥], Katherine A. Minson[⊥], Madeline G. Huey[⊥], Justus M. Huelse[⊥], Debra Hunter[‡§], Thomas S. K. Gilbert[§], Jacqueline Norris-Drouin[†], Michael Miley[§], Laura E. Herring[§], Lee M. Graves[§], Deborah DeRyckere[⊥], H. Shelton Earp[‡§], Douglas K. Graham[⊥], Stephen V. Frye[†‡], Xiaodong Wang[†*] and Dmitri Kireev[†*]

[†]Center for Integrative Chemical Biology and Drug Discovery, Division of Chemical Biology and Medicinal Chemistry, Eshelman School of Pharmacy, University of North Carolina at Chapel Hill, Chapel Hill, North Carolina, 27599-7363

[§]Department of Pharmacology, [‡]Lineberger Comprehensive Cancer Center, Department of Medicine, School of Medicine, University of North Carolina at Chapel Hill, Chapel Hill, North Carolina 27599, United States

[⊥]Aflac Cancer and Blood Disorders Center, Children's Healthcare of Atlanta, and Department of Pediatrics, Emory University, Atlanta, GA 30322

*e-mail: xiaodonw@unc.edu; dmitri.kireev@unc.edu

Table of contents

Supplementary methods

FRASE database

The Uniprot database[1] was searched for human proteins featuring a kinase domain with available 3D structure. The hit list of 3,680 PDB codes (for 358 proteins) was intersected with the list of high-affinity complexes extracted from the PDBbind database[2], yielding 2,800 complexes of 2,100 unique ligands with 230 proteins. Additionally to PDB structures, full kinase-domain sequences were extracted in FASTA format. All PDB files were automatically processed using a Pipeline Pilot[3] protocol, which included removal of extra chains and water molecules. Each sequence was aligned with the fully solved structure of PKA (1APM) using Clustal W[4] and 3D structures superposed using Schrodinger's *structalign* tool[5]. A suite of Pipeline Pilot protocols was then used to locate FRASEs in the global space of aligned structures of ligand-protein complexes. A FRASE was defined as a ligand fragment with all nearby protein residues (*i.e.*, residues having at least one atom within 4.5 Å from the closest ligand's atom). Ligands were fragmented using the "Generate fragments" component of Pipeline Pilot. A total of ~230,000 FRASEs were identified.

FRASE database screening

The following database queries were used to screen the FRASE database for relevant ligand fragments:

1) A two-point query was developed to identify a rigid linker that would connect the potential substituent attachment site to the putative entry into the back-pocket. The attachment site was

defined as the α-carbon of the butyl group of the template ligand **1** (XYZ = 18.5, 88.4, 19.1). The putative entry into the back-pocket, located at ~6 Å from the attachment site, was defined via visual inspection of the pocket's surface representation (22.2, 88.8, 16.4). This query resulted in 1,786 FRASE search hits (representing 73 unique PDB structures and 30 unique proteins). Hits were defined as FRASEs featuring ligand fragments having at least one atom within 1 Å from the query point 1 and 1.5 Å from the query point 2 and protein environments with a *Tscore* of at least 0.5. The above thresholds were tuned to obtain a large, but manageable number of hits (~1-3% of the database to enable their visual triage). The resulting 3D fragment poses are available as a supplementary file, "FRASE-search-init-hits.sdf.gz", in zipped SD format.

2) A one-point query was used to identify an optimal pivoting spacer at the entry into the back-pocket. The point was placed on the continuation of the triple bond of compound **2** (23.2, 88.1, 15.7). This query resulted in 151 FRASE search hits (corresponding to 9 unique PDB structures and 6 unique proteins). The resulting 3D fragment poses are available as a supplementary file, "FRASE-search-aryl-spacer-hits.sdf.gz", in zipped SD format.

3) Two one-point queries were used to explore potential options to fill the back-pocket space. Each query represents an alternative location (27.3, 88.5, 14.0 or 24.4, 91.2, 13.0) of a potential substituent dependent on the orientation of the pivoting spacer. One of the queries yielded 1,837 fragment hits (119 non-redundant, 30 unique PDB, 7 kinases) and the other, 155 hits (21 non-redundant, 11 unique PDB, 8 kinases). The resulting 3D fragment poses for both queries are available as supplementary files, "FRASE-search-bp-fill-1-hits.sdf.gz" and "FRASE-search-bp-fill-2-hits.sdf.gz", in zipped SD format.

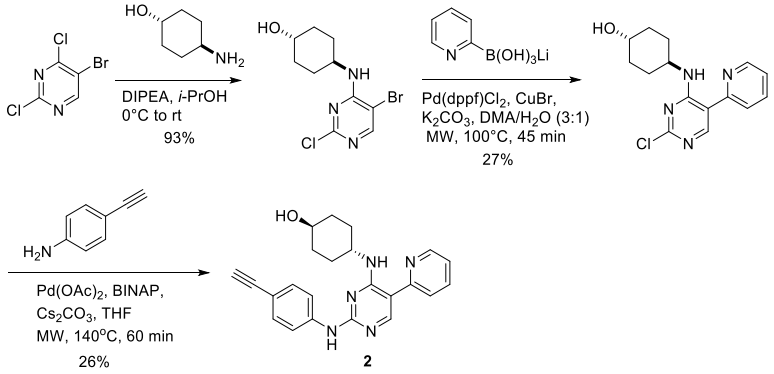
The next step is triage of the initial fragment hits. One notable aspect of ligand fragments in the FRASE database is their significant redundancy due to the fact that a fragment can be cut out of a ligand in multiple overlapping ways (this redundancy was intentionally allowed to maximize the sensitivity of the database search). We reduced this redundancy by clustering on chemical similarity of the ligand fragments and the location of their geometrical centers. The Pipeline Pilot implementation of K-means clustering was used with a minimum similarity of 0.5 and maximum distance of 2 Å between cluster members.

In the supplementary fragment files (see above), the fragments that passed the triage step were flagged with Triage = "Passed" and the fragments that were selected to inspire the ligand design were flagged with Selected = "True".

