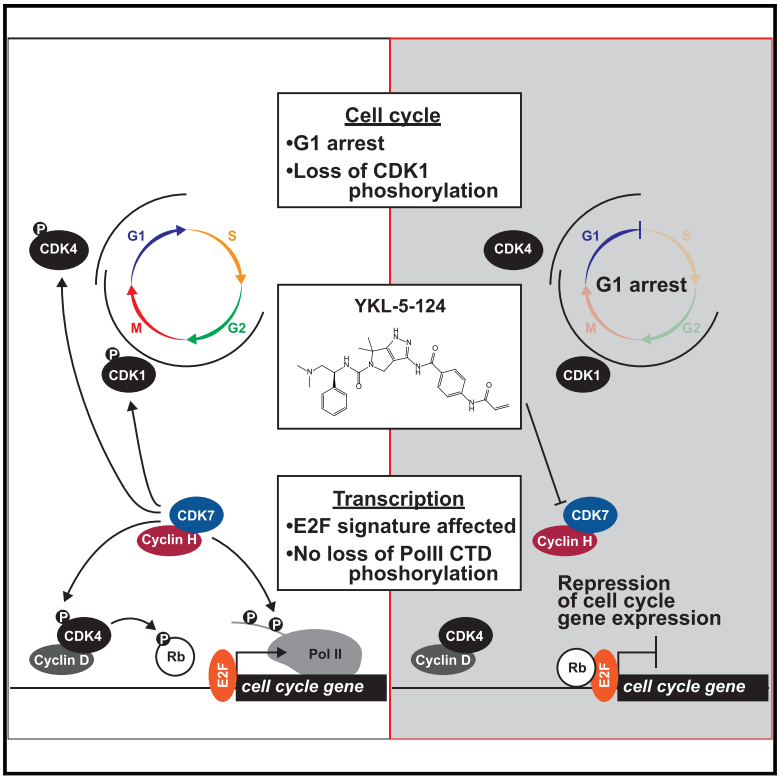
Article

Cell Chemical Biology

Development of a Selective CDK7 Covalent Inhibitor Reveals Predominant Cell-Cycle Phenotype

Graphical Abstract



Highlights

- Development of YKL-5-124 a potent, selective and covalent CDK7 inhibitor
- Selective CDK7 inhibition results in cell-cycle inhibition rather than apoptosis
- YKL-5-124 exhibited little effect on RNA polymerase II phosphorylation status

Authors

Calla M. Olson, Yanke Liang, Alan Leggett, ..., Charles Y. Lin, Nicholas Kwiatkowski, Nathanael S. Gray

Correspondence

nicholasp_kwiatkowski@dfci.harvard.edu (N.K.), nathanael_gray@dfci.harvard.edu (N.S.G.)

In Brief

Olson et al. describe the development and characterization of YKL-5-124, a potent, selective, and covalent CDK7 inhibitor. YKL-5-124 displays biochemical and cellular selectivity for CDK7 over CDK12/13, structurally related kinases. CDK7 inhibition by YKL-5-124 induces a strong cell-cycle arrest and a surprisingly weak effect on RNA Pol II phosphorylation.

Olson et al., 2019, Cell Chemical Biology *26*, 1–12
June 20, 2019
https://doi.org/10.1016/j.chembiol.2019.02.012

CellPress

Please cite this article in press as: Olson et al., Development of a Selective CDK7 Covalent Inhibitor Reveals Predominant Cell-Cycle Phenotype, Cell Chemical Biology (2019), https://doi.org/10.1016/j.chembiol.2019.02.012

Cell Chemical Biology
Article

Development of a Selective CDK7 Covalent Inhibitor Reveals Predominant Cell-Cycle Phenotype

Calla M. Olson,[1,2,3,4,12] Yanke Liang,[1,2,12] Alan Leggett,[1,2] Woojun D. Park,[5] Lianbo Li,[6] Caitlin E. Mills,[7] Selma Z. Elsarrag,[5] Scott B. Ficarro,[1,2,8] Tinghu Zhang,[1,2] Robert Düster,[9] Matthias Geyer,[9] Taebo Sim,[10,11] Jarrod A. Marto,[1,2,8] Peter K. Sorger,[7] Ken D. Westover,[6] Charles Y. Lin,[3,4,5] Nicholas Kwiatkowski,[1,2,*] and Nathanael S. Gray[1,2,13,*]

[1]Department of Cancer Biology, Dana-Farber Cancer Institute, Boston, MA 02215, USA
[2]Department of Biology Chemistry and Molecular Pharmacology, Harvard Medical School, Boston, MA 02215, USA
[3]Therapeutic Innovation Center (THINC@BCM), Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA
[4]Verna & Marrs McLean Department of Biochemistry & Molecular Biology, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA
[5]Department of Molecular & Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA
[6]Departments of Biochemistry and Radiation Oncology, The University of Texas Southwestern Medical Center at Dallas, Dallas, TX 75390, USA
[7]Laboratory of Systems Pharmacology, Harvard Medical School, Boston MA 02115, USA
[8]Blais Proteomics Center, Dana-Farber Cancer Institute, Boston, MA 02115, USA
[9]Institute of Structural Biology, University of Bonn, Sigmund-Freud-Strasse 25, 53127 Bonn, Germany
[10]Chemical Kinomics Research Center, Korea Institute of Science and Technology (KIST), Seoul 02792, Korea
[11]KU-KIST Graduate School of Converging Science and Technology, Korea University, Seoul 136-701, Korea
[12]These authors contributed equally
[13]Lead Contact
*Correspondence: nicholasp_kwiatkowski@dfci.harvard.edu (N.K.), nathanael_gray@dfci.harvard.edu (N.S.G.)
https://doi.org/10.1016/j.chembiol.2019.02.012

SUMMARY

Cyclin-dependent kinase 7 (CDK7) regulates both cell cycle and transcription, but its precise role remains elusive. We previously described THZ1, a CDK7 inhibitor, which dramatically inhibits superenhancer-associated gene expression. However, potent CDK12/13 off-target activity obscured CDK7s contribution to this phenotype. Here, we describe the discovery of a highly selective covalent CDK7 inhibitor. YKL-5-124 causes arrest at the G_1/S transition and inhibition of E2F-driven gene expression; these effects are rescued by a CDK7 mutant unable to covalently engage YKL-5-124, demonstrating on-target specificity. Unlike THZ1, treatment with YKL-5-124 resulted in no change to RNA polymerase II C-terminal domain phosphorylation; however, inhibition could be reconstituted by combining YKL-5-124 and THZ531, a selective CDK12/13 inhibitor, revealing potential redundancies in CDK control of gene transcription. These findings highlight the importance of CDK7/12/13 polypharmacology for anti-cancer activity of THZ1 and posit that selective inhibition of CDK7 may be useful for treatment of cancers marked by E2F misregulation.

INTRODUCTION

Cyclin-dependent kinase 7 (CDK7) is a master regulator of cell-cycle progression and gene transcription (Fisher and Morgan, 1994; Devault et al., 1995; Glover-Cutter et al., 2009; Kelso et al., 2014). In the cytoplasm, CDK7 comprises the enzymatic core of the CDK-activating kinase (CAK), a trimeric complex that also includes CDK7's obligate binding partners, cyclin H and MAT1. CAK phosphorylates key cell-cycle CDKs (1, 2, and 4) leading to their full activation (Larochelle et al., 2007, 2012; Schachter et al., 2013; Bisteau et al., 2013; Gerber et al., 2008). Activation of CDKs by CAK plays an essential role in control over cell-cycle progression: (1) activation of CDK4 results in the phosphorylation of retinoblastoma protein and consequent G_1 progression (Sherr and Roberts, 1999), (2) CDK2 activation promotes progression from G_1 into S phase (Furuno et al., 1999; Satyanarayana and Kaldis, 2009), and (3) phosphorylation of CDK1 positively regulates the G_2/M transition. These results have been demonstrated both genetically and through the use of the "bumped hole" chemical genetic approach, which has enabled selective inhibition of a genetically engineered CDK7 mutant that is sensitive to a bulky ATP-like inhibitor (Larochelle et al., 2007). The predominant cell-cycle phenotype caused by inhibition of CDK7 (and the kinases that depend on CDK7 for activity) is accumulation of G_1 or G_2/M cells and a concomitant reduction in the number of S-phase cells (Larochelle et al., 2007; Ali et al., 2009; Wang et al., 2016; Kwiatkowski et al., 2014), showing that cell proliferation requires CDK7 kinase activity.

In the nucleus, the CAK module is a component of the general transcription factor, TFIIH, a multi-protein complex that is essential for RNA polymerase II (Pol II)-mediated transcription (Adamczewski et al., 1996; Glover-Cutter et al., 2009; Serizawa et al., 1995; Shiekhattar et al., 1995; Chen et al., 2015; Feaver et al., 1994). Phosphorylation of the Pol II C-terminal domain (CTD) at S2/S5/S7 residues is tightly regulated as a means to coordinate transcription initiation, elongation, and termination (Buratowski,

