SUPPORTING INFORMATION

DEVELOPMENT OF A POTENT, SPECIFIC CDK8 KINASE INHIBITOR WHICH PHENOCOPIES CDK8/19 KNOCKOUT CELLS

Michael F.T. Koehler[†*], Philippe Bergeron[†], Elizabeth M. Blackwood[‡], Krista Bowman[§], Kevin R. Clark[|], Ron Firestein[#], James R. Kiefer[§], Klaus Maskos[‖], Mark L. McCleland[#], Linda Orren, Laurent Salphati[⊥], Steve Schmidt[|], Elisabeth V. Schneider[‖°], Jiansheng Wu[◊] and Maureen H. Beresini[|]

[†]Department of Discovery Chemistry, [‡]Department of Translational Oncology, [§]Department of Structural Biology, [|]Department of Biochemical and Cellular Pharmacology, [#]Department of Pathology, Department of Drug Metabolism and Pharmacokinetics [⊥] and [◊]Department of Protein Chemistry, Genentech, Inc., 1 DNA Way, South San Francisco, CA 94080, United States

[‖]Proteros Biostructures GmbH, Bunsenstr. 7a, D-82152 Martinsried, Germany

[°]Max-Planck-Institut für Biochemie, Am Klopferspitz 18a, D-82152 Martinsried, Germany

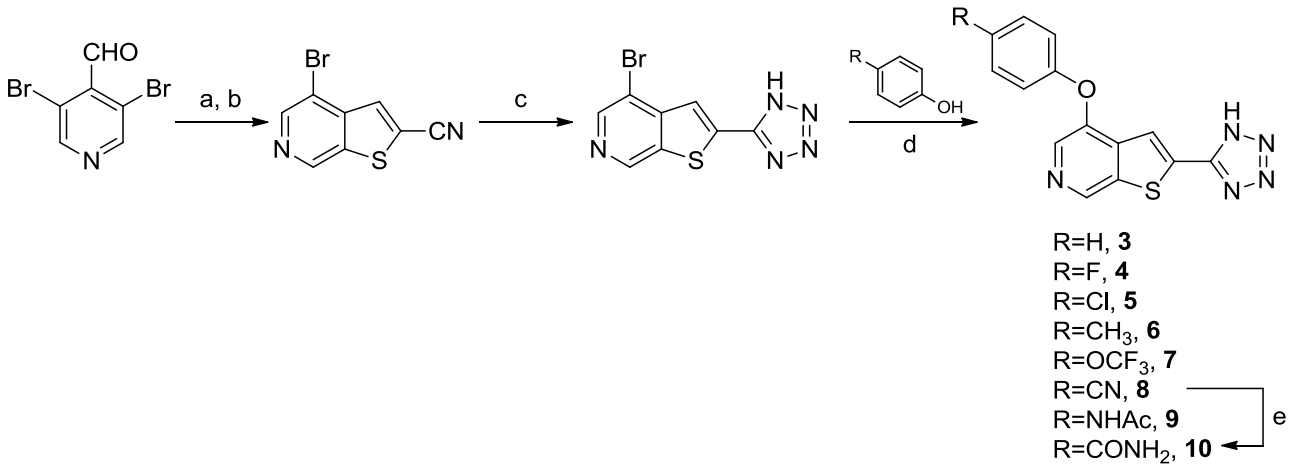
1. Chemistry

Experimental details:

General. Unless otherwise indicated, all reagents and solvents were purchased from commercial sources and were used without further purification Moisture or oxygen sensitive reactions were conducted under an atmosphere of argon or nitrogen gas. Unless otherwise stated, ^{1}H NMR spectra were recorded at 300 or 400 MHz using Varian or Bruker instruments operating at the indicated frequencies. Chemical shifts are expressed in ppm relative to a tetramethylsilane (ppm = 0.00) internal standard. The following abbreviations are used: br = broad signal, s = singlet, d = doublet, dd = doublet of doublets, t = triplet, q = quartet, p = pentet, m = multiplet. Purification by silica gel chromatography was carried out using Isco systems with prepacked cartridges. Chemical purities were >95% for all final compounds, as assessed by LC/MS analysis at UV 220 nm.

Compound 2. Compound 2 was synthesized according to previously described procedures.[1]
^{1}H NMR (400 MHz, DMSO-D_6) δ 9.25 (s, 1H), 8.31 (s, 1H), 8.04 – 7.98 (m, 2H), 7.96 (d, J = 8.3 Hz, 2H), 7.80 (d, J = 8.7 Hz, 2H), 7.75 (d, J = 8.3 Hz, 2H), 7.36 (s, 1H), 7.23 (d, J = 8.7 Hz, 2H). MS (ESI(+)): m/z 415.1 (M+H).

Scheme S1. Synthesis of compounds 3-10.[a]



[a](a) 2-mercaptoacetamide, Cs_2CO_3, THF, 60°C, 2h, 79%; (b) trifluoroacetic anhydride, pyridine, rt, 6h, 70%; (c) NaN_3, NH_4Cl, DMF, 80°C, 2h, 78%; (d) Cs_2CO_3, CuCl, 2,2,6,6-tetramethylheptane-3,5-dione, 120°C sealed tube, 16h; (e) H2O2, K2CO3, DMSO

Compound 3. Compound 3 was synthesized according to Scheme S1 with R=H.
^{1}H NMR (400 MHz, DMSO-D_6) δ 9.215 (s, 1H), 8.20 (s, 1H), 7.97 (s, 1H), 7.45 (t, J = 8 Hz, 2H), 7.24-7.14 (m, 3H). MS (ESI(+)): m/z 296.0 (M+H).

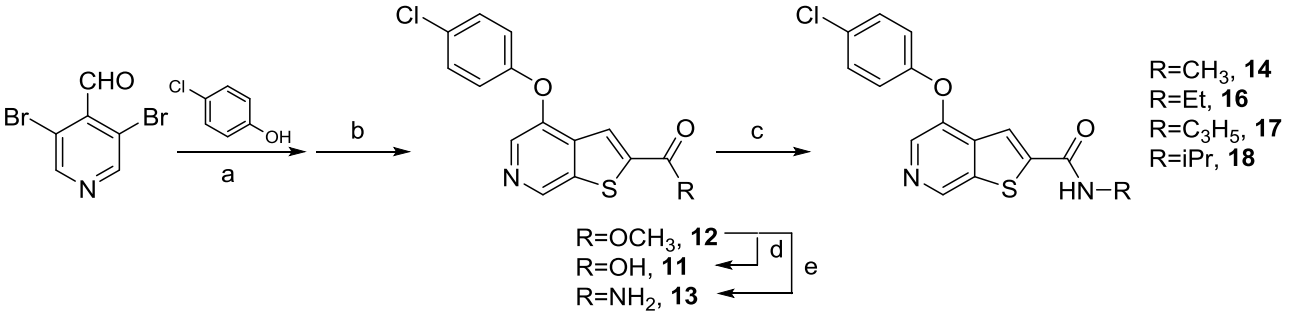
Compound 4. Compound 4 was synthesized according to Scheme S1 with R=F.
^{1}H NMR (400 MHz, DMSO-D_6) δ 9.14 (br s, 1H), 8.15 (br s, 1H), 7.80 (s, 1H), 7.31 – 7.23 (m, 2H), 7.23 – 7.16 (m, 2H), 1.30 – 1.19 (m, 0H). MS (ESI(+)): m/z 314.0 (M+H).

Compound 5. Compound 5 was synthesized according to Scheme S1 with R=Cl.
^{1}H NMR (400 MHz, DMSO-D_6) δ 9.21 (br s, 1H), 8.26 (br s, 1H), 7.84 (s, 1H), 7.47 (d, J = 8.9 Hz, 2H), 7.16 (d, J = 8.9 Hz, 2H). MS (ESI(+)): m/z 330.0 (M+H).

Compound 6. Compound 6 was synthesized according to Scheme S1 with R=CH3.
^{1}H NMR (400 MHz, DMSO-D_6) δ 9.19 (br s, 1H), 8.21 (br s, 1H), 7.79 (s, 1H), 7.23 (d, J = 8.5 Hz, 2H), 7.02 (d, J = 8.5 Hz, 2H), 2.31 (s, 3H). MS (ESI(+)): m/z 310.2 (M+H).