Synthesis of compounds 2-10

General experimental information: Microwave reactions were carried out using a CEM Discover-S reactor with a vertically-focused IR external temperature sensor and an Explorer 72 autosampler. The dynamic mode was used to set up the desired temperature and hold time with the following fixed parameters: PreStirring, 1 min; Pressure, 200 psi; Power, 200 W; PowerMax, off; Stirring, high. Flash chromatography was carried out with pre-packed silica gel disposable columns. Preparative HPLC was performed with the UV detection at 220 or 254 nm. Samples were injected onto a 75 x 30 mm, 5 µM, C18(2) column at rt. The flow rate was 30 mL/min. Various linear gradients were used with A being H_2O + 0.5% TFA and B being MeOH. Analytical thin-layer chromatography (TLC) was performed with silica gel 60 F_{254}, 0.25 mm pre-coated TLC plates. TLC plates were visualized using UV_{254} and phosphomolybdic acid with charring. All ^1H NMR spectra were obtained with a 400 MHz spectrometer using chloroform-d ($CDCl_3$, 7.26 ppm), dimethyl sulfoxide-d_6 (DMSO-d_6, 2.50 ppm), or methanol-d_4 (CD_3OD, 3.31 ppm) as an internal reference. Signals are reported as m (multiplet), s (singlet), d (doublet), t (triplet), q (quartet), and bs (broad singlet); and coupling constants are reported in Hertz (Hz). ^{13}C NMR spectra were

obtained with a 100 MHz spectrometer using $CDCl_3$ (77.2 ppm), DMSO-d_6 (39.5 ppm), or CD_3OD (49.0 ppm) as the internal standard. LC-MS was performed using an analytical instrument with the UV detector set to 220 nm, 254 nm, and 280 nm, and a single quadrupole mass spectrometer using electrospray ionization (ESI) source. Samples were injected (2 µL) onto a 4.6 x 50 mm, 1.8 µM, C18 column at rt. A linear gradient from 10% to 100% B (MeOH with 0.1% acetic acid) in 5.0 min was followed by pumping 100% B for another 2 or 4 minutes with A being H_2O with 0.1% acetic acid. The flow rate was 1.0 mL/min. High-resolution (positive ion) mass spectra (HRMS) were acquired using a LCMS-TOF mass spectrometer.



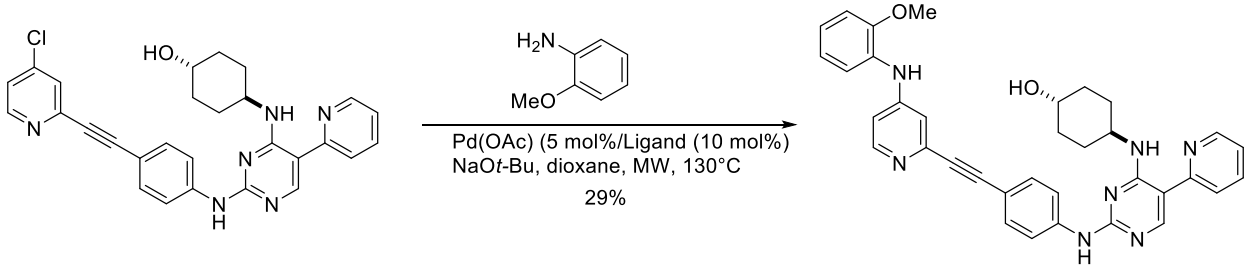
***trans*-4-((2-((4-Ethynylphenyl)amino)-5-(pyridin-2-yl)pyrimidin-4-yl)amino) cyclohexan-1-ol** (**2**) (General Procedure A): A solution of 5-bromo-2,4-dichloropyrimidine (114 g, 500 mmol) and in *i*-PrOH (450 mL) was added DIPEA (91.5 mL, 525 mmol) and a solution of 4-aminocyclohexanol (66.5 g, 500 mmol) in *i*-PrOH (450 mL) slowly at 0℃. The reaction mixture was stirred at room temperature for 12 h, and then the solvent was evaporated under reduced pressure. The residue was dissolved in EtOAc and washed with brine (200 mL×2). The organic layer was dried (Na_2SO_4) and concentrated to provide the crude product. Recrystallization of the crude product from a mixture of EtOAc/Hexane yielded *trans*-4-((5-bromo-2-chloropyrimidin-4-yl)amino)cyclohexan-1-ol (142g, 93% yield).

A solution of *trans*-4-((5-bromo-2-chloropyrimidin-4-yl)amino)cyclohexan-1-ol (3.06 g, 10.0 mmol) in dimethylacetamide (40 mL) was added K_2CO_3 (4.14 g, 30.0 mmol) in water (H_2O, 10 mL) at room temperature. After 3 minutes, lithium 2-pyridyl trihydroxy borate (4.5 g, 30.0 mmol) was added in one portion. The reaction mixture was stirred for 20 min and was added [1,1′-bis(diphenylphosphino)ferrocene]dichloropalladium(II) (816 mg, 1.00 mmol) and copper(I) bromide (284 mg, 2.00 mmol) in one portion. Then the reaction mixture was heated at 100℃ open to air for 45 min. After passing through a pad of celite and concentreated, the residue was purified by an ISCO silica gel column to provide the desired product 4-((2'-chloro-[4,5'-bipyrimidin]-4'-yl)amino)cyclohexanol as a light yellow solid (821 mg, 27% yield).

A solution of 4-((2'-chloro-[4,5'-bipyrimidin]-4'-yl)amino)cyclohexanol (61mg, 0.20 mmol) and 4-ethynylaniline (28 mg, 0.24 mmol) in THF (1.0 mL) was added palladium(II) acetate (4.5 mg, 0.020 mmol), 2,2′-bis(diphenylphosphino)-1,1′-binaphthalene (12.5 mg, 0.020 mmol) and cesium carbonate (156 mg, 0.48 mmol). The resulting mixture was heated under microwave irradiation at 140℃ for 60 min, then allowed to cool to room temperature. The solvent was removed and the residue was purified by an ISCO silica gel column to provide the desired compound **2** as a light yellow solid **(**20 mg, 26% yield, UNC3203A, DZ00177-136). ^{1}H NMR (400 MHz, CD_3OD) δ 8.65–8.61 (m, 1H), 8.43 (s, 1H), 8.08 (d, J = 8.0 Hz, 1H), 7.98–7.92 (m, 1H), 7.81–7.73 (m, 1H), 7.63 (d, J = 8.0 Hz, 1H), 7.61 (d, J = 8.0 Hz, 1H), 7.56–7.52 (m, 1H), 7.45–7.41 (m, 1H), 4.15–4.05 (m, 1H), 3.72–3.63 (m, 1H), 3.51 (s, 1H), 2.25–2.12 (m, 2H), 2.10–2.00 (m, 2H), 1.63–1.35 (s, 4H); MS *m/z* $[M+H]^+$ Calcd for $C_{23}H_{24}N_5O$: 386.20; found: 386.50; LC-MS >95% purity.