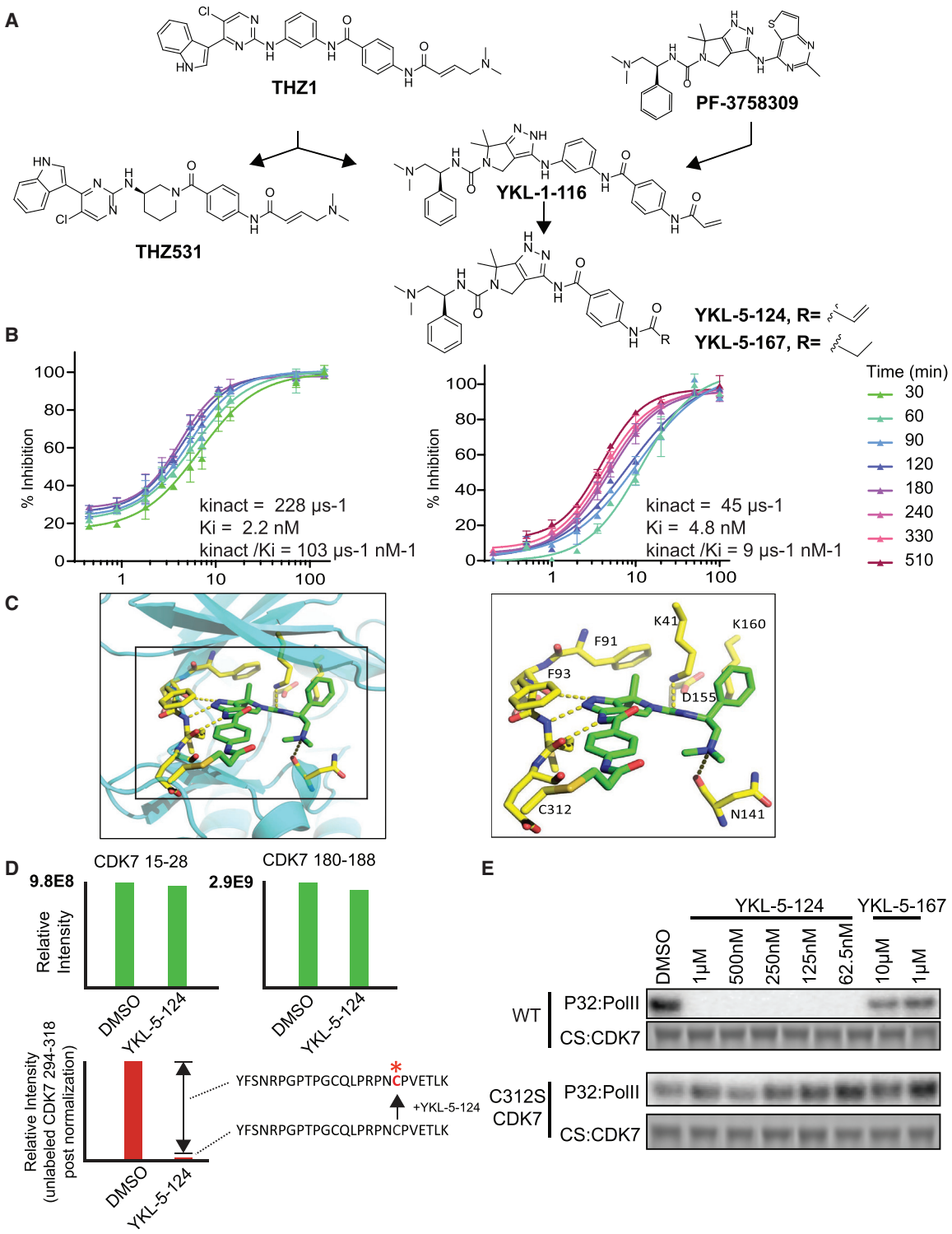


Figure 1. Biochemical Characterization of YKL-5-124

(A) Chemical structures of THZ1, THZ531, YKL-1-116, YKL-5-124, and YKL-5-167.

(B) Analysis of time- and dose-dependent inhibition of CDK7/CycH/MAT1 enzymatic activity by YKL-05-124 or THZ1 reveals a leftward shift in IC_{50} values over time. Kinetics parameters of enzyme inactivation were obtained using nonlinear least-squares regression. Error bars represent the standard deviation (SD) for duplicate measurements.

(legend continued on next page)

2009; Egloff and Murphy, 2008). However, the conclusion that CDK7 kinase activity is required for Pol II CTD phosphorylation, and thus global gene transcription has been controversial. Several reports find that CDK7 is non-essential in transcription both *in vivo* and *in vitro* (Kanin et al., 2007; Glover-Cutter et al., 2009; Serizawa et al., 1993). In contrast, other reports show that CDK7 inhibition reduces serine (S5 and S7) phosphorylation across the 52 heptad repeats constituting the Pol II CTD (Kwiatkowski et al., 2014; Akhtar et al., 2009; Glover-Cutter et al., 2009) with a concurrent reduction in global basal transcription (Kwiatkowski et al., 2014; Kelso et al., 2014; Poss et al., 2016); however, these experiments were mostly conducted using *in vitro* biochemical assays or with less selective chemical probes. Consistent with its role in transcription, CDK7 is also implicated in mRNA processing through the recruitment of capping enzymes, transcription activation through regulating the deposition of histone marks associated with active transcription, pause induction (Nilson et al., 2015, Ebmeier et al., 2017, Glover-Cutter et al., 2009), and pause release through phosphorylation of CDK9, the kinase subunit of the elongation complex pTEFb (Shiekhattar et al., 1995; Serizawa et al., 1995; Marshall et al., 1996). Together these results suggest a role for CDK7 in gene transcription; however, it remains unclear how CDK7 kinase activity ultimately affects steady-state mRNA levels (Marshall et al., 1996; Rodríguez-Molina et al., 2016; Larochelle et al., 2012).

Although the "bumped hole" approach to kinase inhibition has yielded key insights into CDK7's biological role, this approach requires engineering mammalian cell lines and is difficult to use broadly; it also does not provide a route to developing new drugs. Selective small molecules capable of selective inhibition of endogenous CDK7 are therefore required to further delineate CDK7 activities in cell-cycle progression and transcription. Previously, we reported the discovery of THZ1, a small-molecule kinase inhibitor that targets CDK7 by covalently labeling a cysteine (C312) that lies in the CTD of CDK7 and is remote from the active site (Kwiatkowski et al., 2014). We also found that THZ1 inhibited CDK12/13 via the same covalent mechanism of action (Kwiatkowski et al., 2014; Zhang et al., 2016). To separate these activities, we sought chemical probe molecules with a narrower selectivity profile. This has yielded THZ531, a derivative of THZ1 with the same phenylaminopyrimidine core scaffold, which inhibits CDK12/13 approximately 20 times more potently than CDK7 (Zhang et al., 2016; Gao et al., 2018). However, efforts to generate a CDK7-selective inhibitor using the THZ1 scaffold have not been successful.

Here, we describe the development and mechanistic validation of a selective, covalent CDK7 inhibitor, YKL-5-124, based on a previously unexplored CDK-targeting scaffold that originates from the PAK4 inhibitor PF-3758309 (Murray et al., 2010). We performed a series of structured biochemical and cell-based assays to validate a covalent mode of CDK7 inhibition by YKL-5-124 involving C312 modification and determined the compound's selectivity profile. We also employed YKL-5-124 as a tool compound to investigate the functions of CDK7 in regulating cell-cycle progression and gene transcription. Our data confirm the reported CAK activities of CDK7 in regulating cell-cycle progression, but suggest a non-essential role for CDK7 in regulating Pol II CTD phosphorylation and global basal gene expression.

RESULTS

YKL-5-124 Covalently Targets CDK7

Our initial efforts to develop a selective CDK7 inhibitor using THZ1 (Figure 1A) as a lead compound were not productive. We turned our attention to PF-3758309, a PAK4 inhibitor (Murray et al., 2010), since molecules from the series of compounds that gave rise to PF-3758309 are reported to have strong CDK7 inhibitory activity (Rudolph et al., 2015). We envisioned that the hybridization of the covalent warhead from THZ1 and the pyrrolidinopyrazole core from PF-3758309 could lead to more selective CDK7 inhibitors. The first compound in the new series, YKL-1-116 (Figure 1A) (Kalan et al., 2017), showed good selectivity for CDK7 but only moderate potency, and had minimal anti-proliferative effects on cancer cell lines (Kalan et al., 2017). Further iterative optimization has focused on tuning the acidity of the aminopyrazole core, gaining additional hydrophobic interactions with the side-chain residue, and optimizing the length and trajectory of the covalent warhead that targets C312. This focused medicinal chemistry campaign yielded YKL-5-124, a potent and selective covalent CDK7 inhibitor, and YKL-5-167, an inactive analog lacking the acrylamide reactive center and therefore incapable of forming a covalent bond with CDK7- C312 (Figure 1A).

Biochemical evaluation of YKL-5-124 using a fixed time-point, *in vitro* kinase assay indicated that YKL-5-124 inhibited CDK7/Mat1/CycH with an IC_{50} of 9.7 nM, while the other CDKs tested biochemically, CDK2 and CDK9, had IC_{50} values of 1,300 nM and 3,020 nM, respectively. To more precisely assay the contribution of covalent bond formation to compound inhibitory activity, we measured k_{inact}/K_i using a kinetic assay that monitors the shift in the electrophoretic mobility of a CDK7 peptide substrate following incubation with YKL-5-124 (Blackwell et al., 2009; Tan et al., 2017). YKL-5-124 and THZ1 displayed similar k_i values (1.9 nM and 2.1 nM, respectively) showing that they achieved nearly equivalent inhibition of CDK7 (Figures 1B, S1A, and S1B). However, YKL-5-124 exhibited a faster k_{inact} of 103 μs^{-1} nM^{-1} as compared with k_{inact} of 9 μs^{-1} nM^{-1} for of THZ1, demonstrating that YKL-5-124 covalently modifies CDK7 approximately 11-fold faster than THZ1 (Figures 1B, S1A, and S1B).

(C) Left: docking model of YKL-5-124 in the ATP-binding site of CDK7. CDK7 is shown as blue ribbons with key residues highlighted in yellow and YKL-5-124 is modeled as a green structure. Expansion of the box on the left is shown to the right, and key residues that make contact with YKL-5-124 are labeled and highlighted.

(D) Efficiency of labeling was estimated to be approximately 99% gauged by the reduction in signal of triply and quadruply charged YFSNRPGPTPGCQLPRPNCPVETLK ions (residues 294–318). The peptides VPFLPGDSDLDQLTR (residues 180–194) and LDFLGEGQFATVYK (residues 15–28) were used for normalization.

(E) *In vitro* kinase assays of exogenous tagged CDK7 protein immunoprecipitated with FLAG antibody after HeLa cells were treated with YKL-5-124 or YKL-5-167 as indicated. WT, wild-type; C312S, C312 to serine mutation.