Compound 7. Compound 7 was synthesized according to Scheme S1 with R=OCF_3.
^{1}H NMR (400 MHz, DMSO-D_6) δ 9.27 (s, 1H), 8.30 (s, 1H), 7.97 (s, 1H), 7.44 (d, J = 8.2 Hz, 2H), 7.25 (d, J = 9.1 Hz, 2H). (ESI(+)): m/z 380.1 (M+H).
Compound 8. Compound 8 was synthesized according to Scheme S1 with R=CN.
^{1}H NMR (500 MHz, MeOH-D_4) δ 9.04 (s, 1H), 8.18 (s, 1H), 7.95 (d, J = 9.5 Hz, 2H), 7.86 (s, 1H), 7.15 (d, J = 8.5 Hz, 2H). MS (ESI(+)): m/z 321.0 (M+H).
Compound 9. Compound 9 was synthesized according to Scheme S1 with R=NHAc.
^{1}H NMR (500 MHz, MeOH-D_4) δ 8.99 (br s, 1H), 8.04 (s, 1H), 7.99 (br s, 1H), 7.64 (d, J = 7 Hz, 2H), 7.14 (d, J = 7 Hz, 2H), 2.15 (s, 3H). MS (ESI(+)): m/z 353.2 (M+H).
Compound 10. Compound 10 was synthesized according to Scheme S1 with R=$CONH_2$.
^{1}H NMR (500 MHz, MeOH-D_4) δ 9.11 (br s, 1H), 8.26 (br s, 1H), 7.82 (s, 1H), 7.79 (d, J = 5 Hz, 2H), 7.22 (d, J = 7 Hz, 2H). MS (ESI(+)): m/z 339.0 (M+H).

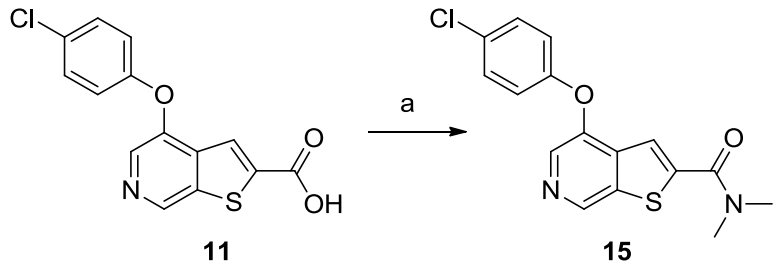
Scheme S2. Synthesis of compounds 11-14 and 16-18[a]



[a](a) 1 eq KOtBu, THF, 0°C, 1h then 70°C, 3h; (b) methyl thioglycolate, Cs2CO3, THF, 70°C, 0.5h; (c) RNH2, 80°C, sealed tube, 16h; (d) LiOH, THF, H2O, RT, 3h, then 2N HCl; (e) NH4OH, MeOH, RT, 16h

Compound 11. Compound 11 was synthesized according to Scheme S2 with R=OH.
^{1}H NMR (400 MHz, DMSO-D_6) δ 9.20 (s, 1H), 8.23 (s, 1H), 7.83 (s, 1H), 7.47 (d, *J* = 8.9 Hz, 2H), 7.16 (d, *J* = 8.9 Hz, 2H). (ESI(+)): m/z 305.9 (M+H).
Compound 12. Compound 12 was synthesized according to Scheme S2 with R=OCH_3.
^{1}H NMR (400 MHz, DMSO-D_6) δ 9.23 (s, 1H), 8.24 (s, 1H), 7.94 (s, 1H), 7.48 (d, *J* = 9.0 Hz, 2H), 7.17 (d, *J* = 8.9 Hz, 2H), 3.91 (s, 3H). (ESI(+)): m/z 320.0 (M+H).
Compound 13. Compound 13 was synthesized according to Scheme S2 with R=NH_2.
^{1}H NMR (400 MHz, DMSO-D_6) δ 9.15 (s, 1H), 8.43 (s, 1H), 8.15 (d, *J* = 3.3 Hz, 2H), 7.84 (s, 1H), 7.47 (d, *J* = 8.9 Hz, 2H), 7.15 (d, *J* = 8.9 Hz, 2H). (ESI(+)): m/z 304.9 (M+H).
Compound 14. Compound 14 was synthesized according to Scheme S2 with R=CH_3.
^{1}H NMR (400 MHz, DMSO-D_6) δ 9.16 (s, 1H), 8.95 (q, 1H), 8.19 (s, 1H), 8.06 (s, 1H), 7.52-7.42 (m, 2H), 7.20-7.09 (m, 2H), 2.80 (d, *J* = 4.5 Hz, 3H) (ESI(+)): m/z 318.8 (M+H).
Compound 16. Compound 16 was synthesized according to Scheme S2 with R=Et.
^{1}H NMR (400 MHz, DMSO-D_6) δ 9.15 (s, 1H), 8.99 (t, *J* = 5.5 Hz, 1H), 8.17 (s, 1H), 8.14 (s, 1H), 7.50 – 7.44 (m, 2H), 7.17 – 7.11 (m, 2H), 3.32 – 3.26 (m, 2H), 1.13 (t, *J* = 7.2 Hz, 3H). (ESI(+)): m/z 333.1 (M+H).
Compound 17. Compound 17 was synthesized according to Scheme S2 with R=cPr.
^{1}H NMR (400 MHz, DMSO-D_6) δ 9.15 (s, 1H), 8.96 (d, *J* = 4.2 Hz, 1H), 8.15 (s, 2H), 8.13 (s, 0H), 7.52 – 7.43 (m, 2H), 7.18 – 7.09 (m, 2H), 2.89 – 2.81 (m, 1H), 0.78 – 0.66 (m, 4H), 0.63 – 0.56 (m, 2H). (ESI(+)): m/z 345.2 (M+H).
Compound 18. Compound 18 was synthesized according to Scheme S2 with R=iPr.
^{1}H NMR (400 MHz, DMSO-D_6) δ 9.14 (s, 1H), 8.78 (d, *J* = 7.7 Hz, 1H), 8.23 (s, 1H), 8.14 (s, 1H), 7.55 – 7.43 (m, 2H), 7.22 – 7.10 (m, 2H), 4.13 – 4.02 (m, 1H), 1.18 (d, *J* = 6.5 Hz, 6H). (ESI(+)): m/z 347.1 (M+H).

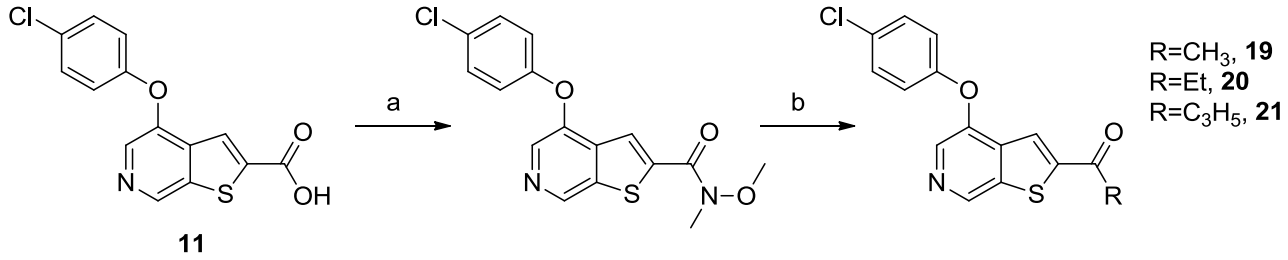
Scheme S3. Synthesis of compound 15.[a]



[a](a) dimethylamine hydrochloride, EDCI, HOBt, DIPEA

Compound 15. Compound 15 was synthesized according to Scheme S3.
^{1}H NMR (400 MHz, DMSO-D_6) δ 9.15 (s, 1H), 8.99 (t, *J* = 5.5 Hz, 1H), 8.17 (s, 1H), 8.14 (s, 1H), 7.50 – 7.44 (m, 2H), 7.17 – 7.11 (m, 2H), 3.32 – 3.26 (m, 2H), 1.13 (t, *J* = 7.2 Hz, 3H). (ESI(+)): m/z 333.1 (M+H).