***trans*-4-((2-((4-((4-Fluoropyridin-2-yl)ethynyl)phenyl)amino)-5-(pyridin-2-yl)pyrimidin-4-yl)amino)cyclohexan-1-ol** (**3**): The title compound **3** (200 mg, 6.6% yield over three steps) was prepared according to general procedure A from 4-((4-fluoropyridin-2-yl)ethynyl)aniline (335 mg, 1.58 mmol) and 2-pyridyl trihydroxy borate (2570 mg, 17.5 mmol) as a yellow solid. ^{1}H NMR (CD_3OD, 400 MHz): δ 8.70 (t, J = 8.0 Hz, 1H), 8.64 (d, J = 4.0 Hz, 1H), 8.60–8.55 (m,1H), 8.52–8.47 (m, 1H), 8.00–7.93 (m, 2H), 7.85–7.70 (m, 4H), 7.51–7.41 (m, 2H), 4.14–4.07 (m, 1H), 3.69–3.67 (m, 1H), 2.18 (d, J = 8.0 Hz, 2H), 2.04 (d, J = 8.0 Hz, 2H), 1.61–1.53 (m, 2H), 1.50–1.40 (m, 2H); MS m/z 481.20 $[M+1]^+$.

***trans*-4-((2-((4-((4-Chloropyridin-2-yl)ethynyl)phenyl)amino)-5-(pyridin-2-yl)pyrimidin-4-yl)amino)cyclohexan-1-ol**: The title compound (23 mg, 5.8% yield over three steps) was prepared according to general procedure A from 4-((4-chloropyridin-2-yl)ethynyl)aniline (55 mg, 0.24 mmol) as a yellow solid. ^{1}H NMR (CD_3OD, 400 MHz): 8.67–8.64 (m, 1H), 8.60–8.58 (m, 1H), 8.51 (s, 1H), 7.99–7.97 (m, 2H), 7.92 (dd, J_1 = 8.0 Hz, J_2 = 4.0 Hz, 1H), 7.77–7.65 (m, 4H), 7.64 (d, J = 8.0 Hz, 1H), 7.47–7.43 (m,1H), 4.16–4.09 (m, 1H), 3.72–3.66 (m, 1H), 2.22–2.19 (m, 2H), 2.09–2.05 (m, 2H), 1.63–1.42 (m, 4H);. MS *m/z* 497.17 $[M+H]^+$.



***trans*-4-((2-((4-((4-((2-Methoxyphenyl)amino)pyridin-2-yl)ethynyl)phenyl) amino)-5-(pyridin-2-yl)pyrimidin-4-yl)amino)cyclohexan-1-ol** (**4**) (General Procedure B): To a mixture of *trans*-4-((2-((4-((4-chloropyridin-2-yl)ethynyl)phenyl)amino)-5-(pyridin-2-yl)pyrimidin-4-yl)amino)cyclohexan-1-ol (50 mg, 0.10 mmol), 2-methoxyaniline (37 mg, 0.30 mmol), $Pd_2(dba)_3$ (4.6 mg, 0.0050 mmol), XPhos (4.7 mg, 0.010 mmol) and sodium *tert*-butoxide (30 mg, 0.30 mmol) in a mixture of dioxane (1.0 mL) was heated under microwave irradiation at 130 ℃ for 60 min under nitrogen atmosphere. The resulting mixture was cooled to room temperature, filtrated through a pad of celite, and concentrated. The residue was purified through preparative HPLC to afford the title product as a yellow solid (20 mg, 29% yield) (UNC4507A). ^{1}H NMR (CD_3OD, 400 MHz): δ 8.62 (d, J = 4.0 Hz, 1H), 8.48 (s, 1H), 8.02 (s, 1H), 7.93 (d, J = 8.0 Hz, 1H), 7.71 (d, J = 8.0 Hz, 1H), 7.66–7.60 (m, 4H), 7.56–7.51 (m, 2H), 7.40 (dd, J_1 = 8.0 Hz, J_2 = 4.0 Hz, 1H), 7.32–7.28 (m, 1H), 7.23–7.19 (m, 1H), 7.14–7.11 (m, 1H),7.03–6.97 (m, 1H), 4.12–4.07 (m, 1H), 3.85 (s, 3H), 3.71–3.66 (m, 1H), 2.19 (d, J = 8.0 Hz, 2H), 2.04 (d, J = 8.0 Hz, 2H), 1.61–1.39 (m, 4H); MS *m/z* 584.30 $[M+H]^+$.

***trans*-((5-(Pyridin-2-yl)-2-((4-((4-(pyrimidin-5-yl)pyridin-2-yl)ethynyl)phenyl) amino)pyrimidin-4-yl)amino)cyclohexan-1-ol** (**5**): The title compound **5** (17 mg, 31% yield) was prepared according to general procedure B from K_3PO_4 (42.4 mg, 0.20 mmol) (instead of sodium *tert*-butoxide) and pyrimidin-5-ylboronic acid (25 mg, 0.20 mmol) as a yellow solid. ^{1}H NMR (DMSO-D_6, 400 MHz): δ 11.27 (s, 1H), 10.92 (d, J = 8.0 Hz, 1H), 9.49 (s, 1H), 9.43 (s, 1H), 9.26–9.16 (m, 2H), 8.65–8.61 (m, 1H), 8.45–8.35 (m, 1H), 8.10–7.95 (m, 2H), 7.85–7.75 (m, 1H), 7.75–7.65 (d, J = 8.0 Hz, 1H), 7.47–7.42 (m, 1H), 7.36–7.34 (m, 1H), 7.24–7.21 (m, 2H), 7.10–7.07 (m, 1H), 4.07–4.02 (m, 2H), 2.08–1.98 (m, 2H), 1.93–1.87 (m, 2H), 1.50–1.32 (m, 4H); MS *m/z* 541.30 $[M+H]^+$.