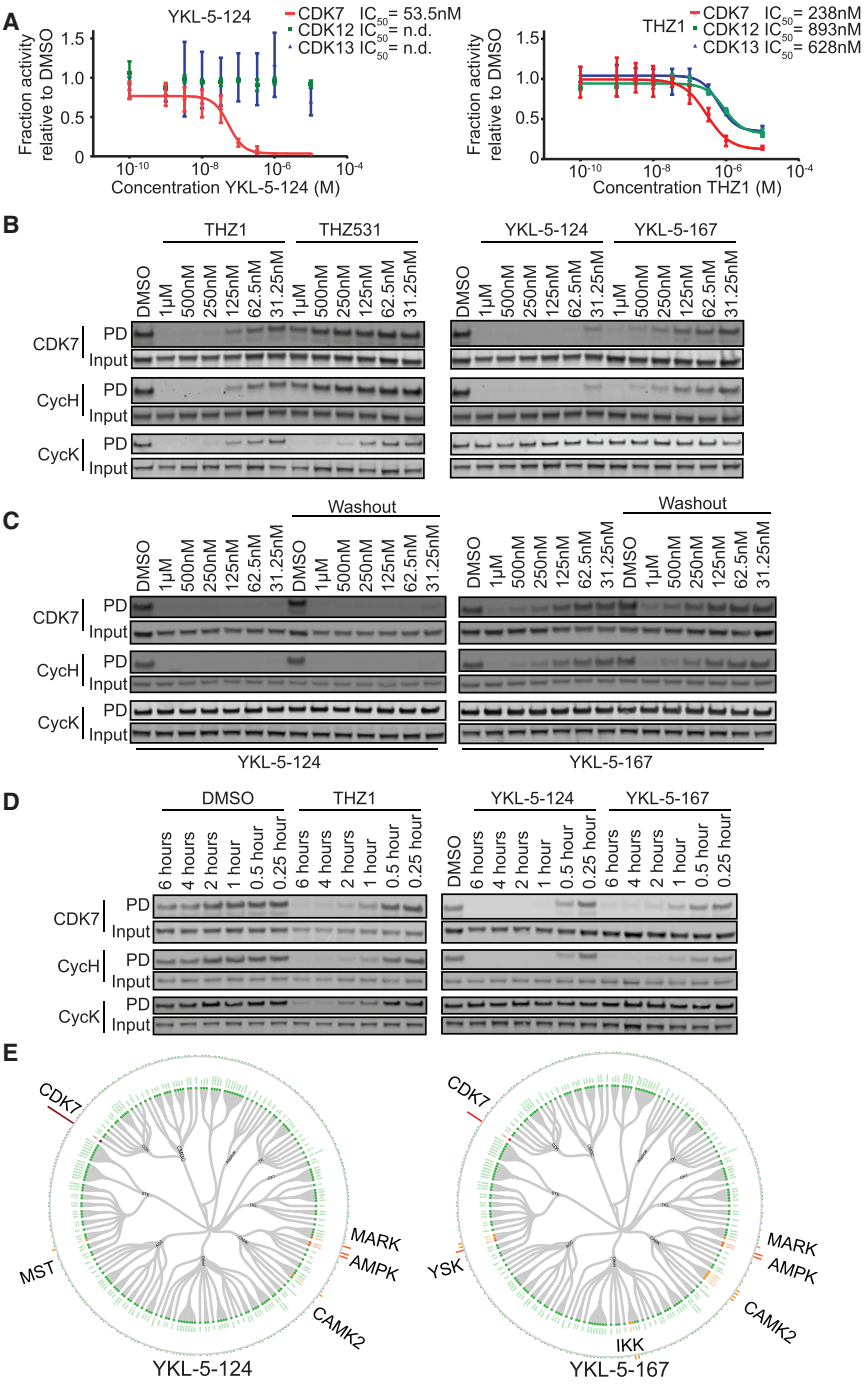


Figure 2. Cellular Engagement of CDK7 by YKL-5-124

(A) *In vitro* biochemical dose curves of CDK7, CDK12, and CDK13, with YKL-5-124 (left) and THZ1 (right) treatments. Error bars represent SD of triplicate experiments.

(B) Competitive pull-down assays in HAP1 cells treated with a concentration course of THZ1, YKL-5-124, or YKL-5-167 for 6 h. Western blots show the pull-down or input of CDK7, cyclin H, and cyclin K.

(C) Competitive pull-down assays in HAP1 cells treated for 6 h with a concentration course of YKL-5-124 or YKL-5-167 with or without washout after 3-h treatment.

(D) Competitive pull-down assays in HAP1 cells treated with DMSO, THZ1, or YKL-5-124 for the time indicated before lysing and treating with bioTHZ1 to pull down all available CDK7. PD, pull-down.

(E) KiNativ target engagement kinase tree showing kinases engaged over 65% with 1 μM YKL-5-124 treatment (left) or YKL-5-167 treatment (right) in Jurkat cells.

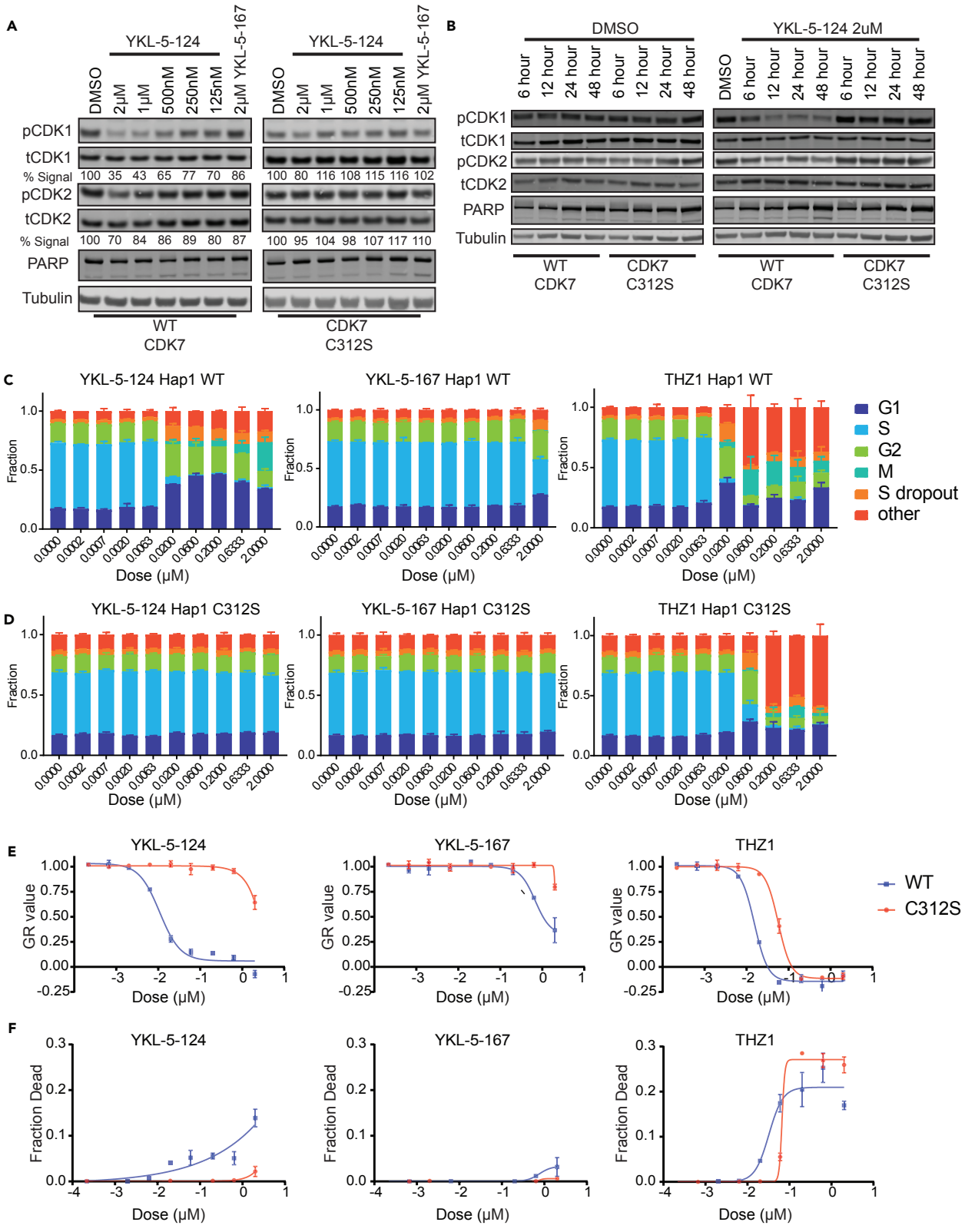
See also Table S1.

Molecular docking studies place the reactive acrylamide of YKL-5-124 near residue CDK7-C312, the same cysteine targeted by THZ1 (Figure 1C). These docking results are confirmed by X-ray structures of kinases other than CDK7 complexed with inhibitors containing a tetrahydropyrrolopyrazole core such as maternal embryonic leucine zipper kinase (MELK, PDB: 5MAG and 4BKY) and serine/threonine-protein kinase PAK 4 (4APP) wherein the same binding mode is observed (Klaeger et al., 2017; Canevari et al., 2013; Guo et al., 2012). The docking results additionally highlighted the possibility that YKL-5-124 can make two additional contacts with CDK7 at K41 and Q141, which may contribute to binding affinity and potency (Figure 1C). Mass spectrometry analysis of recombinant CDK7-cyclin H-MAT1 trimeric complex incubated with YKL-5-124 demonstrated covalent labeling of CDK7 with an efficiency of ~99% (Figures 1D, S1C, and S1D) and capillary electrophoresis-mass spectrometry (CE-MS) confirmed the site of labeling to be C312 (Figure S1E).

To determine whether YKL-5-124 is capable of inhibiting CDK7 in cells, we immunoprecipitated HeLa cells with wild-type (WT) or engineered CDK7 with the C312S mutation from vehicle- or YKL-5-124-treated cells and subjected the precipitates to *in vitro* kinase assays. YKL-5-124 blocked the *in vitro* phosphorylation of RNA Pol II substrate by immunopurified CDK7 at concentrations as low as 62.5 nM, whereas the reversible YKL-5-167 analog demonstrated little or no inhibition of phosphorylation at 10 μM (Figure 1E). Mutation of the reactive cysteine to a less nucleophilic serine (C312S) abolished the inhibitory activity of YKL-5-124 in the *in vitro* kinase assay (Figure 1E). Taken together, these data show that YKL-5-124 is a nanomolar covalent inhibitor that targets C312 on CDK7 with faster kinetics than THZ1 and 100-fold greater selectivity for CDK7 than CDK9 and CDK2.