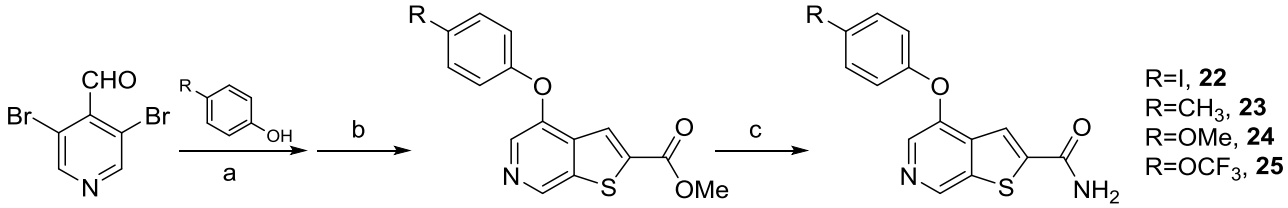
Scheme S4. Synthesis of compounds 19-21[a]



[a](a) N,O-dimethylhydroxylamine, EDCI, HOBt, N-methyl morpholine, DMF, RT, 15 min; (b) 2 eq. RMgBr, THF, 0°C to reflux

Compound 19. Compound 19 was synthesized according to Scheme S4, with R=CH_3.
^{1}H NMR (400 MHz, DMSO-D_6) δ 9.17 (s, 1H), 8.34 (s, 1H), 8.13 (s, 1H), 7.55 – 7.43 (m, 2H), 7.27 – 7.17 (m, 2H), 2.71 (s, 3H). (ESI(+)): m/z 303.8 (M+H).
Compound 20. Compound 20 was synthesized according to Scheme S4, with R=Et.
^{1}H NMR (400 MHz, DMSO-D_6) δ 9.17 (s, 1H), 8.30 (s, 1H), 8.14 (s, 1H), 7.53 – 7.47 (m, 2H), 7.24 – 7.18 (m, 2H), 3.18 (q, *J* = 7.2 Hz, 2H), 1.10 (t, *J* = 7.2 Hz, 3H). (ESI(+)): m/z 317.9 (M+H).
Compound 21. Compound 21 was synthesized according to Scheme S4, with R=cPr.
^{1}H NMR (400 MHz, DMSO-D_6) δ 9.19 (s, 1H), 8.57 (s, 1H), 8.16 (s, 1H), 7.55 – 7.45 (m, 2H), 7.26 – 7.17 (m, 2H), 3.20 – 3.09 (m, 1H), 1.12 (d, *J* = 6.0 Hz, 4H). (ESI(+)): m/z 329.9 (M+H).

Scheme S5. Synthesis of compounds 22-25[a]

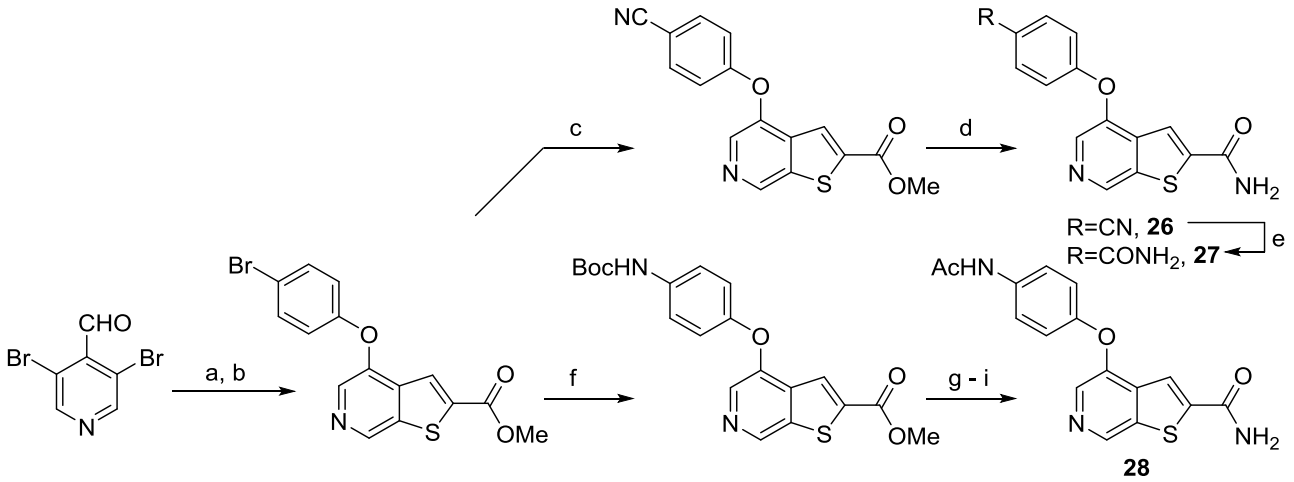


[a](a) 1 eq KOtBu, THF, 70°C, 3h; (b) methyl thioglycolate, Cs_2CO_3, THF, 70°C, 0.5h; (c) NH_3, MeOH, 16h

Compound 22. Compound 22 was synthesized according to Scheme S5.
^{1}H NMR (400 MHz, DMSO-D_6) δ 9.16 (s, 1H), 8.97 – 8.88 (m, 1H), 8.20 (s, 1H), 8.04 (s, 1H), 7.73 (d, J = 8.7 Hz, 2H), 6.92 (d, J = 8.7 Hz, 2H), 2.79 (d, J = 4.6 Hz, 3H). (ESI(+)): m/z 410.9 (M+H).
Compound 23. Compound 23 was synthesized according to Scheme S5, with R=CH_3.
^{1}H NMR (400 MHz, DMSO-D_6) δ 9.08 (s, 1H), 8.44 (br s, 1H), 8.21 (s, 1H), 8.00 (s, 1H), 7.84 (br s, 1H), 7.29 – 7.20 (m, 2H), 7.09 – 7.00 (m, 2H), 2.32 (s, 3H). (ESI(+)): m/z 285.0 (M+H).
Compound 24. Compound 24 was synthesized according to Scheme S5, with R=OCH_3.
^{1}H NMR (400 MHz, DMSO-D_6) δ 9.04 (s, 1H), 8.44 (br s, 1H), 8.27 (s, 1H), 7.91 (s, 1H), 7.85 (br s, 1H), 7.18 – 7.11 (m, 2H), 7.05 – 6.98 (m, 2H), 3.77 (s, 3H). (ESI(+)): m/z 301.1 (M+H).
Compound 25. Compound 25 was synthesized according to Scheme S5, with R=OCF_3.

^{1}H NMR (400 MHz, DMSO-D_6) δ 9.17 (s, 1H), 8.44 (br s, 1H), 8.19 (s, 2H), 8.17 (s, 1H), 7.85 (br s, 1H), 7.47 – 7.38 (m, 2H), 7.27 – 7.17 (m, 2H). (ESI(+)): m/z 355 (M+H).

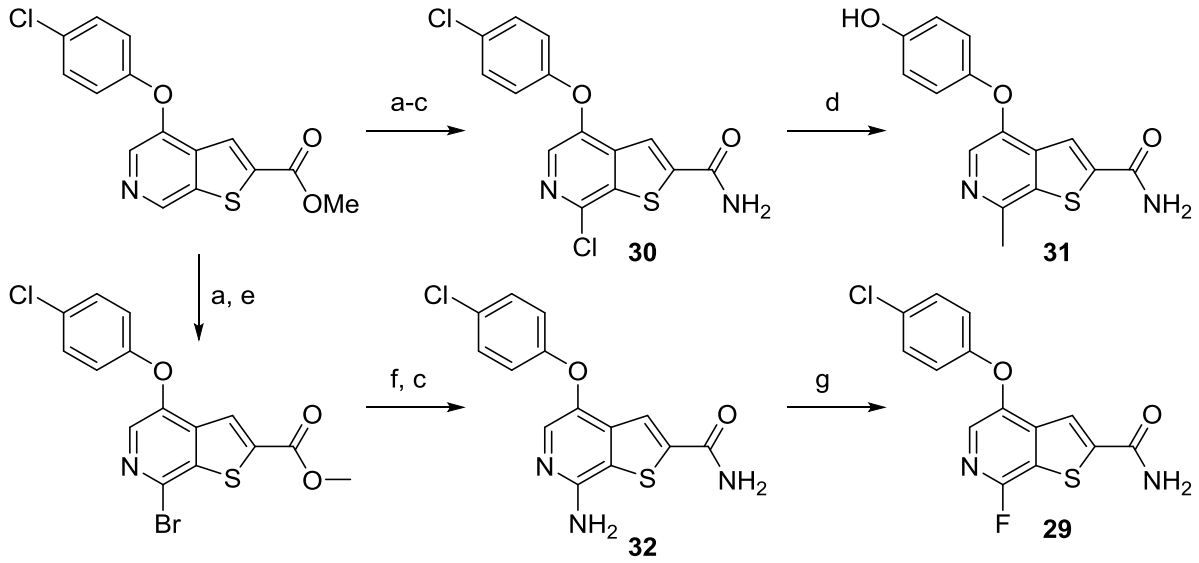
Scheme S6. Synthesis of compounds 26-28[a]