***trans*-4-((2-((4-((4-(1*H*-Pyrazol-4-yl)pyridin-2-yl)ethynyl)phenyl)amino)-5-(pyridine-2-yl)pyrimidin-4-yl)amino)cyclohexan-1-ol** (**6**): The title compound **6** (16 mg, 30% yield) was prepared according to general procedure B from K_3PO_4 (42.4 mg, 0.20 mmol) (instead of sodium *tert*-butoxide) and 1,1-dimethylethyl ester 4-(4,4,5,5-tetramethyl-1,3,2-dioxaborolan-2-yl)-1*H*-pyrazole-1-carboxylic acid (60 mg, 0.20 mmol) as a yellow solid. ^{1}H NMR (DMSO-D_6, 400 MHz): δ 11.24 (brs, 1H), 10.81 (brs, 1H), 8.94–8.90 (m, 1H), 8.70 (s, 1H), 8.63 (d, J = 4.0 Hz, 1H), 8.25–8.23 (m, 2H), 8.12 (br, 1H), 8.04–7.96 (m, 2H), 7.77 (d, J = 8.0 Hz, 2H), 7.70–7.60 (m, 3H), 7.45–7.42 (m, 1H), 4.01–3.93 (m, 2H), 2.07–2.04 (m, 2H), 1.91–1.88 (m, 2H), 1.51–1.30 (m, 4H); MS *m/z* 529.25 $[M+H]^+$.

***trans*-4-((2-((4-((4-(1,3-Dioxolan-2-yl)pyridin-2-yl)ethynyl)phenyl)amino)-5-(4-((4-methyl piperazin-1-yl)methyl)phenyl)pyrimidin-4-yl)amino)cyclohexan-1-ol** (**7**): The title compound **7** (2.31 g, 27% yield over three steps) was prepared according to general procedure A from 1-methyl-4-(4-(4,4,5,5-tetramethyl-1,3,2-dioxaborolan-2-yl)benzyl)piperazine (3.96 g, 12.5 mmol) and 4-((4-(1,3-dioxolan-2-yl)pyridin-2-yl)ethynyl)aniline (1.60 g, 6.0 mmol) as a yellow solid. ^{1}H NMR (400 MHz, CD_3OD) δ 8.77 (dd, J_1= 2.0 Hz, J_2= 8.0 Hz, 1H), 8.13–8.12 (m, 1H), 7.92 (dd, J_1= 2.0 Hz, J_2= 8.0 Hz, 1H), 7.81–7.78 (m, 4H), 7.75 (d, J= 8.0 Hz, 2H), 7.69 (s, 1H), 7.54 (d, J= 8.0 Hz, 2H), 6.00 (s, 1H), 4.42 (s, 2H), 4.11 (s, 4H), 3.81–3.47 (m, 10H), 3.00 (s, 3H), 2.06–1.96 (m, 4H), 1.58–1.45 (m, 2H), 1.43–1.27 (m, 2H), MS *m/z* 646.00 $[M+H]^+$.

***N*-Cyclopropyl-2-((4-((5-(2-fluoro-4-(morpholinosulfonyl)phenyl)-4-((*trans*-4-hydroxycyclohexyl)amino)pyrimidin-2-yl)amino)phenyl)ethynyl)isonicotinamide** (**8**): The title compound **8** (20 mg, 5.9% yield over three steps) was prepared according to general procedure A from 4-((3-fluoro-4-(4,4,5,5-tetramethyl-1,3,2-dioxaborolan-2-yl)phenyl)sulfonyl) morpholine (167 mg, 0.45 mmol) and 2-((4-aminophenyl)ethynyl)-*N*-cyclopropylisonicotinamide (125 mg, 0.45 mmol) as a yellow solid. ^{1}H NMR (400 MHz, CD_3OD) δ 8.72 (d, J= 4.0 Hz, 1H), 8.05 (s, 1H), 7.82–7.78 (m, 2H), 7.75–7.65 (m, 7H), 4.17–4.07 (m, 1H), 3.76–3.72 (m, 4H), 3.57–3.45 (m, 1H), 3.09–3.05 (m, 4H), 2.92–2.86 (m, 1H), 1.98 (t, J= 12.0 Hz, 4H), 1.52–1.40 (m, 2H), 1.40–1.24 (m, 2H), 0.86–0.80 (m, 2H), 0.69–0.64 (m, 2H), MS *m/z* 712.30 $[M+H]^+$.

***N*-Cyclopropyl-6-((4-((4-((*trans*-4-hydroxycyclohexyl)amino)-5-(4-((4-methyl piperazin-1-yl)methyl)phenyl)pyrimidin-2-yl)amino)phenyl)ethynyl) picolinamide** (**9**): The title compound **9** (70 mg, 13% over three steps) was synthesized according to general procedure A from 1-methyl-4-(4-(4,4,5,5-tetramethyl-1,3,2-dioxaborolan-2-yl)benzyl)piperazine (237 mg, 0.75 mmol) and 6-((4-aminophenyl)ethynyl)-N-cyclopropylpicolinamide (100 mg, 0.36 mmol) as an orange solid. ^{1}H NMR (400 MHz, D_2O) δ 7.81–7.73 (m, 3H), 7.51–7.46 (m, 3H), 7.44–7.38 (m, 2H), 7.33 (d, J= 2.0 Hz, 2H), 7.24 (d, J= 2.0 Hz, 2H), 3.97 (s, 2H), 3.74–3.64 (m, 1H), 3.63–3.52 (m, 1H), 3.48–3.27 (m, 4H), 3.24–3.06 (m, 4H), 2.84 (s, 3H), 2.71–2.65 (m, 1H), 1.99–1.86 (m, 4H), 1.33–1.19 (m, 4H), 0.80–0.73 (m, 2H), 0.56–0.51 (m, 2H); MS *m/z* $[M+H]^+$ Calcd for $C_{39}H_{45}N_8O_2$: 657.40; found: 657.00; LC-MS >95% purity.