YKL-5-124 Specifically Targets CDK7

To confirm that YKL-5-124 does not inhibit CDKs 12 and 13, we performed *in vitro* kinase reactions with CDK7, CDK12, and CDK13 at an ATP concentration of 1 mM (Figure 2A). We found



(legend on next page)

that YKL-5-124 had an IC_{50} of 53.5 nM for CDK7, yet displayed no inhibition of CDK12 or CDK13 at the concentrations tested. In comparison, we found that THZ1 was equipotent on CDKs 7, 12, and 13. To determine the cellular selectivity of YKL-5-124 for CDK7 versus CDK12/13, we used HAP1 and Jurkat cells. HAP1 is a near-haploid human cell line derived from the male chronic myelogenous leukemia (CML) cell line KBM-7; Jurkat is a human T lymphocyte cell line previously used to characterize THZ1 (Esslетzbichler et al., 2014; Andersson et al., 1995; Kwiatkowski et al., 2014). Lysates from cells treated with increasing concentrations of THZ1, YKL-5-124, or YKL-5-167 were treated with biotinylated THZ1 (bioTHZ1) to affinity purify CDK complexes, which were then analyzed by western blotting using antibodies against CDK7, CycH, and CycK, the obligate binding partner of CDK12/13. THZ1 bound to both CDK7 and CDK12/13 based on its ability to block pull-down of all three proteins (Figures 2B and S2); treatment of cells with YKL-5-124 as a competitor at a concentration of about 30 nM blocked pull-down of CDK7-cyclin H but had no effect on the pull-down of cyclin K-CDK12/13 (Figures 2B and S2A). YKL-5-124 displayed engagement of CDK7 3 h after compound was washed out of cells down to 125 nM, further confirming that YKL-5-124 functions in an irreversible manner (Figures 2C and S2B). In contrast, YKL-5-167, which cannot react with CDK7-C312, was 100-fold less effective as a competitor for CDK7-cyclin H in bioTHZ1 pull-down experiments and inactive following washout. To study the kinetics of target engagement, we treated cells with drug for 30 min to 6 h prior to pull-down. Treatment with 100 nM YKL-5-124 reduced CDK7-cyclin H binding to bioTHZ1 by >50% at 30 min, whereas THZ1 required twice as long, consistent with the faster k_{inact} of YKL-5-124 (Figures 2D and S2C). YKL-5-167 required 6 h to achieve equivalent engagement of CDK7 in HAP1 cells and was even less active in Jurkat cells, requiring 10-fold higher concentrations and longer times to function as a competitor (Figures 2D and S2C).

To assess target selectivity, we performed KiNativ profiling, an MS-based method that measures the ability of a test compound to block binding of kinases to a desthiobiotin-ATP probe. This showed that CDK7 is the only target bound by 1 μM YKL-5-124 at more than 65% in Jurkat extracts pretreated with YKL-5-124 for 6 h (Figure 2E and Table S1). Furthermore, YKL-5-167 exhibited an activity profile similar to that of YKL-5-124 except that CDK7 was only weakly engaged (Figure 2E and Table S1). To further demonstrate covalent binding of YKL-5-124 for CDK7, we used CRISPR to create an isogenic pair of cell lines expressing either WT or CDK7-C312S (Figure S2D) and then used bioTHZ1 to perform pull-down experiments from cell extracts. We found that the C312S mutation prevented binding of CDK7-cyclin H to beads but had no effect on binding by CDK12/13-CycK (Figure S2E). Collectively, these data show that YKL-5-124 is a CDK7-selective agent, and provide additional support for the C312-dependent covalent mode of binding.

YKL-5-124 Disrupts Cell-Cycle Progression through Inhibition of CDK7 CAK Activity

CDK7 functions as the CAK for several CDKs, including CDK1 and CDK2, activating them by phosphorylating key threonine residues in the T loops of each kinase (Larochelle et al., 2007; Gerber et al., 2008). We therefore examined the effect of exposing HAP1 cells for 24 h to 125 nM to 2 μM YKL-5-124 or 2 μM YKL-5-167 and then blotted cell extracts for the levels of total and phospho-CDK1(T161) and CDK2(T160). Blotting for poly(ADP-ribose) polymerase (PARP) cleavage was used to monitor the extent of apoptosis. We found that YKL-5-124 inhibited CDK1 T-loop phosphorylation, and to a lesser extent CDK2 T-loop phosphorylation (Figure 3A) in a concentration-dependent fashion while YKL-5-167 had no detectable effect. Reduction in the phosphorylation of CDK1 and CDK2 T loops was not observed in HAP1 cells expressing CDK7-C312S, implying that the effect of YKL-5-124 on T-loop phosphorylation is mediated by CDK7 inhibition (Figure 3A). When HAP1 cells were exposed to 2 μM YKL-5-124 for 6, 12, 24, and 48 h, inhibition of T-loop phosphorylation was observed within 12 h (Figure 3B) The absence of PARP cleavage across all conditions showed that apoptosis was not induced by YKL-5-124 exposure. Similar data were obtained in Jurkat cells although the effect on CDK2 T-loop phosphorylation was greater than on the CDK1 T loop (Figure S3A), suggesting variability among cell lines in CAK-substrate modification.

To monitor effects on cell-cycle progression, we exposed HAP1 cells to YKL-5-124 or THZ1 for 72 h and performed cell-cycle analysis by pulsing live cells for 1 h with 5-ethynyl-2′-deoxyuridine (EdU), then fixed cells and stained them with Hoechst and phospho-histone H3 (pHH3) antibodies, followed by fluorescence microscopy. YKL-5-124 caused a dose-dependent increase in G_1- and G_2/M-phase cells and a corresponding loss of S-phase cells (Figure 3C, left panel). Concurrent with S-phase loss, we observed a notable increase in "S dropout," or cells that have intermediate DNA content, but no EdU stain indicating

Figure 3. Cell-Cycle Effects from YKL-5-124 Treatment

(A) Western blot analysis of proteins shown after treatment of HAP1 WT and C312S cells with YKL-5-124 at concentrations indicated for 24 h (pCDK1, phospho T161 CDK1; tCDK1, total CDK1; pCDK2, phospho T160 CDK2; tCDK2, total CDK2), with % Signal indicating percentage compared with DMSO treatment in the relevant cell line.

(B) Western blot analysis of proteins shown after treatment of HAP1 WT and C312S cells with 2 μM YKL-5-124 for time points indicated (pCDK1, phospho T161 CDK1; tCDK1, total CDK1; pCDK2, phosphor T160 CDK2; tCDK2, total CDK2).

(C) Cell-cycle analysis using Hoechst, EdU, and pHH3 staining across concentrations of YKL-5-124 (left), YKL-5-167 (middle), or THZ1 (right) at 72 h in WT cells. Error bars represent SD of technical duplicates.

(D) Cell-cycle analysis using Hoechst, EdU, and pHH3 staining across concentrations of YKL-5-124 (left), YKL-5-167 (middle), or THZ1 (right) at 72 h in CDK7 C312S cells. Error bars represent SD of technical duplicates.

(E) GR value across concentrations in WT (red) and C312S (blue) HAP1 cells with YKL-5-124 (left), YKL-5-167 (middle), or THZ1 (right). Error bars represent SD of technical duplicates.

(F) Percentage of LDR-positive cells across concentrations in WT (red) and C312S (blue) HAP1 cells with YKL-5-124 (left), YKL-5-167 (middle), or THZ1 (right). Error bars represent SD of technical duplicates.

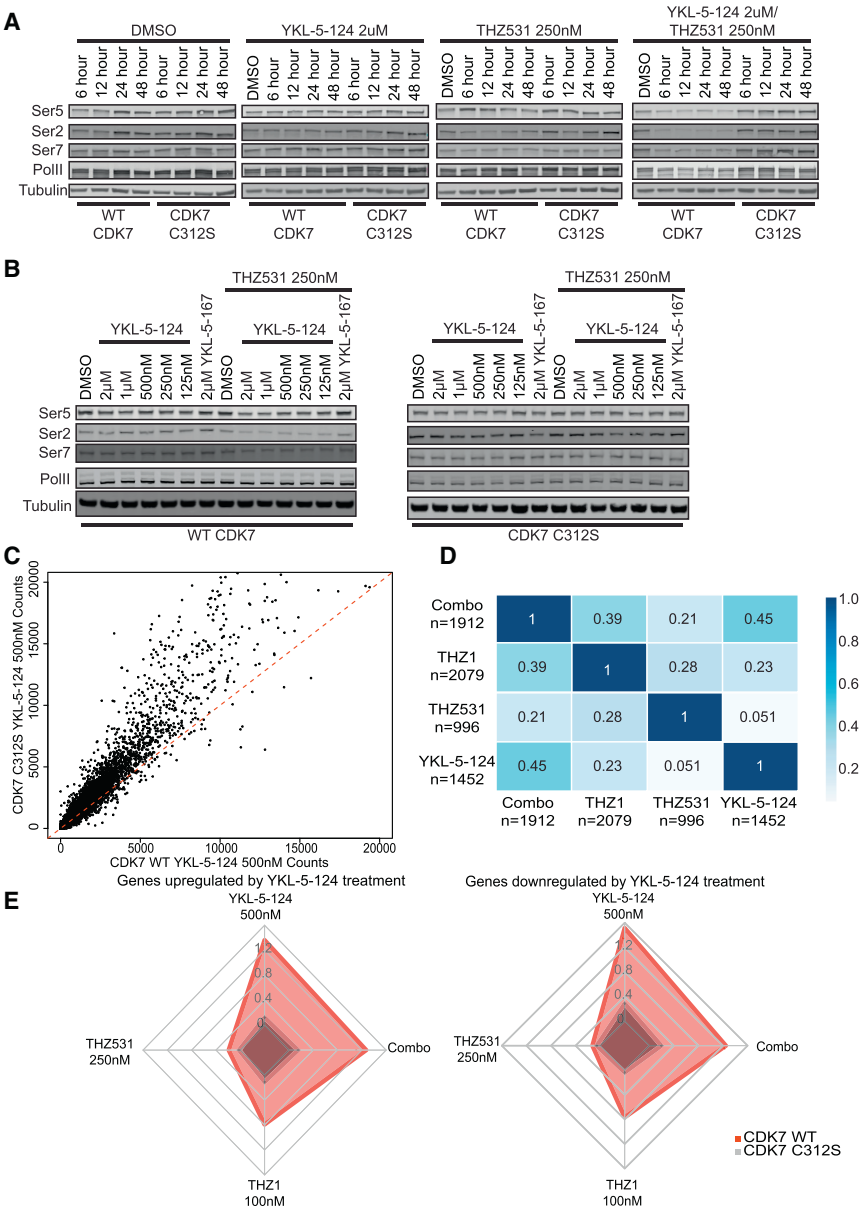


Figure 4. Transcriptional Effects from YKL-5-124 Treatment

(A) Western blot analysis of proteins shown after treatment of HAP1 WT and C312S cells with YKL-5-124 at concentrations indicated for 24 h and co-treatment with 250 nM THZ531 (right).

(B) Western blot analysis of proteins shown after treatment of HAP1 WT and C312S cells with 2 μM YKL-5-124 for time points indicated and co-treatment with 250 nM THZ531 (right).

(C) Scatterplot of FPKM (fragments per kilobase per million reads) of genes comparing YKL-5-124 500 nM treatment in WT and C312S cells.