[a](a) 4-bromophenol, 1 eq KOtBu, THF, 70°C, 3h; (b) methyl thioglycolate, Cs2CO3, THF, 70°C, 0.5h; (c) 2.5 eq CuCN, 10 mol% Pd(OAc)2, Na2CO3, DMF, 125°C; (d) NH_3, MeOH, rt; (e) H_2O_2, K_2CO_3, DMSO; (f) 3 eq NH_2Boc, 10 mol% Pd(OAc)2, 10 mol% X-Phos, Cs_2CO_3, dioxane, 16h; (g) HCl, EtOAc; (h) AcCl, TEA, DCM, 2h; (i) NH_3, MeOH, RT

Compound 26. Compound 26 was synthesized according to Scheme S6.
^{1}H NMR (400 MHz, DMSO-D_6) δ 9.24 (s, 1H), 8.42 (br s, 1H), 8.35 (s, 1H), 8.05 (s, 1H), 7.91 – 7.86 (m, 2H), 7.85 (br s, 1H), 7.24 – 7.19 (m, 2H). (ESI(+)): m/z 296.1 (M+H).
Compound 27. Compound 27 was synthesized according to Scheme S6.
^{1}H NMR (400 MHz, DMSO-D_6) δ 9 9.18 (s, 1H), 8.44 (br s, 1H), 8.23 (s, 1H), 8.12 (s, 1H), 7.98 – 7.89 (m, 3H), 7.84 (br s, 1H), 7.33 (br s, 1H), 7.19 – 7.07 (m, 2H). (ESI(+)): m/z 314.1 (M+H).
Compound 28. Compound 28 was synthesized according to Scheme S6.
^{1}H NMR (400 MHz, DMSO-D_6) δ 10.01 (br s, 1H), 9.07 (s, 1H), 8.44 (br s, 1H), 8.23 (s, 1H), 7.99 (s, 1H), 7.84 (br s, 1H), 7.68 – 7.59 (m, 2H), 7.16 – 7.07 (m, 2H), 2.05 (s, 3H). (ESI(+)): m/z 328.1 (M+H).

Scheme S7. Synthesis of compounds 29-31[a]



[a](a) *m*-CPBA, DCM, rt, 16h; (b) 6 eq $POCl_3$, DCM, RT, 16h, 54% over two steps; (c) NH_3, MeOH, RT, 1h, 73%; (d) 3 eq trimethylboroxine, 10 mol % $Pd(PPh_3)_4$, 2 eq Cs_2CO_3, DMF, 100°C, 16h; (e) 7 eq $POBr_3$, DCM, rt, 16h; (f) NH_3, H_2O, Cu_2SO_4, 110°C then $SOCl_2$, MeOH, reflux; (g) $NaNO_2$, HF/pyridine

Compound 29. Compound 29 was synthesized according to Scheme S7.
1H NMR (400 MHz, DMSO-D6) δ 8.53 (br s, 1H), 8.19 (d, J = 3.2 Hz, 1H), 7.97 (br s, 1H), 7.85 (d, J = 1.9 Hz, 1H), 7.49 – 7.43 (m, 2H), 7.18 – 7.11 (m, 2H). (ESI(+)): 323.0 (M+H).
Compound 30. Compound 30 was synthesized according to Scheme S7.
^{1}H NMR (400 MHz, DMSO-D_6) δ 8.51 (br s, 1H), 8.23 (s, 1H), 8.03 (s, 1H), 7.96 (br s, 1H), 7.51 – 7.45 (m, 2H), 7.23 – 7.17 (m, 2H), 3.50 – 3.36 (m, 1H). (ESI(+)): m/z 338.8 (M+H).
Compound 31. Compound 31 was synthesized according to Scheme S7.
^{1}H NMR (400 MHz, DMSO-D_6) δ 8.60 (br s, 1H), 8.30 (s, 1H), 8.21 (s, 1H), 8.01 (br s, 1H), 7.54 – 7.47 (m, 2H), 7.24 – 7.17 (m, 2H), 2.88 (s, 3H). (ESI(+)): m/z 318.8 (M+H).
Compound 32. Compound 32 was synthesized according to Scheme S7.
^{1}H NMR (400 MHz, DMSO-D_6) δ 8.32 (br s, 1H), 8.14 (br s, 1H), 7.85 (s, 1H), 7.72 – 7.67 (m, 2H), 7.40 – 7.33 (m, 2H), 6.98 – 6.90 (m, 2H), 6.60 (br s, 1H). (ESI(+)): m/z 319.9 (M+H).

Equilibrium solubility determinations:
Sample preparation: For routine equilibrium solubility determination, samples are prepared by adding 4 µL of a 10 mM DMSO stock solution of each compound into 196 µL of PBS buffer (pH 7.4) in Millipore Multiscreen 96-well filter plates, yielding a compound concentration of 200 µM and 2% DMSO. The filter plate is sealed with aluminum sealing film and shaken at room temperature for 24 hours at 1000 rpm. When shaking has completed, the solutions are filtered into a clean 96-well plate utilizing a vacuum manifold. The filtrate samples are diluted twice by a factor of 2 and 4 times using PBS pH7.4 buffer and then transferred to a 384-well plate for analysis by UHPLC-CLND (CLND: Chemiluminescence nitrogen detector). 5 µL of each sample was injected into the UHPLC-CLND for analysis.
Data processing and solubility determination: Agilent Chemstation software was used for all data acquisition and processing. Samples were detected and analyzed using UV 254 nm and CLND channels. UV 254 nm is used primarily to confirm the sample purity but in rare cases, is also used to quantify the concentration of compounds with no nitrogens. This requires the additional step of creating a compound specific calibration curve. The identity of CLND target peaks of each compound were confirmed by LCMS. Sample quantification was accomplished by CLND peak intensity, the calibration curve, and the number of nitrogen contained in the compound. Note that the number of nitrogens in a compound used for this quantification needs to be corrected for the actual configuration of the nitrogens in the compound. One calibration curve was used for up to one week period. The solubility determination and quantification was automatically processed in Microsoft Excel using scripts.

2. X-ray co-crystal structure of CDK8:Cyclin C in complex with compound 22.

Protein production and crystallization

Human CDK8 and Cyclin C were cloned and expressed as previously described[2]. Protein production and purification were essentially performed according to previously published protocols[1]. Co-crystals at a protein concentration of 11.3 mg / ml with 1 mM compound **22** were obtained in 20% PEG 3350 and 0.20 M sodium formate at 20°C (hanging-drop) and shock-frozen with 25% ethylene glycol as cryoprotectant.

Structure determination

Diffraction data were collected at the Swiss Light Source (SLS, Villigen, Switzerland), beamline X06SA / PXI, with a wavelength of 0.9996Å at 100K, and processed using XDS[3] and SCALA/Collaborative Computational Project, number 4 (CCP4)[4]. Structures were solved by molecular replacement (MOLREP),[5] subsequent model building and refinement (including TLS refinement) was performed with COOT[6], Refmac[7], and Phenix[8]. The R_{free} validation was based on a subset of about 3.6% of the reflections omitted during refinement. Waters were included at stereochemically reasonable sites. Final refinement cycles led to a model with R_{work} value 17.3% and R_{free} value 22.7%. The refined model includes CDK8 residues 1-362 and Cyclin C residues 1-264. Additional residues from the expression construct were visible at the amino terminal ends of both proteins. All main-chain angles of non-glycine residues fall into the conformationally most favored (97.7%) or additionally allowed (2.3%) regions of the Ramachandran plot. The orientation of the methylamide of the compound likely exists in both orientations, separate by 180° rotation about the bond to the core. In the displayed orientation, a hydrogen bond is made to Lys252. In the flipped orientation, a hydrogen bond could be made to a molecule to formate observed to bind at partial occupancy in the active site. In this orientation, the apical methyl group would clash with the observed position of the lysine sidechain, necessitating its movement. The electron density most strongly supports the modeled conformation; however, it is the higher energy conformer, relative to the presumed solution state of **22**. The resolution of the data leaves open the possibility that the formate molecule and compound are not present simultaneously.