***N*-(2-((4-((4-(*trans*-4-Hydroxycyclohexyl)amino)-5-(pyridin-2-yl)pyrimidin-2-yl)amino)phenyl)ethynyl)pyridin-4-yl)isobutyramide** (**10**): The title compound **10** (218 mg, 10% over three steps) was synthesized according to general procedure A from lithium 2-pyridyl trihydroxy borate (1300 mg, 9.0 mmol) and *N*-(2-((4-aminophenyl)ethynyl)pyridin-4-yl)isobutyramide (334 mg, 1.2 mmol) as a yellow solid. ^{1}H NMR (400 MHz, CD_3OD) δ 8.65 (dt, J_1= 2.0 Hz, J_2= 4.0 Hz, 1H), 8.52 (d, J= 8.0 Hz, 1H), 8.52 (s, 1H), 8.36 (d, J= 4.0 Hz, 1H), 8.02–7.96 (m, 3H), 7.86–7.77 (m, 4H), 7.45 (dd, J_1= 4.0 Hz, J_2= 8.0 Hz, 1H), 4.17–4.07 (m, 1H), 3.75–3.65 (m, 1H), 2.79–2.70 (m, 1H), 2.24–2.16 (m, 2H), 2.11–2.04 (m, 2H), 1.64–1.53 (m, 2H), 1.53–1.40 (m, 2H), 1.25 (d, J = 4.0 Hz, 6H); MS *m/z* $[M+H]^+$ Calcd for $C_{32}H_{34}N_7O_2$: 548.30; found: 548.00; LC-MS >95% purity.

Microfluidic Capillary Electrophoresis Assay

Microfluidic Capillary Electrophoresis (MCE) Assay Activity assays were performed in a 384 well, polypropylene microplate in a final volume of 50 μL of 50 mM Hepes, Ph 7.4 containing 10 mM MgCl2, 1.0 mM DTT, 0.01% Triton X-100, 0.1% Bovine Serum Albumin (BSA), containing 1.0 μM fluorescent substrate (see below) and ATP at the Km for each enzyme (Table S1). All reactions were terminated by addition of 20 μL of 70 mM EDTA. After a 180 min incubation, phosphorylated and unphosphorylated substrate peptides (Table 5) were separated in buffer supplemented with 1 x CR-8 on a LabChip EZ Reader equipped with a 12-sipper chip. Data were analyzed using EZ Reader software. The following substrates were used for (i) Mer: 5-FAM-EFPIYDFLPAKKK-CONH at 10 μM (1.7 nM protein, 22.3 μM ATP); Axl: 5-FAM-KKKKEEIYFFF-CONH2 at 10 μM (16 nM protein, 200 μM ATP); Tyro3: 5-FAM-EFPIYDFLPAKKK-CONH2 at 10 μM (5 nM protein, 40 μM ATP); Flt3: 5-FAM-KKKKEEIYFFF-CONH2 at 10 μM (0.3 nM protein, 275 μM ATP).

X-rays crystallography

Crystallization: Crystals of Mer in complex with compound **8** were obtained by vapor diffusion from sitting drops at 12 °C. Protein at 32.5 mg/mL in crystallization buffer (20 mM Tris pH 8.0, 500 mM sodium chloride, 2 mM β-mercaptoethanol) was incubated with **8** (dissolved in DMSO) to give a final concentration of 2.5 mM and slowly rocked overnight. This solution was mixed 1:1 with and equilibrated against crystallization solution containing 27-33% (v/v) Peg 400, 200 mM magnesium chloride, 100 mM Tris pH 8.5. Plate-like crystals grew to final dimensions of up to 1000 x 400 x 50 μm over 10 days. Prior to diffraction data collection, crystals were vitrified by plunging into liquid nitrogen. The Mer:**8** crystals displayed the symmetry of space group $P2_1$ with cell parameters a = 51.4 Å, b = 92.0 Å, c = 69.7 Å, β = 101.67°, contained two molecules in the asymmetric unit, and diffracted X-rays to a minimum Bragg spacing of about 2.90 Å.

Structure determination: Data were collected at the Southeast Regional Collaborative Access Team (SER-CAT) 22-ID beamline at the Advanced Photon Source, Argonne National Laboratory at a wavelength of 1.07426 Å and a temperature of 100K. Data were processed using the program HKL2000. The diffraction quality of our crystals was non-uniform: when the thin edge was exposed to X-rays, the resulting diffraction spots were radially smeared, which was accompanied by a loss of resolution and a somewhat reduced completeness in the high-resolution shells.

The structure of the Mer:**8** complex was determined by molecular replacement with the program Phaser [9]. The search model was generated from the coordinates of Mer in complex with ADP (PDB entry 3BRB) with all non-protein atoms removed. Refinement was carried out using the program Phenix [9], interspersed with manual revisions of the model using the program Coot [10]. Refinement consisted of conjugate-gradient minimization and calculation of individual atomic displacement and translation/libration/screw (TLS) parameters. To avoid any model bias, coordinates for **8** were not included until the remainder of the model (including water molecules and ions) was completed. Inclusion of high-resolution, albeit weaker, data increased the stability and convergence of the refinement process. For data collection and refinement statistics see Table 2. The current model contains two molecules of the Mer kinase domain with only one monomer bound to compound **8**, 5 chloride ions, 1 magnesium ion, and 3 water molecules. Residues that could not be identified in the electron density were: chain A, 596-598, 621-632, 658-666, 745-762; chain B, 620-636, 657-666, 743-762, 775-776. The model exhibits excellent geometry as determined by MolProbity [11]. Ramachandran analysis identified 96.3% favored, 3.7% allowed, and 0.0% disallowed residues. The coordinates and structure factors have been deposited in the RCSB Protein Data Bank under accession number **6MEP**.