(D) Jaccard plot comparing the genes downregulated by 500 nM YKL-5-124, 100 nM THZ1, 250 nM THZ531, or YKL-5-124 and THZ531 co-treatment.

(E) Radar plot comparing the effects of different drug treatments on genes upregulated (left) or downregulated (right) by treatment with YKL-5-124 at 500 nM on WT (red) and C312S (gray) cell lines.

inhibition of DNA synthesis. As a control we showed that cell-cycle distribution was not affected by exposure of HAP1 cells expressing CDK7-C312S to YKL-5-124, further supporting the selectivity of YKL-5-124 and that CDK7 is mediating the effects of YKL-5-124 (Figures 3C and 3D, left panel). The reversible compound, YKL-5-167, also showed no effect on cell-cycle distribution.

To quantify drug-induced phenotypes we used growth rate inhibition (GR) metrics, which measure the effect of a drug on viable cell number corrected for the rate of cell division (Hafner et al., 2016, 2018); GR_{50} is a measure of potency and GR_{max} a measure of efficacy at high concentration. The GR_{max} value for THZ1 was negative, indicating cytotoxicity, whereas that of YKL-5-124 was zero, indicating a cytostatic response (Figure 3E) (Hafner et al., 2016). For THZ1, the GR_{50} value was 3.6-fold higher in CDK7-C312S-expressing than in WT cells but the GR_{max} value was unchanged. In contrast, for YKL-5-124 the GR_{50} value was at least 150-fold higher in CDK7-C312S than in WT cells and GR_{max} was too high to be estimated accurately. Staining cells with LIVE/DEAD dye (Hafner et al., 2016) confirmed that THZ1 induced substantially more cell death than YKL-5-124 and that killing was relatively unaffected by CDK7-C312S mutation (whereas cell killing by YKL-5-124 was substantially reduced in CDK7-C312S cells). From these data we conclude that YKL-5-124 primarily induces G_1/S cell cycle and a low level of cell killing in a CDK7-dependent manner, whereas THZ1 is cytotoxic in a largely CDK7-independent manner, as indicated by the potent effect in CDK7-C312S cells (Figure 3F). Cell viability data based on ATP content were obtained in Jurkat cells where treatment with THZ1 was able to inhibit proliferation to zero, indicating total loss of cell viability at the highest concentrations, while treatment with YKL-5-124 could only inhibit proliferation to within 20% of the original cell count as measured by cell titer Glo (Figure S3C). Together, these data show that YKL-5-124 is primarily cytostatic at G_1/S and that arrest in CDK7 dependent; in contrast, THZ1 is cytotoxic in a CDK7-independent manner.

YKL-5-124 Treatment Has No Discernible Effect on RNA Pol II CTD Phosphorylation

CDK7 is reported to be the kinase responsible for phosphorylating S5 in the RNA Pol II CTD heptapeptide repeat; therefore, we asked whether exposure of cells to YKL-5-124 would reduce Pol II CTD phosphorylation. Surprisingly, in both HAP1 (Figure 4A) and Jurkat (Figure S4A) cells we found that YKL-5-124 concentrations 100-fold above GR_{max} (up to 2 μM) had no effect on the phosphorylation of S5 and other Pol II CTD serine residues

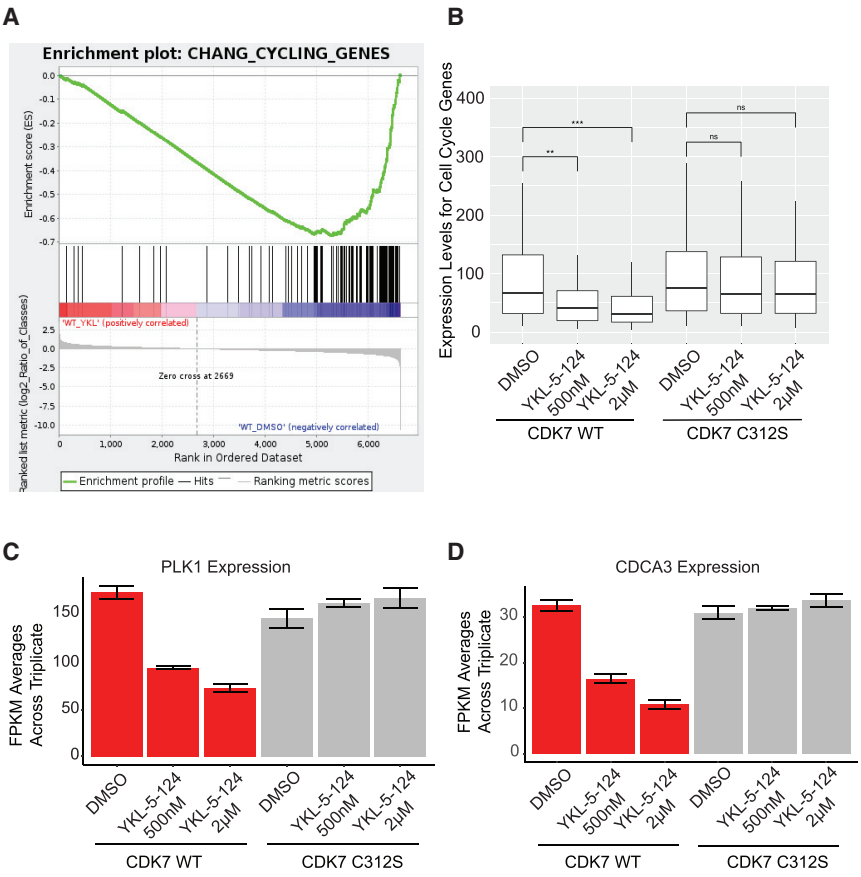


Figure 5. Cell-Cycle Transcriptional Effects from YKL-5-124 Treatment

(A) GSEA of treatment with 500 nM YKL-5-124.

(B) Bar plot quantitating the expression levels on cell-cycle genes with DMSO, 500 nM YKL-5-124, or 2 μM YKL-5-124 treatment in WT and C312S cell lines, with error bars representing SD.

(C) Bar plot quantitating PLK1 expression with DMSO, 500 nM YKL-5-124, or 2 μM YKL-5-124 treatment in WT and C312S cell lines, with error bars representing SD.

(D) Bar plot quantitating CDCA3 expression with DMSO, 500 nM YKL-5-124, or 2 μM YKL-5-124 treatment in WT and C312S cell lines, with error bars representing SD.

cells. However, inhibition of CTD phosphorylation by THZ1 has previously been demonstrated at nanomolar doses (Kwiatkowski et al., 2014). This inconsistency led us to ask whether loss of Pol II CTD phosphorylation requires inhibition of CDK7 and CDK12/13 by THZ1. To test this hypothesis, we combined YKL-5-124 with a previously described inhibitor of CDK12/13, THZ531, and found that the combination was active as a concentration-dependent inhibitor of CTD S2/S5/S7 phosphorylation (Figures 4A, 4B, S4A, and S4B). Inhibition of CDK7 was required for this effect, since it was not observed in CDK7-C312S cells. We conclude that the CTD of Pol II is modified by a combination of CDK7 and Cdk12/13 (Figures 4A and 4B).

Inhibition of CDK7 by YKL-5-124 Elicits a Transcriptional Signature Distinct from that of THZ1

To measure the effects of CDK7 inhibition of YKL-5-124 on gene transcription, we performed RNA sequencing. After 24 h of exposure of WT HAP1 cells to 500 nM YKL-5-124, differential expression of 1,452 genes was observed (a $\log_2$ fold change >1.5) compared with 583 genes in CDK7-C312S cells (Figures 4C and S4D). Under DMSO-only control conditions the transcript profiles of WT and CDK7-C312S were highly similar, demonstrating that the mutation does not appreciably affect baseline transcription (Figure S4C). Next, we compared differentially expressed genes in cells exposed to YKL-5-124, THZ1, THZ531, or a combined of YKL-5-124 and THZ531 by Jaccard analysis, a 0–1 bounded measure of gene set overlap (Levandowsky and Winter, 1971). In the WT HAP1 cells, the effects of 500 nM YKL-5-124 were most similar to those of a combination of 500 nM YKL-5-124 plus 250 nM THZ531 (Jaccard similarity index = 0.45) somewhat less similar to 100 nM THZ1 (index = 0.23) and least similar to 250 nM THZ531 alone (index = 0.05; Figure 4D). Genes that were upregulated or downregulated by YKL-5-124 are shown as radar plots in Figure 4E and were largely similar to those differentially regulated by 100 nM THZ1. The number of genes differentially regulated by YKL-124 was substantially reduced to 583 in CDK7-C312S cells, showing that the effect was mediated by inhibition of CDK7. Moreover, the similarity between YKL-5-124 and the YKL-5-124/THZ531 combination fell in CDK7-C312S relative to WT cells (index = 0.29 versus 0.45) and rose relative to THZ531 alone (index = 0.23 versus 0.05; Figure S4D). These data suggest that CDK7 inhibition alone is not sufficient to affect global basal transcription and that simultaneous inhibition of CDK12/13 is required. Interestingly, our data show that the converse is also true: CDK12/13 inhibition without CDK7 inhibition is insufficient to block global basal transcription.

CDK7 Inhibition Leads to a Strong Cell-Cycle Transcriptional Profile

Gene set enrichment analysis (GSEA) of expression data from WT HAP1 cells treated with YKL-5-124 showed strong leading-edge enrichment for downregulation of E2F expression programs (Figure 5A) (Subramanian et al., 2005). For THZ531, we observed no significant enrichment for E2F gene expression and instead saw downregulation of genes normally repressed by EZH2 (Lu et al., 2010) (Figure S4E). When we focused specifically on a curated set of cell-cycle regulated genes (Chang et al., 2004) (which includes the E2F program) differentially expressed in WT and CDK7-C312S cells at 500 nM or 2 μM, we observed a statistically significant decrease in their expression (Figure 5B). Among genes in the E2F3 cell-cycle program, PLK1 and CDCA3 both exhibited dose-dependent decreases in transcription with YKL-5-124 treatment (Figures 5C and 5D).Taken together, we find a strong inhibition of cell-cycle transcriptional program in cells treated with YKL-5-124 for 24 h.