Table S1: Crystallographic data and refinement statistics.

PDB code	5CEI
Data collection[a]	
Beamline	SLS PXI/X06SA
Wavelength (Å)	0.99996
Detector	PILATUS 6M
Resolution range (Å)	83.72 – 2.24
No. unique reflections[a]	40,329 / 9,552
No. total reflections	167,682 / 39,718
Redundancy	4.2 / 4.2
Completeness (%)	98.4 / 98.6
Mean I/σ_I	5.7 / 2.3
R_{sym} (%)[b]	10.3 / 44.9
Refinement	
Resolution range (Å)	47.8-2.24 (2.45-2.24)
No. reflections (R_{free} set)	1,402
R_{cryst} , R_{free} (%)	17.3, 22.7
No. protein, ligand atoms in ASU	5,140
No. water molecules in ASU	181
Rmsd from ideal values	
Bond lengths (Å)	0.007
Bond angles (°)	0.975
Wilson B-factor for data (Å^2)	36.8
Mean non-water temperature factor (Å^2)	48.1
Ramachandran plot analysis	
Favored region (%)	97.7
Allowed region (%)	2.3
Outlier region (%)	0

[a]Data for highest resolution bin (2.45 - 2.24Å resolution) presented after the slash.

[b]$R_{sym} = (\Sigma |I - \langle I \rangle| / \Sigma I)$, where $\langle I \rangle$ is the average intensity for multiple measurements.

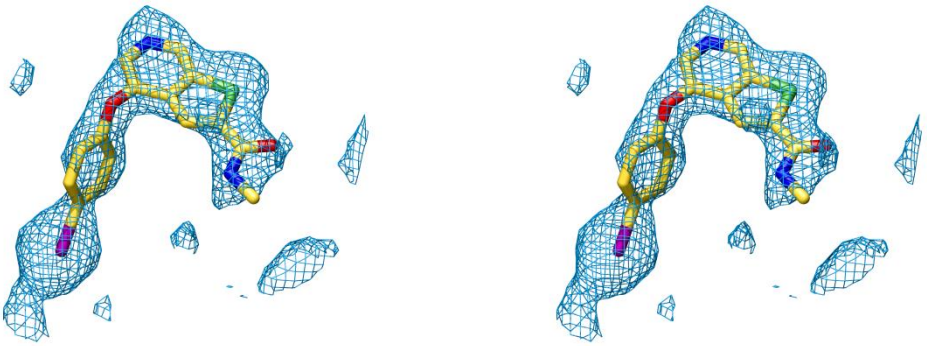


Figure S1: Divergent eye stereo diagram of the unbiased F_o-F_c difference electron density map of the structure of **22** bound at the CDK8 active site. Final inhibitor coordinates are shown, but the map was calculated prior to the addition of the inhibitor to the model and contoured at 2.3σ. The small additional density peak just below the iodine atom likely corresponds to a partially occupied solvent molecule. That portion of the density is negligible at 3σ.

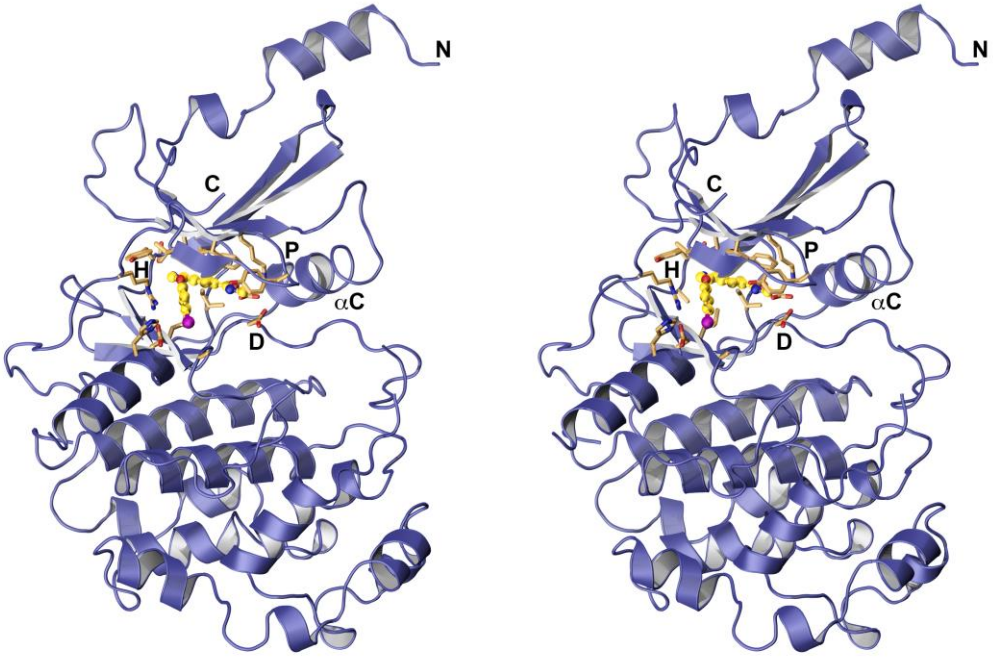


Figure S2: Crystal structure of the complex of CDK8:Cyclin C and **22**. Divergent eye stereo diagram of the CDK8 kinase with **22** (gold carbons) bound at the active site (beige carbons). Key structural features of the hinge (H), DMG loop (D), C helix (αC), P loop (P) as well as the amino (N) and carboxy (C) termini of the protein are labeled. Cyclin C was omitted for clarity but is a protein of roughly equivalent size to CDK8 that binds on the right face of the kinase (relative to this view), forming interactions with CDK8 from the N-terminus to the loop just below the αC helix.

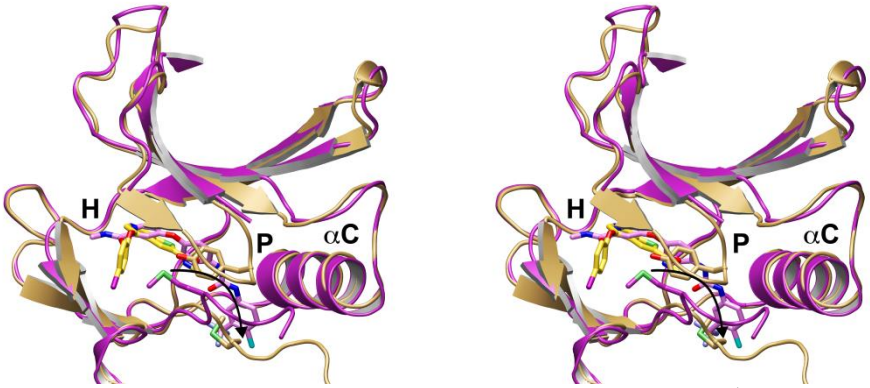


Figure S3: Divergent eye stereo diagram of the DMG-in complex with **22** (sand and gold) and the DMG-out complex[1] (accession code 3RGF) with sorafenib (purple and pink). For clarity, structural subregions of the active site are marked by single letters (H = hinge, P = P loop, αC = C helix) and only two amino acids are shown (Met174, of the DMG motif, and Tyr32 , which folds under the P loop in the complex with **22**.)