MIB/MS kinome profiling

Sample preparation: NOMO-1 human AML cells were grown in RPMI-1640 + 10% FBS. Cells were treated for 1 hour with 100 nM or 1 μM of compounds **1** and **10**. One hour after treatment, cells were lysed on ice with MIB lysis buffer [50 mM HEPES (pH 7.5), 0.5% Triton X-100, 150 mM NaCl, 1 mM EDTA, 1 mM EGTA, 10 mM sodium fluoride, 2.5 mM sodium orthovanadate, 1X protease inhibitor cocktail (Roche), 1% phosphatase inhibitor cocktail 2 (Sigma-Aldrich), and 1% of phosphatase inhibitor cocktail 3 (Sigma-Aldrich)]. Lysates were sonicated 3X10 seconds at 40% power then centrifuged 10,000g for 10 min at 4°C. Supernatant was collected and passed through a 0.2 μm syringe filter (Corning 0.2 micron SFCA syringe filter). Lysates not immediately used were promptly stored at -80°C. Columns of mixed kinase inhibitor-conjugated sepharose beads (MIBs) consisting of 2 inhibitors UNC1917 (**1**) and UNC4042 (**10**) were prepared. 5 mg of lysate from each experimental condition was brought to 1M NaCl and passed over the columns by gravity flow. Columns were subsequently washed with 5ml of a high salt wash buffer and 5ml of a low salt wash buffer [50 mM HEPES (pH 7.5), 0.5% Triton X-100, 1 mM EDTA, 1 mM EGTA, and 10 mM sodium fluoride, and 1 M NaCl or 150 mM NaCl, respectively]. Columns were washed a final time with 1ml of 0.1% SDS buffer. Kinases were eluted from the column by boiling with 1ml of elution buffer [0.5% SDS, 1% 2 mercaptoethanol, and 0.1 M Tris (pH 6.8)] for 10 min. Eluates were reduced, alkylated, then concentrated with Amicon spin concentrators (10k cutoff, Millipore). Detergent was removed from the concentrated eluate by chloroform-methanol extraction and a series of methanol washes. Protein pellets were resuspended in 50 mM HEPES (pH 8.0) and digested overnight at 37°C with sequencing grade modified trypsin (Promega). Peptides were desalted using PepClean C18 spin columns (Thermo Scientific).

LC-MS/MS Analysis: Samples were analyzed by LC/MS/MS using an Easy nLC 1200 coupled to a QExactive HF mass spectrometer (Thermo Scientific). Samples were injected onto an Easy Spray PepMap C18 column (75 μm id × 25 cm, 2 μm particle size) (Thermo Scientific) and separated over a 120 min method. The gradient for separation consisted of 5–38% mobile phase B at a 250 nl/min flow rate, where mobile phase A was 0.1% formic acid in water and mobile phase B consisted of 0.1% formic acid in 80% acetonitrile. The QExactive HF was operated in data-dependent mode where the 15 most intense precursors were selected for subsequent fragmentation. Resolution for the precursor scan (m/z 350–1600) was set to 120,000 with a target value of 3×10^6 ions. MS/MS scans resolution was set to 15,000 with a target value of 1×10^5 ions, 75 ms max IT. The normalized collision energy was set to 27% for HCD, with an isolation widow of 1.6 m/z. Dynamic exclusion was set to 30 s and precursors with unknown charge or a charge state of 1 and ≥ 8 were excluded.

Data Analysis: Raw data files were processed using MaxQuant version 1.6.3.4 and searched against the reviewed human database (downloaded Jan 2019, containing 20,414 entries), using Andromeda within MaxQuant. Enzyme specificity was set to trypsin, up to two missed cleavage sites were allowed, carbamidomethylation of C was set as a fixed modification and oxidation of *M* and acetyl of *N*-term were set as variable modifications. A 1% FDR was used to filter all data and match between runs was enabled. A minimum of two peptides was required for label-free quantitation using the LFQ intensities. Further analyses were performed in Perseus version 1.6.0.2 and Excel. Fold changes were calculated using the average LFQ intensity of the compound divided by the average LFQ intensity of the DMSO control per sample set. ANOVA was performed on the UNC4042 (**10**) data and a p-value of ≤0.05 was considered significant. A full set of MIB-MS results is available as a supplementary Excel file.

In vivo pharmacokinetics (Sai Life Sciences)

Plasma pharmacokinetics of compound **10** was investigated in a group of 9 male Swiss Albino mice following a single intraperitoneal dose administration. Animals were administered intraperitoneally with **10** solution formulation in 10% DMSO, 10% Tween 80 and normal saline at 3 mg/kg dose. Blood samples (approximately 60 mL) were collected from retro orbital plexus under light isoflurane anesthesia such that the samples were obtained at 0.08, 0.25, 0.5, 1, 2, 4, 6, 8 & 24 hrs (*i.p.*). The blood samples were collected from set of three mice at each time point in labeled micro centrifuge tube containing K2EDTA as anticoagulant. Plasma samples were separated by centrifugation of whole blood and stored below -70 ºC until bioanalysis. All samples were processed for analysis by protein precipitation using acetonitrile (ACN) and analyzed with fit for purpose LC/MS/MS method (LLOQ: 1.23 ng/mL). Pharmacokinetic parameters were calculated using the non-compartmental analysis tool of Phoenix WinNonlin (Version 6.3). The overall pharmacokinetic parameters are summarized in **Supplementary Table S4**.

Cell Lines and Culture Conditions

A549 and Kasumi-1 cell lines were obtained from the American Type Culture Collection. OCI-AML5, NOMO-1, and COLO-699 cell lines were obtained from the Deutsche Sammlung von Mikroorganismen und Zellkulturen (DSMZ). Cell line identities were confirmed by short-tandem repeat microsatellite loci analysis. Cells were cultured in RPMI-1640 medium supplemented with 10% fetal bovine serum (FBS; Atlanta Biologics), 100u/ml penicillin, and 100μg/ml streptomycin at 37°C and in 5% CO2.