DISCUSSION

Here we performed target deconvolution on the previously reported CDK7/12/13 inhibitor, THZ1, parsing the CDK7 and CDK12/13 activities to (1) isolate and examine the relative contribution of CDK7 inhibition to the overall THZ1 phenotype, and (2) further explore the consequences of CDK7 pharmacological inhibition. To this end, we leveraged a covalent inhibitor strategy to target CDK7 and other kinases with similarly positioned cysteines to the exclusion of the broader kinome, followed by scaffold diversification to gain further selectivity for CDK7, an approach we previously deployed to selectively inhibit a diverse array of kinase targets (Zhang et al., 2012, 2016; Kwiatkowski et al., 2014). This method culminated in the development of YKL-5-124, an irreversible CDK7 inhibitor derived from the previously reported pyrrolopyrazole-based PAK4 inhibitor, PF-3758309 (Murray et al., 2010). Covalent targeting of CDK7 at C312 is obligatory for YKL-5-124 activity as mutation to the less nucleophilic serine completely rescued both CDK7 activity and cellular proliferation, highlighting both a clear mechanism of action and exceptional selectivity enabling YKL-5-124 use across cellular systems.

Unique among CDKs, CDK7 has been reported to play major roles in both the cell cycle and transcription regulation, through its CAK- and TFIIH-related functions, respectively (Gerber et al., 2008, Ebmeier et al., 2017, Feaver et al., 1994; Adamczewski et al., 1996; Glover-Cutter et al., 2009; Kelso et al., 2014; Rodríguez-Molina et al., 2016). However, many mechanistic details remain elusive and the range of CDK7 functions is incompletely understood. In this study, we employed our newly developed CDK7-selective chemical probe, YKL-5-124, to examine the effects of CDK7 pharmacological inhibition on both the cell cycle and global transcription. Our treatment of HAP1 CML and Jurkat T cell acute lymphoblastic leukemia cells with YKL-5-124 reduced both the levels of activated CDK1/2 and the number of cells in S phase, consistent with previous findings (Larochelle et al., 2007). Interestingly, whereas reduced CDK1/2 activation occurred at concentrations exceeding 1 μM YKL-5-124, strong anti-proliferative effects were observed beginning at 100 nM, suggesting that additional substrates of CDK7 that are important mediators of cell-cycle progression remain to be discovered. RNA-sequencing analysis following treatment with YKL-5-124 revealed a strong cell-cycle signature enriched for E2F3 target genes, indicating that CDK7 may indeed impinge on the cell cycle by various mechanisms. Expression of the inhibitor-refractory C312S mutation rescued all these phenotypes, indicating that these are on-target effects of CDK7 inhibition.

Surprisingly, YKL-5-124 treatment alone did not globally downregulate RNA Pol II CTD phosphorylation or Pol II-mediated gene expression. This stands in direct contrast to our previous inhibitor, THZ1, which elicited robust downregulation of RNA Pol II CTD phosphorylation and gene expression (Kwiatkowski et al., 2014). However, we were able to recapitulate THZ1-mediated effects on gene expression and CTD phosphorylation when YKL-5-124 was combined with our previously characterized CDK12/13 inhibitor, THZ531 (Zhang et al., 2016). This indicates that combined inhibition of CDK7, CDK12, and CDK13 is central to the anti-transcriptional effects of THZ1. Similar results have been observed when the CDK7 "bumped hole" inhibitor, 3MB-PP1, was combined with the pan-CDK inhibitor, flavopiridol, which strongly inhibits CDK9, another Ser2 CTD kinase (Larochelle et al., 2007). This suggests that phosphorylation of the CTD by CDK7 may be dispensable for basal global gene expression and that CDKs 9/12/13 or other kinases such as ERK1/2 (Bonnet et al., 1999) may largely compensate for loss of CDK7. Alternatively, CDK7's role may become essential when Ser2 CTD phosphorylation and corresponding Pol II elongation is first perturbed. This does not exclude the possibility that: (1) CDK7 inhibition leads to potent, but transient loss of Pol II CTD phosphorylation and gene expression, which is quickly reset following engagement of compensatory mechanisms; (2) CDK7 is solely responsible for Ser5 or Ser7 phosphorylation at specific genes whose expression changes following its inhibition; (3) CDK7 kinase activity is not required for basal expression of most genes, but may be involved in gene induction; and (4) CDK7 may regulate gene expression by other means including by transcription factor phosphorylation. These findings highlight the power of using selective kinase inhibitors in combination with complementary genetic tools and systems-level analyses to investigate complex signaling pathways.

SIGNIFICANCE

CDK7 has long been reported to act as both the CDK-activating kinase (CAK) and an RNA polymerase II (Pol II) C-terminal domain (CTD) kinase, thereby imparting top-down regulation of the cell cycle and gene transcription, respectively. However, results and interpretations thereof have been complicated by the use of non-selective CDK-targeting small molecules. Here we describe the development of YKL-5-124, a potent and selective covalent inhibitor of CDK7. Using YKL-5-124, we confirmed a strong cell-cycle phenotype in cells following CDK7 inhibition with reduction in cycling cells, CAK activity, and cell-cycle-regulated gene expression. However, cells exhibited little to no change in global Pol II CTD phosphorylation or general loss of transcription factor genes, which are hallmarks of THZ1 treatment. Furthermore, while YKL-5-124 treatment of HAP1 and Jurkat cells led to a strong cell-cycle arrest, it did not induce cell death like THZ1. The ability to recapitulate THZ1 phenotypes using potent by separate inhibitors of CDK7 and CDK12/13 signal the importance of this polypharmacology to the transcriptional activity of THZ1 and derivative compounds. Therefore, YKL-5-124 represents a compound that enables rapid inactivation of CDK7 to study the consequences of selective CDK7 inhibition on both the cell cycle and gene transcription. As YKL-5-124 does not indiscriminately induce cell death, it also presents a unique therapeutic opportunity to find cancer cell states that are uniquely dependent on CDK7 inhibition for survival. As YKL-5-124 demonstrated suppression of the cell cycle at both the gene expression and post-translational levels, therapeutic strategies aimed at targeting cancers that are driven by direct misregulation of cell-cycle factors may show promise.

STAR★METHODS

Detailed methods are provided in the online version of this paper and include the following:

- KEY RESOURCES TABLE
- CONTACT FOR REAGENT AND RESOURCE SHARING
- EXPERIMENTAL MODEL AND SUBJECT DETAILS
  - Cell Lines
  - Constructs
- METHOD DETAILS
  - Genome Editing
  - Immunoblotting
  - Commercial *In Vitro* Kinase Assays
  - *In Vitro* Kinase Assays
  - Radioactive *In Vitro* Kinase Assays from Cellular CDK7
  - Pulldown Assays
  - Covalent Inhibition Kinetics Using Mobility Shift Assay
  - Mass Spectrometry Analysis
  - Kinativ Sample Preparation and Analysis
  - Proliferation Assays
  - Cell Cycle and Growth Rate Analysis
  - RNA Sequencing
  - RNA-seq Analysis
  - Chemical Synthesis
- QUANTIFICATION AND STATISTICAL ANALYSIS
- DATA AND SOFTWARE AVAILABILITY

SUPPLEMENTAL INFORMATION

Supplemental Information can be found online at https://doi.org/10.1016/j.chembiol.2019.02.012.

ACKNOWLEDGMENTS

We would like to thank Richard A. Young and Nancy M. Hannett for their advice and support on the manuscript. C.Y.L., W.D.P., and S.Z.E. are supported by the Cancer Prevention Research Institute of Texas (RR150093). C.Y.L. is also supported by NIH and NCI (1R01CA215452-01) and is a Pew-Stewart Scholar for Cancer Research (Alexander and Margaret Stewart Trust). P.K.S. and N.S.G. are supported by NIH U54-HL127365, and C.M.O., Y.L., T.Z., N.K., and N.S.G. are supported by NIH R01 CA179483-03, the Koch Institute-DF/HCC Bridge Project, and the Hale Center for Pancreatic Cancer. M.G. was supported by the Deutsche Forschungsgemeinschaft (GE-976/9-2).

AUTHOR CONTRIBUTIONS

Conceptualization, C.M.O., Y.L., N.K., and N.S.G.; Investigation, C.M.O., Y.L., A.L., L.L., C.E.M., S.B.F., and R.D.; Formal Analysis, W.D.P. and S.Z.E.; Writing – Original Draft, C.M.O., N.K, and N.S.G.; Writing – Review & Editing, C.M.O., N.K., N.S.G., A.L., L.L., C.E.M., S.B.F., J.A.M., P.K.S., K.D.W., and C.Y.L.,; Funding Acquisition, N.K., T.Z., and N.S.G.; Supervision, J.A.M., P.K.S., K.D.W., C.Y.L., and M.G.

DECLARATION OF INTERESTS

C.Y.L. an equity holder and inventor of IP licensed to Syros Pharmaceuticals and N.S.G. is a founder, science advisor, and equity holder in C4, Syros, Petra, Soltego, B2S, and Gatekeeper Pharmaceuticals. The Gray lab receives or has received research funding from Novartis, Takeda, Astellas, Taiho, Janssen, Kinogen, Voronoi, Her2llc, Deerfield, and Sanofi. N.S.G., T.Z., and N.K. are inventors on a patent application covering THZ1, which is licensed to a company co-founded by N.S.G. The other authors declare no competing interests.

Received: September 25, 2018
Revised: December 19, 2018
Accepted: February 18, 2019
Published: March 21, 2019

REFERENCES

Adamczewski, J.P., Rossignol, M., Tassan, J.P., Nigg, E.A., Moncollin, V., and Egly, J.M. (1996). Mat1, cdk7 and cyclin H form a kinase complex which is UV light-sensitive upon association with TFIIH. EMBO J. *15*, 1877–1884.

Akhtar, M.S., Heidemann, M., Tietjen, J.R., Zhang, D.W., Chapman, R.D., Eick, D., and Ansari, A.Z. (2009). TFIIH kinase places bivalent marks on the carboxy-terminal domain of RNA polymerase II. Mol. Cell *34*, 387–393.

Alexander, W.M., Ficarro, S.B., Adelmant, G., and Marto, J.A. (2017). Multiplierz v2.0: a Python-based ecosystem for shared access and analysis of native mass spectrometry data. Proteomics *17*, https://doi.org/10.1002/pmic.201700091.

Ali, S., Heathcote, D.A., Kroll, S.H.B., Jogalekar, A.S., Scheiper, B., Patel, H., Brackow, J., Siwicka, A., Fuchter, M.J., Periyasamy, M., et al. (2009). The development of a selective cyclin-dependent kinase inhibitor that shows anti-tumor activity. Cancer Res. *69*, 6208–6215.