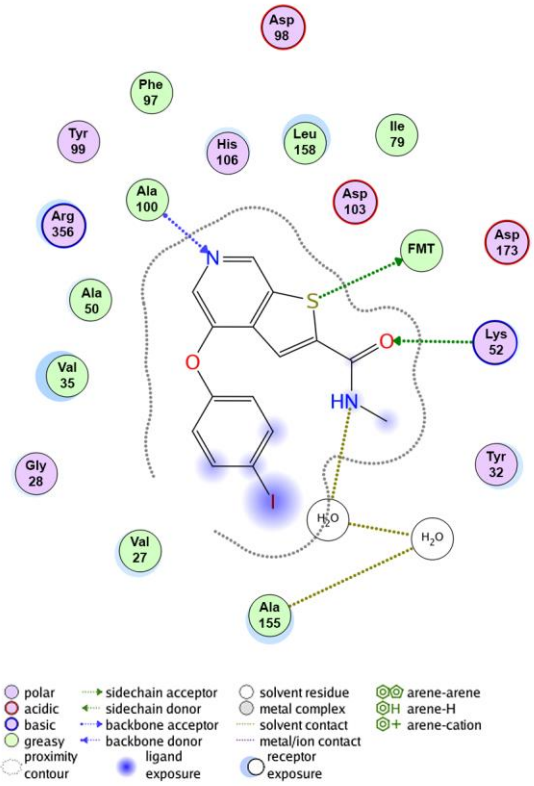


Figure S4: Binding site interactions of **22** with CDK8. The dashed line describes the interaction surface of the protein, drawn at a 0 kcal/mol energy isocontour. FMT represents the formate ion located in the active site. Hydrogen bonding interactions with it and the residues Ala100 and Lys52 are indicated with dashed lines.

3. HCT116 Proliferation Assays

Cell Culture: HCT116 cells (ATCC Catalog# CCL-247) are maintained in medium containing RPMI 1640, 10% FBS, 2 mM L-Glutamine, and 10 mM HEPES buffer pH 7.2. All cell culture reagents were purchased from Invitrogen/Gibco. Cells are cultured at 37°C at 5% CO_2 and split/maintained as recommended by ATCC.

Assay Procedure: Cells are harvested and plated into sterile cell culture treated 384 well plates (Greiner catalog #781091) at a density of 1000 cells/well in 50 ul culture medium containing either 10% FBS (High FBS) or 0.625% FBS (Low FBS) and placed in a 37°C at 5% CO_2 incubator overnight. The following day, test compounds are serially diluted in dimethyl sulfoxide (DMSO) and added to cells in culture medium (final DMSO concentration 0.5%, final assay volume 50 ul). Assay plates are then incubated for 72 hour at 37°C and 5% CO_2. After 72 hours, 25 ul of reconstituted Promega Cell Titer-Glo reagent (Promega Catalog# G7572) is added to all wells. Plates are then read on a Perkin Elmer Envision Multi-label plate reader using luminescence mode. Percent inhibition of proliferation by varying concentrations of test compounds is calculated relative to untreated controls. EC50 values are calculated using the 4 parameter logistic nonlinear regression dose-response model.

Cell proliferation assays: Cell proliferation was quantified by either CellTiter-Glo (Promega) or in real time using IncuCyte analysis (Essen Bioscience). For IncuCyte analysis, cells were imaged by phase every 4 hours and cell proliferation was quantified by percent confluence.

4. Immunoblot

Antibodies used for immunoblot: CDK8 (Santa Cruz Biotechnology; sc-1521 and Bethyl; A302-501A), CYCLIN C (Bethyl; A301-989A), p727-STAT1 (Cell Signaling; 8826), STAT1 (BD Biosciences; 610115), and ACTIN (MP Biomedicals; 691002). A custom rabbit CDK19 antibody was described previously.[9]
Phospho-STAT1 analysis: HCT116 cells were treated with 0, 1 or 10 ng/ml IFN-γ (R&D Systems). Cdk8 kinase inhibitor was added at 2 μM. Cell lysates were harvested 24 hours after treatment and processed for immunoblot.

5. CRISPR knockout of Cdk8 and Cdk19

Custom guide RNAs (gRNA) to human Cdk8 (target sequences: gRNA-1: gtttaagattgtttacctgt and gRNA-2 gaagccagttcagttacctc) and human Cdk19 (target sequences: gRNA-1: ccttcctgtcactgtgagaa and gRNA-2: tgcctccaaacaagaagcca), were cloned into pLKO2 (Sigma; SHC201) based on published design strategy.[10]
Cdk8 and Cdk19 knockout cell lines were generated by transiently co-transfecting pTKneo-Cas9 and gene specific gRNAs using Lipofectamine LTX (Invitrogen). Untransfected cells were eliminated with a brief 24 hour Puromycin treatment (2 μg/ml) two days after transfection. To isolate clonal populations, cells were single cell FACS sorted 6 days after transfection into 96 well plates (one cell per well). Subsequent clones were screened by PCR to identify gene deletions. Genomic DNA was isolated from cells growing in 96 well plates. Cells were washed 1X with HBSS, lysed in 20 μl of buffer (Sigma L3289), incubated at 75°C for 15 minutes, and neutralized with 180 μl of buffer (Sigma N9784). 2.5 ul of DNA was used for PCR (Phusion GC; NEB). Human Cdk8 oligos (tggtgatcgcagaacattgtg and ttctacagttggccttgtatc) were used to identify gene deletion events and oligos (taggatgccactgaattgtac and aagatagcatattcgtcctcc) were used to determine zygosity. Human Cdk19 oligos (cagatatacctaagcacatag and agctataagggtaaaatgatg) were used to identify gene deletion and oligos (cagatatacctaagcacatag and ttctgtgcattaacaatatg) were used to determine zygosity. Cdk8/Cdk19 double knockout cells were generated sequentially (Cdk8 followed by Cdk19). Genotypes were confirmed by immunoblot.

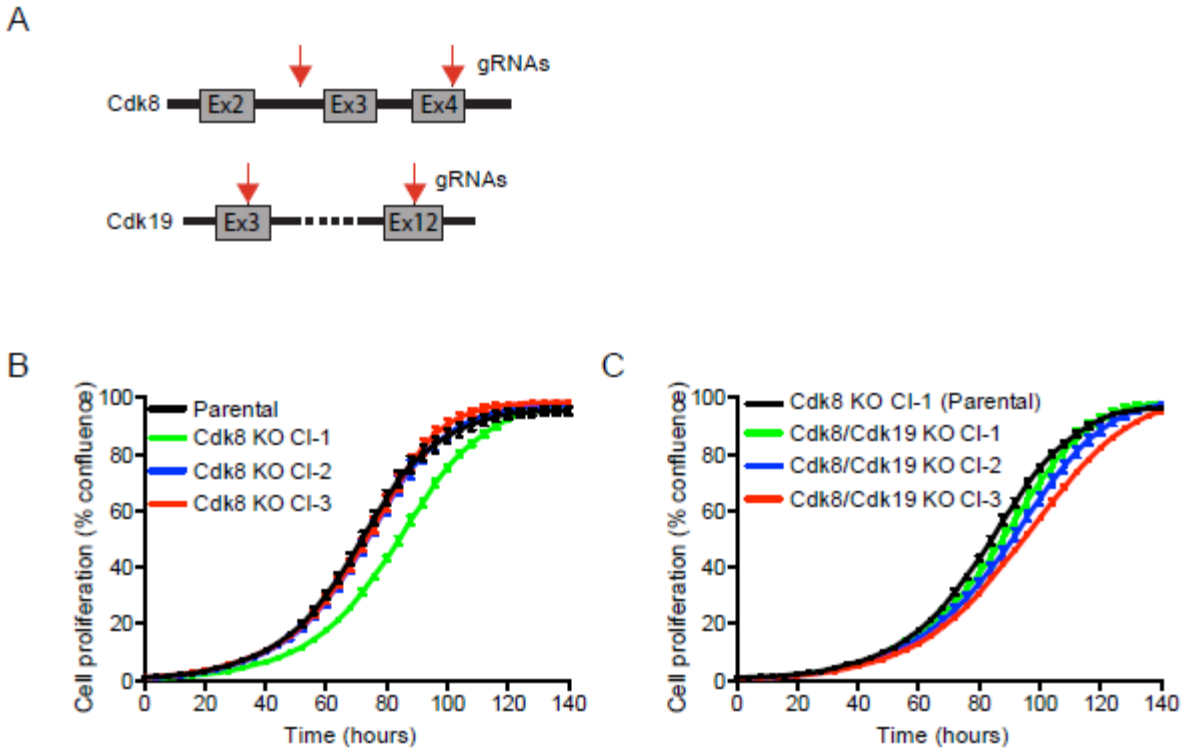


Figure S5: A) Targets of CRISPR guide RNAs targeting deletion of CDK8 and CDK19 in this study. B) Growth curve of three independent clones lacking CDK8. C) Growth curve of three independent CDK8/CDK19 clones along with that of the CDK8 clone which they were derived from.