Cell Based Kinase Inhibition Assays

697 B-ALL cells and A549 NSCLC cells were cultured in the presence of **10** or DMSO vehicle for 1.0 h. Pervanadate solution was prepared fresh by combining 20 mM sodium orthovanadate in 0.9x PBS in a 1:1 ratio with 0.3% (w/w) hydrogen peroxide in PBS for 15-20 min at room temperature. Cultures were treated with 120 μM pervanadate for 3 min prior to preparation of cell lysates in 50 mM HEPES pH 7.5, 150 mM NaCl, 10 mM EDTA, 10% glycerol, and 1% Triton X-100, supplemented with protease inhibitors (Roche Molecular Biochemicals, #11836153001). MERTK or AXL proteins were immunoprecipitated with anti-MERTK (R&D Systems, #MAB8912) or anti-AXL (R&D Systems, #AF154) antibody and Protein G agarose beads (InVitrogen). Phospho-proteins were detected by western blot using an anti-phospho-MERTK antibody raised against a peptide derived from the tri-phosphorylated activation loop of MERTK [25] (Phopshosolutions, Inc) or an anti-phosphotyrosine antibody (4G10 Platinum, Millipore, #05-1050). Membranes were stripped and total proteins were detected using anti-MERTK (Abcam, #ab52968) or anti-AXL (R&D Systems, #AF154) antibodies. Phosphorylated and total proteins were quantitated by densitometry using Image J software and IC_{50} values were calculated by non-linear regression.

Soft Agar Colony Formation Assays

OCI-AML5, NOMO-1, A549 or COLO-699 cells were cultured in 1.5 mL of 0.35% soft agar containing 1x RPMI-1640 medium and 10% FBS and overlaid with 2.0 mL of 1x RPMI medium containing 10% FBS and **10** or DMSO vehicle. Medium and **10** or vehicle were refreshed 3 times per week. Colonies were stained with nitrotetrazolium blue chloride (Sigma Aldrich, #N6876) and counted after 14 days.

Apoptosis Assays

Kasumi-1 and Nomo-1 cells were cultured in the presence of **10** or DMSO vehicle for 72 hours, then stained on ice with 1μM PO-PRO®-1 iodide (Invitrogen) and 1.5μM propidium iodide (PI) (Invitrogen) in PBS for 15-20 minutes. Fluorescence was detected and analyzed using a FC 500 flow cytometer with CXP data analysis software (Beckman Coulter).

In Vivo Pharmacodynamic Assay

Six to eight-week old NOD.Cg-*Prkdc*^{scid} *Il2rg*^{tm1Wjl}/SzJ (NSG) mice were purchased from Jackson Laboratories or bred in-house and maintained under sterile conditions. Mice were injected intravenously with 2x10^6 697 B-cell ALL cells. Twenty-one days after xenograft transplantation, mice were treated with a single 60 mg/kg dose of **10** or an equivalent dose (10mL/kg) of vehicle (10% DMSO and 10% Tween 80 in saline). Two hours later, mice were euthanized and bone marrow was flushed from femurs with RPMI medium containing 20% fetal bovine serum and 240μM freshly prepared pervanadate (see above), then incubated at room temperature for 10 minutes prior to preparation of cell lysates for detection of phosphorylated and total MERTK proteins as described above. All animal studies were approved by the Emory University Animal Care and Use Committee.

Supplementary Tables

Table S1: MCE assay data for compounds **2-10**

ID	Compound ID	Axl IC_{50} (nM)	Mer IC_{50} (nM)	Tyro3 IC_{50} (nM)	Flt3 IC_{50} (nM)
2	UNC3203A	58±27*	4±3	10±9	8±8
3	UNC4391A	1,400±720	240±10	400±90	80±38
4	UNC4507A	100±0	170±0	110±0	360±0
5	UNC4362A	560±47	81±15	140±23	49±22
6	UNC4360A	2,500±250	150±53	660±240	40±17
7	UNC3437A	377±252	45±49	100±130	430±270
8	UNC3614A	46±29	140±20	6,800±6,000	1,200±750
9	UNC3321A	2,200±920	240±73	170±96	280±42
10	UNC4042A	8±4	5±2	7±5	29±16
11	UNC3399B	5±2	3±1	11±4	73±16
12	UNC4337A	51±24	3±2	66±21	6±3
13	UNC2491A	84±18	1±0.3	33±11	5±0.5

*Mean ± Standard deviation

Table S2: Crystallographic data collection and refinement statistics

Data collection	
Space group	$P2_1$
Cell dimensions: *a*, *b*, *c* (Å), β (°)	51.4, 92.0, 69.7, 101.7
Resolution (Å)	37.07–2.90 (2.92–2.90)
R_{merge} (%)[a]	14.2 (77.1)
$I/\sigma I$	11.0 (2.0)
Cut-off (σ)	-3
Unique reflections	14,260 (337)
Completeness (%)	99.9 (98.5)
Redundancy	4.2 (3.7)
Wilson B-factor (Å^2)	43.5
Refinement	
Resolution (Å)	37.07–2.89 (3.05–2.89)
No. of reflections (work/free)	11,634/1,000 (1026/88)
Cut-off (σ)	0
R_{work} / R_{free}	20.4/26.9 (21.0/27.5)
No. of atoms	
Protein	7,763
Compound **8**	91
Ions	6
Water	3
B-factors (Å^2)	
Protein	43.7
Compound **8**	59.6
Ions	40.4
Water	15.2
R.m.s. deviations	
Bond lengths (Å)	0.002
Bond angles (°)	0.534
Ramachandran[b]	
Favored (%)	96.3
Generally Allowed (%)	3.7
Disallowed (%)	0.0
Missing residues	A: 596-598, 621-632, 658-666, 745-762 B: 620-636, 657-666, 743-762, 775-776

Values in parentheses denote highest resolution shell

[a] $R_{merge} = 100\ \Sigma_h\Sigma_i | I_{h,i} - \langle I_h \rangle | / \Sigma_h \Sigma_i I_{h,i}$, where the outer sum (*h*) is over the unique reflections and the inner sum (*i*) is over the set of independent observations of each unique reflection.

[b] As defined by the validation suite MolProbity [11].

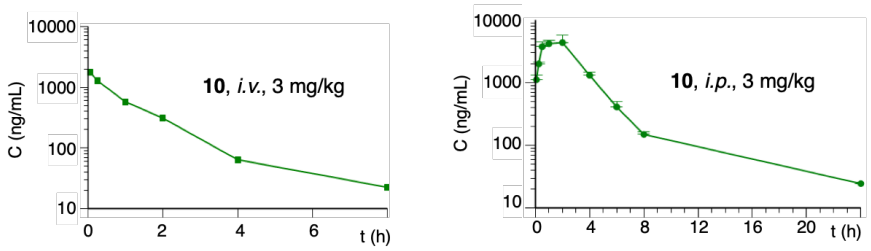
Table S3: Key results of the in vivo pharmacokinetic study for compound **10**.