Andersson, B., Collins, V., Kurzrock, R., Larkin, W.D., Childs, C., Ost, A., Cork, A., Trujillo, J.M., et al. (1995). KBM-7, a human myeloid leukemia cell line with double Philadelphia chromosomes lacking normal c-ABL and BCR transcripts. Leukemia *9*, 2100–2108.

Askenazi, M., Parikh, J.R., and Marto, J.A. (2009). mzAPI: a new strategy for efficiently sharing mass spectrometry data. Nat. Methods *6*, 240–241.

Bieniossek, C., Imasaki, T., Takagi, Y., and Berger, I. (2012). MultiBac: expanding the research toolbox for multiprotein complexes. Trends Biochem. Sci. *37*, 49–57.

Bisteau, X., Paternot, S., Colleoni, B., Ecker, K., Coulonval, K., de Groote, P., Declercq, W., Hengst, L., and Roger, P.P. (2013). CDK4 T172 phosphorylation is central in a CDK7-dependent bidirectional CDK4/CDK2 interplay mediated by p21 phosphorylation at the restriction point. PLoS Genet. *9*, e1003546.

Blackwell, L.J., Birkos, S., Hallam, R., van de Carr, G., Arroway, J., Suto, C.M., and Janzen, W.P. (2009). High-throughput screening of the cyclic AMP-dependent protein kinase (PKA) using the caliper microfluidic platform. Methods Mol. Biol. *565*, 225–237.

Bonnet, F., Vigneron, M., Bensaude, O., and Dubois, M.F. (1999). Transcription-independent phosphorylation of the RNA polymerase II C-terminal domain (CTD) involves ERK kinases (MEK1/2). Nucleic Acids Res. *27*, 4399–4404.

Bösken, C.A., Farnung, L., Hintermair, C., Merzel Schachter, M., Vogel-Bachmayr, K., Blazek, D., Anand, K., Fisher, R.P., Eick, D., and Geyer, M. (2014). The structure and substrate specificity of human Cdk12/Cyclin K. Nat. Commun. *5*, 3505.

Buratowski, S. (2009). Progression through the RNA polymerase II CTD cycle. Mol. Cell *36*, 541–546.

Canevari, G., Re Depaolini, S., Cucchi, U., Bertrand, J.A., Casale, E., Perrera, C., Forte, B., Carpinelli, P., and Felder, E.R. (2013). Structural insight into maternal embryonic leucine zipper kinase (MELK) conformation and inhibition toward structure-based drug design. Biochemistry *52*, 6380–6387.

Chang, H.Y., Sneddon, J.B., Alizadeh, A.A., Sood, R., West, R.B., Montgomery, K., Chi, J.T., van de Rijn, M., Botstein, D., and Brown, P.O. (2004). Gene expression signature of fibroblast serum response predicts human cancer progression: similarities between tumors and wounds. PLoS Biol. *2*, E7.

Chen, F., Gao, X., and Shilatifard, A. (2015). Stably paused genes revealed through inhibition of transcription initiation by the TFIIH inhibitor triptolide. Genes Dev. *29*, 39–47.

Cong, L., Ran, F.A., Cox, D., Lin, S., Barretto, R., Habib, N., Hsu, P.D., Wu, X., Jiang, W., Marraffini, L.A., and Zhang, F. (2013). Multiplex genome engineering using CRISPR/Cas systems. Science *339*, 819–823.

Devault, A., Martinez, A.M., Fesquet, D., Labbé, J.C., Morin, N., Tassan, J.P., Nigg, E.A., Cavadore, J.C., and Dorée, M. (1995). MAT1 ('menage à trois') a

new RING finger protein subunit stabilizing cyclin H-cdk7 complexes in starfish and Xenopus CAK. EMBO J. *14*, 5027–5036.

Ebmeier, C.C., Erickson, B., Allen, B.L., Allen, M.A., Kim, H., Fong, N., Jacobsen, J.R., Liang, K., Shilatifard, A., Dowell, R.D., et al. (2017). Human TFIIH kinase CDK7 regulates transcription-associated chromatin modifications. Cell Rep. *20*, 1173–1186.

Egloff, S., and Murphy, S. (2008). Cracking the RNA polymerase II CTD code. Trends Genet. *24*, 280–288.

Essletzbichler, P., Konopka, T., Santoro, F., Chen, D., Gapp, B.V., Kralovics, R., Brummelkamp, T.R., Nijman, S.M.B., and Bürckstümmer, T. (2014). Megabase-scale deletion using CRISPR/Cas9 to generate a fully haploid human cell line. Genome Res. *24*, 2059–2065.

Feaver, W.J., Svejstrup, J.Q., Henry, N.L., and Kornberg, R.D. (1994). Relationship of CDK-activating kinase and RNA polymerase II CTD kinase TFIIH/TFIIK. Cell *79*, 1103–1109.

Ficarro, S.B., Alexander, W.M., and Marto, J.A. (2017). mzStudio: a dynamic digital canvas for user-driven interrogation of mass spectrometry data. Proteomes *5*, https://doi.org/10.3390/proteomes5030020.

Ficarro, S.B., Zhang, Y., Lu, Y., Moghimi, A.R., Askenazi, M., Hyatt, E., Smith, E.D., Boyer, L., Schlaeger, T.M., Luckey, C.J., and Marto, J.A. (2009). Improved electrospray ionization efficiency compensates for diminished chromatographic resolution and enables proteomics analysis of tyrosine signaling in embryonic stem cells. Anal. Chem. *81*, 3440–3447.

Fisher, R.P., and Morgan, D.O. (1994). A novel cyclin associates with M015/CDK7 to form the CDK-activating kinase. Cell *78*, 713–724.

Furuno, N., den Elzen, N., and Pines, J. (1999). Human cyclin A is required for mitosis until mid prophase. J. Cell Biol. *147*, 295–306.

Gao, Y., Zhang, T., Terai, H., Ficarro, S.B., Kwiatkowski, N., Hao, M.-F., Sharma, B., Christensen, C.L., Chipumuro, E., Wong, K.-K., and Marto, J.A. (2018). Overcoming resistance to the THZ series of covalent transcriptional CDK inhibitors. Cell Chem. Biol. *25*, 135–142.e5.

Gerber, H.B., Pikman, Y., and Fisher, R.P. (2008). The CDK-activating kinase (CAK) Csk1 is required for normal levels of homologous recombination and resistance to DNA damage in fission yeast. PLoS One *3*, e1492.

Glover-Cutter, K., Larochelle, S., Erickson, B., Zhang, C., Shokat, K., Fisher, R.P., and Bentley, D.L. (2009). TFIIH-associated Cdk7 kinase functions in phosphorylation of C-terminal domain Ser7 residues, promoter-proximal pausing, and termination by RNA polymerase II. Mol. Cell. Biol. *29*, 5455–5464.

Greifenberg, A.K., Honig, D., Pilarova, K., Duster, R., Bartholomeeusen, K., Bosken, C.A., Anand, K., Blazek, D., and Geyer, M. (2016). Structural and functional analysis of the Cdk13/Cyclin K complex. Cell Rep. *14*, 320–331.

Guo, C., Mcalpine, I., Zhang, J., Knighton, D.D., Kephart, S., Johnson, M.C., Li, H., Bouzida, D., Yang, A., Dong, L., et al. (2012). Discovery of pyrroloaminopyrazoles as novel PAK inhibitors. J. Med. Chem. *55*, 4728–4739.

Hafner, M., Mills, C.E., Gerosa, L., Chung, M., Niepel, M. & Sorger, P.K., DeepDyeDrop: an image-based approach to quantify the phenotypic response of cancer cells to therapeutics (conference presentation). SPIE BiOS, 2018. SPIE.

Hafner, M., Niepel, M., Chung, M., and Sorger, P.K. (2016). Growth rate inhibition metrics correct for confounders in measuring sensitivity to cancer drugs. Nat. Methods *13*, 521–527.

Kalan, S., Amat, R., Schachter, M.M., Kwiatkowski, N., Abraham, B.J., Liang, Y., Zhang, T., Olson, C.M., Larochelle, S., Young, R.A., et al. (2017). Activation of the p53 transcriptional program sensitizes cancer cells to Cdk7 inhibitors. Cell Rep. *21*, 467–481.

Kanin, E.I., Kipp, R.T., Kung, C., Slattery, M., Viale, A., Hahn, S., Shokat, K.M., and Ansari, A.Z. (2007). Chemical inhibition of the TFIIH-associated kinase Cdk7/Kin28 does not impair global mRNA synthesis. Proc. Natl. Acad. Sci. U S A *104*, 5812–5817.

Kelso, T.W.R., Baumgart, K., Eickhoff, J., Albert, T., Antrecht, C., Lemcke, S., Klebl, B., and Meisterernst, M. (2014). Cyclin-dependent kinase 7 controls mRNA synthesis by affecting stability of preinitiation complexes, leading to altered gene expression, cell cycle progression, and survival of tumor cells. Mol. Cell. Biol. *34*, 3675–3688.

Klaeger, S., Heinzlmeir, S., Wilhelm, M., Polzer, H., Vick, B., Koenig, P.A., Reinecke, M., Ruprecht, B., Petzoldt, S., Meng, C., et al. (2017). The target landscape of clinical kinase drugs. Science *358*, https://doi.org/10.1126/science.aan4368.

Kuzmic, P. (1996). Program DYNAFIT for the analysis of enzyme kinetic data: application to HIV proteinase. Anal. Biochem. *237*, 260–273.

Kwiatkowski, N., Zhang, T., Rahl, P.B., Abraham, B.J., Reddy, J., Ficarro, S.B., Dastur, A., Amzallag, A., Ramaswamy, S., Tesar, B., et al. (2014). Targeting transcription regulation in cancer with a covalent CDK7 inhibitor. Nature *511*, 616–620.

Larochelle, S., Amat, R., Glover-Cutter, K., Sansó, M., Zhang, C., Allen, J.J., Shokat, K.M., Bentley, D.L., and Fisher, R.P. (2012). Cyclin-dependent kinase control of the initiation-to-elongation switch of RNA polymerase II. Nat. Struct. Mol. Biol. *19*, 1108–1115.

Larochelle, S., Merrick, K.A., Terret, M.-E., Wohlbold, L., Barboza, N.M., Zhang, C., Shokat, K.M., Jallepalli, P.V., and Fisher, R.P. (2007). Requirements for Cdk7 in the assembly of Cdk1/Cyclin B and activation of Cdk2 revealed by chemical genetics in human cells. Mol. Cell *25*, 839–850.