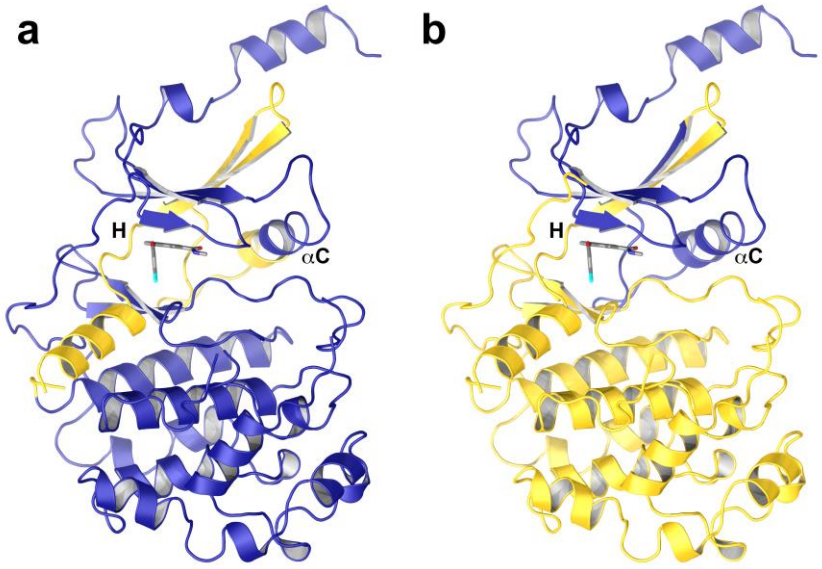


Figure S6: CRISPR deletions of CDK8 and CDK19 in this study. Kinase domain of CDK8 (blue) with residues excised by CRISPR colored gold. a) CDK8 CRISPR construct, b) CDK19 CRISPR construct (displayed on a structure of CDK8, as there is no known structure of CDK19, and the two proteins share a high degree of sequence identity)

Both CRISPR constructs would completely eliminate kinase activity and inhibitor binding for the kinases. In the case of the CDK8 construct, part of the N-lobe and the C helix (αC) as well as all of the hinge region (H) are deleted. In the case of the CDK19 construct, the majority of the protein is deleted.

6. Kinase Selectivity Data

Experimental details: Compounds were evaluated for selectivity of kinase inhibition in biochemical assays with a panel of recombinant kinases at Life Technologies (Madison, WI). Compounds were assayed at a single concentration with N = 2. Assay formats included Z-Lyte activity assays, Adapta activity assays, and LanthaScreen ATP-site binding assays. For activity assays, ATP concentrations were set at Km values. The Z-Lyte assays use doubly-labeled FRET peptide substrates. Phosphorylation of the substrate renders it insensitive to cleavage by a site specific protease. FRET between the fluorophores in the intact substrate yields a fluorescent signal. Thus, substrate phosphorylation reduces the assay signal, and inhibition by compound increases it. The Adapta assays monitor kinase activity through ADP generation. After the kinase reaction, the ADP is detected by competition with an AlexaFluor647-labeled ADP tracer for binding to Europium-labeled anti-ADP. The ADP produced by the kinase reaction reduces the TR-FRET signal generated by binding of the ADP tracer to the antibody, and inhibition by compounds increases it. Finally, the LanthaScreen assays are competitive ATP-site binding assays. Binding of an AlexaFluor647-labeled tracer to a tagged kinase is detected by TR-FRET with a Europium-labeled anti-Tag antibody. Compound competes with the labeled tracer for binding to the kinase, and consequently, binding of the compound to the kinase reduces the TR-FRET signal. Data are expressed as percent inhibition of the kinase activity or binding interaction. The specific kinases were chosen to provide a broad representation of the kinome. For the initial compounds **1** and **2**, the size of the kinase panel was modest, but it was greatly increased for the optimized compounds **13** and **32**. Compounds **1** and **2** were assayed at a concentration of 0.1 μM, while compounds **13** and **32** were assayed at a concentration of 1 μM. All values listed are percent inhibition.

KINASE	Compound **2**	Compound **1**
Abl	-2.35	-4.65
Aurora_A	12.1	5.25
B-Raf	20.8	21.3
CDK1/cyclinB	11.25	8.8
CDK2/cyclinA	11.7	2.15
CDK5/p25	2.4	1.75
CDK7/cyclinH	6	4.45
CDK8/cyclinC	104.65	104.7
CDK9/cyclinT1	58.9	43.5
CK1_gamma1	64.25	51.55
CK2_alpha1	47.25	22.55
CSF1R	28	9.1
Cot	72.7	39.4
DDR1	-2.55	-2.35
DMPK	69.35	61.15
DYRK1A	47.8	24.1
ERK2	4.2	-1.15
EphA3	-3.2	-0.6
EphB4	13.6	7.45
FGFR1	-8.65	-7
GSK3_alpha	96	90.05
HIPK4	82.7	98.3
Haspin	23	0.8
KDR	7.35	6.85
KHS1	-5.9	1.95
Kit	14.2	7.9
LIMK2	1.85	0.85
Lck	18.85	14.05
Lyn	11.3	9.3
MAP4K4	54.75	45.5
MAPKAPK2	1.85	-0.4

MEK1	13	7.95
MYLK(smMLCK)	7.6	-2.3
Mink1	39.45	3.1
MuSK	0.85	-0.2
NEK4	88.95	47.05
PASK	92.6	61.45
PDGFR_alpha	7.25	3.85
PI3K-A	23.55	14.15
PLK2	79.7	8.05
RSK4	-1.4	-1.75
Ret	7.4	6.5
STK16	80.55	22.35
STK33	3.4	-0.75
Srm	-5.65	0.95
TAO1	5.95	8.3
Tie2	2.9	1.25
TrkC	8	5.95
ZAK	4.45	3.05
ZIPK	25.45	5.1
p38_alpha	12.25	-3.7

	compound 13	compound 32
CDK8/cyclinC	104.00	103.80
ACVR1B	-3.90	4.00
AKT1	1.40	3.10
AKT2	-2.50	4.90
ALK2	-2.70	5.50
ARK5	-0.30	-13.80
ASK1	1.90	4.50
Abl	-1.10	9.60
Aurora_A	0.80	10.70
Aurora_B	4.00	36.20
Axl	-0.80	4.00
B-Raf	16.40	19.30
BTK	-1.50	8.10
Blk	-6.80	4.30
Bmx	12.60	5.00
BrSK1	-2.50	3.50
Brk	-6.40	3.90
CAMKK1	9.20	-4.10
CAMKK2	10.00	8.60
CDK1/cyclinB	9.30	27.90
CDK2/cyclinA	26.30	17.90
CDK5/p25	18.40	18.50
CDK7/cyclinH	17.10	8.10
CDK9/cyclinT1	55.60	82.00
CHK1	7.10	-4.20
CHK2	-1.00	3.50
CK1_alpha1	3.80	36.50
CK1_delta	9.30	31.80
CK1_epsilon1	8.20	58.40
CK1_gamma1	33.90	53.80
CK1_gamma2	38.60	75.80