Compound	Dose (mg/kg)	Route	T_{max} (hr)	$^{a}C_0/C_{max}$ (ng/mL)	AUC_{last} (hr*ng/mL)	AUC_{inf} (hr*ng/mL)	$T_{1/2}$ (hr)	CL (mL/min/kg)	Vss (L/kg)
10	3	*i.p.*	2	435	1661	1678	4.97	-	-
		i.v.	-	1914	2086	2133	1.68	23.4	2.13
		p.o.	-	N/C	N/C	N/C	-	-	-

[a]back extrapolated concentration for the *i.v.* group.
N/C: Not calculated due to insufficient data
The metrics include: maximum concentration of drug observed in plasma (C_{max}), time when C_{max} was observed (T_{max}), areas under curve calculated until last time point (AUC_{last}) and extrapolated to infinity (AUC_{inf}), half-life time ($T_{1/2}$), clearance (CL), steady state distribution volume (V_{ss}). Compound **10** was administered via intravenous (i.v.), intraperitoneal (i.p.) and oral (p.o.) routes. Drug concentration in plasma was measured;

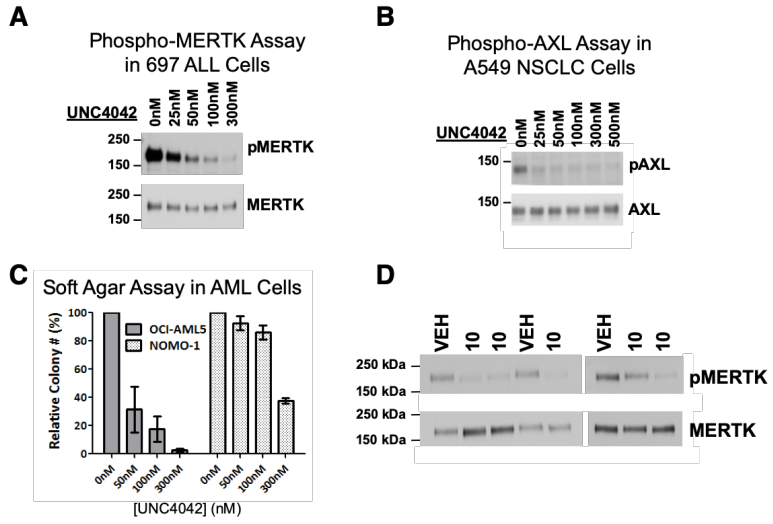
Supplementary Figures

Figure S1: *In vivo* pharmacokinetics



Time course of plasma concentration observed for **10** administrated to mice by *i.v.* (left) and *i.p.* (right) routes.

Figure S2: Supplementary cellular and *in vivo* data



A,**B**. Compound **10** (UNC4042) inhibits MERTK and AXL phosphorylation. 697 ALL cells (**A**) or A549 NSCLC cells (**B**) were cultured with **10** or vehicle for 1 hour and then treated with pervanadate phosphatase inhibitor to stabilize phosphorylated proteins. MERTK (**A**) or AXL (**B**) were

immunoprecipitated from cell lysates and phosphorylated and total MERTK or AXL proteins were detected by immunoblot; **C.** Compound **10** (UNC4042) decreases colony forming potential. AML cells (OCI-AML5 and NOMO-1) were cultured in soft agar overlaid with medium containing **10** or vehicle only for 10-14 days. Media and **10** were refreshed every 3 days; **D**. Compound **10** inhibits MERTK phosphorylation in bone marrow leukemia cells *in vivo*. NSG mice were inoculated with 697 human leukemia cells and leukemic mice were treated with a single dose of 60 mg/kg **10** or saline vehicle (VEH) by intraperitoneal injection. Bone marrow cells were isolated from femurs 2 hours later and incubated with pervanadate phosphatase inhibitor for 10 minutes to stabilize phosphoproteins. MERTK was immunoprecipitated and phosphorylated and total MERTK proteins were detected by immunoblot.

Supplementary references

1. Apweiler, R. *et al.* UniProt: the universal protein knowledgebase. *Nucleic Acids Res.* **32,** D115–D119 (2004).
2. Wang, R., Fang, X., Lu, Y. & Wang, S. The PDBbind database: collection of binding affinities for protein-ligand complexes with known three-dimensional structures. *J. Med. Chem.* **47,** 2977–2980 (2004).
3. *Pipeline Pilot. ver. 9.1, Biovia Inc.* (2014).
4. Larkin, M. A. *et al.* Clustal W and Clustal X version 2.0. *Bioinformatics* **23,** 2947–2948 (2007).
5. *Maestro Suite, Schrodinger LLC* (2017).
6. Henikoff, S. & Henikoff, J. G. Amino acid substitution matrices from protein blocks. *Proc. Natl. Acad. Sci. U. S. A.* **89,** 10915–10919 (1992).
7. Hert, J. *et al.* Comparison of fingerprint-based methods for virtual screening using multiple bioactive reference structures. *J Chem Inf Comput Sci* **44,** 1177–1185 (2004).
8. Zhang, W. H. *et al.* Pseudo-cyclization through intramolecular hydrogen bond enables discovery of pyridine substituted pyrimidines as new mer kinase inhibitors. *J. Med. Chem.* **56,** 9683–9692 (2013).
9. McCoy, A. J. *et al.* Phaser crystallographic software. *J. Appl. Crystallogr.* **40,** 658–674 (2007).
10. Emsley, P. & Cowtan, K. Coot: Model-building tools for molecular graphics. *Acta Crystallogr. Sect. D Biol. Crystallogr.* **60,** 2126–2132 (2004).
11. Chen, V. B. *et al.* MolProbity : all-atom structure validation for macromolecular crystallography. *Acta Crystallogr. Sect. D Biol. Crystallogr.* **66,** 12–21 (2010).