CK2_alpha1	4.20	1.90
CLK1	7.00	10.90
CLK2	12.40	36.60
CLK3	-1.60	11.90
CSF1R	8.80	47.30
CSK	-2.30	15.70
CaMKI	7.00	10.10
CaMKII_beta	-4.20	6.90
CaMKI_delta	2.40	6.40
CamKII_alpha	5.90	1.40
CamKIV	-2.50	15.10
Cot	8.70	28.30
DAPK1	7.80	3.40
DCAMKL2	-3.50	3.00
DDR1	-9.70	7.00
DMPK	16.10	72.90
DRAK1	9.40	33.80
DYRK1A	9.50	9.50
DYRK3	12.80	8.90
DYRK4	1.50	7.40
EGFR	2.80	1.80
ERK2	3.30	8.10
EphA1	0.20	4.60
EphA3	-0.90	8.90
EphA7	5.30	3.60
EphA8	0.20	10.60
EphB1	1.50	13.20
EphB3	3.50	9.30
ErbB2	-4.70	-1.60
ErbB4	-1.40	7.80
FAK	3.50	11.10
FGFR1	1.00	5.20
FGFR3	5.10	0.10
FGFR4	0.40	17.60
Fes	4.30	12.10
Fgr	2.50	60.70
Flt1	4.50	10.60
Flt3	2.60	40.50
Flt4	16.40	14.70
Frk	6.60	6.70
GRK2	4.30	6.60
GRK3	-0.50	2.70
GRK5	5.00	0.60
GRK6	10.30	16.40
GSK3_alpha	63.30	52.90
GSK3_beta	60.50	66.10
HIPK1	19.80	32.80
HIPK2	24.60	26.30
HIPK4	60.60	23.90
Hyl	2.30	1.20
IGF1R	-1.30	12.90
IKK_alpha	13.70	15.60
IKK_beta	-0.70	16.50
IKK_epsilon	3.00	-1.30
IRAK1	4.90	11.30
IRAK4	1.60	-0.60
IRR	1.90	2.70
InsR	3.00	7.70
Itk	5.90	5.00

JAK1	2.10	-1.50
JAK2	0.20	2.30
JAK3	0.20	2.50
JNK1_alpha1	4.90	9.50
JNK2	8.00	19.90
JNK3	4.50	5.00
KDR	1.00	10.40
KHS1	-5.30	11.10
Kit	13.80	7.30
LIMK1	2.30	10.80
LRRK2	11.20	25.60
LTK	3.00	-3.50
Lck	4.00	66.20
Lyn	5.50	31.90
MAP4K4	16.40	26.40
MAPKAPK2	11.80	4.00
MAPKAPK3	-7.00	11.80
MARK1	-2.30	1.10
MARK3	-3.00	4.80
MEK1	1.10	25.60
MEK3	15.80	13.30
MEKK2	-3.50	11.30
MELK	10.30	18.60
MKK6	-0.30	19.00
MLK1	-1.50	18.00
MLK2	5.90	5.70
MRCK_alpha	5.10	35.70
MSK1	2.20	8.00
MSSK1	1.50	4.90
MST1	1.90	1.60
MST2	5.50	13.20
MST3	0.20	32.90
MST4	-8.80	0.50
MYLK(smMLCK)	5.20	11.60
MYLK3(caMLCK)	14.40	33.80
Mer	-2.40	7.20
Met	2.50	0.20
Mink1	-8.50	31.40
MuSK	1.10	1.90
NEK1	4.90	6.40
NEK4	5.40	30.00
NEK6	12.30	5.00
NEK9	-6.70	-5.30
NLK	19.00	19.10
PAK1	-3.50	18.20
PAK3	1.80	7.60
PAK4	5.00	8.70
PAK6	6.50	5.30
PASK	20.80	34.60
PDGFR_alpha	1.80	11.00
PDK1(direct)	8.70	-3.90
PI3K-A	0.70	11.30
PI3K-G	-18.60	1.40
PIM1	-1.60	-2.10
PKA	3.10	5.30
PKC_alpha	-0.40	19.60
PKC_beta1	2.00	4.80
PKC_delta	0.70	-1.40
PKC_epsilon	5.00	9.10

PKC_eta	8.30	7.90
PKC_theta	1.20	15.60
PKC_zeta	-0.70	-1.20
PKD1	7.90	15.10
PKG1_alpha	6.80	9.80
PLK1	11.70	21.50
PLK2	8.80	-2.90
PLK3	11.50	13.80
PRAK	2.60	16.20
PRK1	11.80	15.10
PhK_gamma1	-1.80	-2.10
PhK_gamma2	-1.50	5.00
PrKX	43.50	26.50
RIPK2	12.50	8.20
ROCK1	2.60	13.10
ROCK2	0.50	4.50
Ret	4.50	12.60
Ron	1.90	4.00
Ros	1.40	4.20
Rse	5.80	6.80
Rsk1	7.30	11.90
Rsk2	3.90	5.20
Rsk3	2.50	6.90
SGK1	9.50	10.90
SGK2	6.50	11.20
SGK3	-1.80	7.30
SLK	15.80	13.30
SPHK1	-14.20	-6.80
SRPK1	-0.70	7.50
STK16	15.60	6.20
STK33	3.30	14.60
Src	-0.20	16.40
Srm	9.90	3.00
Syk	-7.50	23.40
TAK1-TAB1	12.20	51.80
TAO1	4.50	-2.00
TBK1	2.90	14.10
TEC	0.10	0.40
TNK2	7.50	37.30
TSSK1	6.50	11.50
TTK	11.10	54.40
TYK2	-1.20	-0.10
Tie2	-3.40	-1.40
TrkA	-11.50	6.30
TrkB	-4.20	11.30
TrkC		
WEE1	-0.80	6.40
WNK2	2.80	14.60
YSK1	2.90	-4.90
Yes	3.40	13.60
ZAK	1.50	5.70
ZAP-70	2.50	1.10
ZIPK	3.00	4.70
eEF-2K	-1.20	4.50
mTOR	-1.90	9.90
p38_alpha		
p38_alpha(direct)	-2.40	31.70
p38_beta	8.00	11.60

p38_delta	3.30	18.00
p38_gamma	-2.10	27.90
p70S6K	-2.30	2.10

7. Table S2: Complete in vitro profile obtained for tool compounds 13 and 32.

	CDK8 IC_{50} (nM)	CDK9 IC_{50} (nM)	MDCK permeability AB x 10^{-6} cm/sec	MDCK permeability BA x 10^{-6} cm/sec	TPSA ($Å^2$)	PPB % bound (H/R/M)	kinetic solubility (μM)
13	6.2	2200	15.7	5.4	65	98/97.5/97.8	6.3
32	1.5	1100	18.2	17.2	91	98.1/97.7/95.9	5.5

REFERENCES

(1) George, D.; Friedman, M.; Allen, H.; Argiriadi, M.; Barberis, C.; Clabbers, A.; Cusack, K.; Dixon, R.; Fix-Stenzel, S.; Gordon, T.; Janssen, B.; Jia, Y.; Moskey, M.; Quinn, C.; Salmeron, J-A.; Wishart, N.; Woller, K.; Yu, Z. Discovery of Thieno[2,3-C]Pyridines as Potent COT Inhibitors. Bioorg. Med. Chem. Lett. 2008, 18, 4952-4955.

(2) Schneider, E. V. *et al.* The Structure of CDK8/CycC Implicates Specificity in the CDK/Cyclin Family and Reveals Interaction with a Deep Pocket Binder. *J. Mol. Biol.* 412, 251–266 (2011).

(3) Kabsch, W. Integration, scaling, space-group assignment and post-refinement. *Acta Cryst.* D66, 133-144 (2010).

(4) Dodson, E. J., Winn, M. & Ralph, A. Collaborative Computational Project, number 4: providing programs for protein crystallography. *Meth. Enzymol.* 277, 620–633 (1997).

(5) Vagin, A. & Teplyakov, A. Molecular Replacement with MOLREP. *Acta Cryst.* D66, 22-25 (2010).

(6) Emsley, P. & Cowtan, K. Coot: Model-building Tools for Molecular Graphics. *Acta Cryst.* D60, 2126-2132 (2004).

(7) Vagin, A.A. et. al. REFMAC5 dictionary: Organization of prior chemical knowledge and guidelines for its use. *Acta Cryst.* D60, 2184–2195 (2004).

(8) Adams, P.D. et. al. PHENIX: A comprehensive Python-based system for macromolecular structure solution. *Acta Cryst*. D66, 21-21 (2010).

(9) McCleland, M.L.; Soukup, T.M.; Liu, S.D.; Esensten, J.H.; de Sousa, E.; Melo, F.; Yaylaoglu, M; Warming, S.; Roose-Girma, M.; Firestein, R. Cdk8 deletion in the Apc(Min) murine tumour model represses EZH2 activity and accelerates tumourigenesis. J Pathol. **2015**, *237, in press*.

(10) Mali, P.; Yang, L.; Esvelt, K. M.; Aach, J.; Guell, M.; DiCarlo, J. E.; Norville, J. E.; Church, G. M. RNA-Guided Human Genome Engineering via Cas9. Science, 2013, 339, 823–826.