Levandowsky, M., and Winter, D. (1971). Distance between sets. Nature *234*, 34.

Lu, C., Han, H.D., Mangala, L.S., Ali-Fehmi, R., Newton, C.S., Ozbun, L., Armaiz-Pena, G.N., Hu, W., Stone, R.L., Munkarah, A., et al. (2010). Regulation of tumor angiogenesis by EZH2. Cancer Cell *18*, 185–197.

Marshall, N.F., Peng, J., Xie, Z., and Price, D.H. (1996). Control of RNA polymerase ii elongation potential by a novel carboxyl-terminal domain kinase. J. Biol. Chem. *271*, 27176–27183.

Murray, B.W., Guo, C., Piraino, J., Westwick, J.K., Zhang, C., Lamerdin, J., Dagostino, E., Knighton, D., Loi, C.-M., Zager, M., et al. (2010). Small-molecule p21-activated kinase inhibitor PF-3758309 is a potent inhibitor of oncogenic signaling and tumor growth. Proc. Natl. Acad. Sci. U S A *107*, 9446–9451.

Nilson, Kyle A., Guo, J., Turek, M.E., Brogie, J.E., Delaney, E., Luse, D.S., and Price, D.H. (2015). THZ1 reveals roles for Cdk7 in Co-transcriptional capping and pausing. Mol. Cell *59*, 576–587.

Patricelli, M.P., Nomanbhoy, T.K., Wu, J., Brown, H., Zhou, D., Zhang, J., Jagannathan, S., Aban, A., Okerberg, E., Herring, C., et al. (2011). In situ kinase profiling reveals functionally relevant properties of native kinases. Chem. Biol. *18*, 699–710.

Perkins, D.N., Pappin, D.J., Creasy, D.M., and Cottrell, J.S. (1999). Probability-based protein identification by searching sequence databases using mass spectrometry data. Electrophoresis *20*, 3551–3567.

Poss, Z.C., Ebmeier, C.C., Odell, A.T., Tangpeerachaikul, A., Lee, T., Pelish, H.E., Shair, M.D., Dowell, R.D., Old, W.M., and Taatjes, D.J. (2016). Identification of mediator kinase substrates in human cells using cortistatin A and quantitative phosphoproteomics. Cell Rep. *15*, 436–450.

R Core Team (2014). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. http://www.R-project.org/.

Rodríguez-Molina, J.B., Tseng, S.C., Simonett, S.P., Taunton, J., and Ansari, A.Z. (2016). Engineered covalent inactivation of TFIIH-kinase reveals an elongation checkpoint and results in widespread mRNA stabilization. Mol. Cell *63*, 433–444.

Rudolph, J., Crawford, J.J., Hoeflich, K.P., and Wang, W. (2015). Inhibitors of p21-activated kinases (PAKs). J. Med. Chem. *58*, 111–129.

Satyanarayana, A., and Kaldis, P. (2009). Mammalian cell-cycle regulation: several Cdks, numerous cyclins and diverse compensatory mechanisms. Oncogene *28*, 2925–2939.

Schachter, M.M., Merrick, K.A., Larochelle, S., Hirschi, A., Zhang, C., Shokat, K.M., Rubin, S.M., and Fisher, R.P. (2013). A Cdk7-Cdk4 T-loop phosphorylation cascade promotes G1 progression. Mol. Cell *50*, 250–260.

Serizawa, H., Conaway, J.W., and Conaway, R.C. (1993). Phosphorylation of C-terminal domain of RNA polymerase II is not required in basal transcription. Nature *363*, 371–374.

Serizawa, H., Makela, T.P., Conaway, J.W., Conaway, R.C., Weinberg, R.A., and Young, R.A. (1995). Association of Cdk-activating kinase subunits with transcription factor TFIIH. Nature *374*, 280–282.

Sherr, C.J., and Roberts, J.M. (1999). CDK inhibitors: positive and negative regulators of G1-phase progression. Genes Dev. *13*, 1501–1512.

Shiekhattar, R., Mermelstein, F., Fisher, R.P., Drapkin, R., Dynlacht, B., Wessling, H.C., Morgan, D.O., and Reinberg, D. (1995). Cdk-activating kinase complex is a component of human transcription factor TFIIH. Nature *374*, 283–287.

Subramanian, A., Tamayo, P., Mootha, V.K., Mukherjee, S., Ebert, B.L., Gillette, M.A., Paulovich, A., Pomeroy, S.L., Golub, T.R., Lander, E.S., and Mesirov, J.P. (2005). Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. Proc. Natl. Acad. Sci. U S A *102*, 15545–15550.

Tan, L., Gurbani, D., Weisberg, E.L., Hunter, J.C., Li, L., Jones, D.S., Ficarro, S.B., Mowafy, S., Tam, C.-P., Rao, S., et al. (2017). Structure-guided development of covalent TAK1 inhibitors. Bioorg. Med. Chem. *25*, 838–846.

Wang, B.-Y., Liu, Q.-Y., Cao, J., Chen, J.-W., and Liu, Z.-S. (2016). Selective CDK7 inhibition with BS-181 suppresses cell proliferation and induces cell cycle arrest and apoptosis in gastric cancer. Drug Des. Devel. Ther. *10*, 1181–1189.

Zhang, T., Inesta-Vaquera, F., Niepel, M., Zhang, J., Ficarro, Scott B., Machleidt, T., Xie, T., Marto, Jarrod A., Kim, N., Sim, T., et al. (2012). Discovery of potent and selective covalent inhibitors of JNK. Chem. Biol. *19*, 140–154.

Zhang, T., Kwiatkowski, N., Olson, C.M., Dixon-Clarke, S.E., Abraham, B.J., Greifenberg, A.K., Ficarro, S.B., Elkins, J.M., Liang, Y., Hannett, N.M., et al. (2016). Covalent targeting of remote cysteine residues to develop CDK12 and CDK13 inhibitors. Nat. Chem. Biol. *12*, 876–884.

STAR★METHODS

KEY RESOURCES TABLE

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Antibodies		
Anti-CDK7	Cell Signaling Technology	Cat#2916; RRID: AB_2077142
Anti-Cyclin H	Bethyl Labs	Cat#A301-674A; RRID: AB_1210922
Anti-Cyclin K	Bethyl Labs	Cat#A301-939A; RRID: AB_1547934
Anti-CDK2	Bethyl Labs	Cat#A301-812; RRID: AB_1233061
Anti-Phospho-CDK2 T160	Cell Signaling Technology	Cat#2561; RRID: AB_2078685
Anti-CDK1	Bethyl Labs	Cat#A303-663; RRID: AB_11205291
Anti-Phospho-cdc2 T161	Cell Signaling Technology	Cat#9114; RRID: AB_2074652
Anti-Tubulin	Cell Signaling Technology	Cat#3873; RRID: AB_1904178
Anti-PARP	Cell Signaling Technology	Cat#9542; RRID: AB_2160739
Anti-Phospho-CTD Ser2	Millipore	Cat#04-1571; RRID: AB_11212363
Anti-Phospho-CTD Ser5	Millipore	Cat#04-1572; RRID: AB_10615822
Anti-Phospho-CTD Ser7	Millipore	Cat#04-1570-I; RRID: AB_10618152
Anti-Total RNA Polymerase II	Santa Cruz Biotechnology	Cat#sc-899; RRID: AB_632359
Anti-phospho-histone H3 Alexa 488 S10	Cell Signaling Technology	Cat#3377; RRID: AB_1549592
Goat IRDye 680RD anti-Rat IgG (H+L)	LI-COR Biosciences	Cat#926-68076; RRID: AB_10956590
Goat IRDye 680RD anti-Mouse IgG	LI-COR Biosciences	Cat#926-68070; RRID: AB_10956588
Goat IRDye 800CW anti-Rabbit IgG	LI-COR Biosciences	Cat#926-32211; RRID: AB_621843
Chemicals, Peptides, and Recombinant Proteins		
Biotinylated-THZ1	Kwiatkowski et al., 2014	N/A
YKL-5-124	This paper	
YKL-5-167	This paper	
THZ1	Kwiatkowski et al., 2014	N/A
THZ531	Zhang et al., 2016	N/A
CellTiter-Glo Luminescent Cell Viability Assay	Promega	Cat#G7571
M2-Agarose Resin	Sigma	Cat#A2220
cOmplete, Mini Protease Inhibitor Cocktail	Roche	Cat#11836153001
PhosSTOP Phosphotase Inhibitor Tablets	Roche	Cat#04906837001
Mycoalert	Lonza	Cat#LT07-318
CDK7/Cyclin H/ MAT1	Thermo Fisher Scientific	Cat#PR6749B
[γ-32P] ATP	Perkin-Elmer	Cat#NEG002A
CDK12/Cyclin K	Bösken et al., 2014	N/A
CDK13/Cyclin K	Greifenberg et al., 2016	N/A
CDK7/CycH/Mat1	Carna Biosciences	Cat#04-108
CTD3 peptide	Carna Biosciences	Cat#04-108MS
EdU	Lumiprobe	Cat#10540
LIVE/DEAD Far Red Dead Cell Stain	Thermo Fisher Scientific	Cat#L10120
Cy3-azide	Lumiprobe	Cat#11030
RNAPII (RPB1)	abcam	Cat#ab81888
Odyssey Blocking Buffer	LI-COR Biosciences	Cat#927-50100
Hoechst 33342	Sigma Aldrich	Cat#B2261
X-tremeGENE 9	Sigma Aldrich	Cat# 6365779001

(*Continued on next page*)

Continued

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Critical Commercial Assays		
Selectscreen- Adapta Eu- CDK7/CycH/MNAT1	Invitrogen	N/A
Selectscreen- Adapta Eu- CDK9/CycT1	Invitrogen	N/A
Selectscreen- Z'LYTE- CDK2/CycA	Invitrogen	N/A
Deposited Data		
RNA-Seq Data	This paper	Accession Number GEO: GSE124607
Experimental Models: Cell Lines		
Human: Jurkat cells (male)	Laboratory of Thomas Looke	N/A
Human: HAP1 cells (male)	Horizon Discovery	C859
Human: Hela cells (female)	Kwiatkowski et al., 2014	N/A
Oligonucleotides		
GCCAAGACCAAACTGTCCAG- sgRNA against CDK7	This paper	N/A
CCAAGACCAAACTGTCCAGTGG- WT specific amplification	This paper	N/A
ACCAAACTCACCGGTAGAAAC- mutation specific amplification	This paper	N/A
TGGGTCAGATCTTCCAAATAACT- genomic CDK7 primer	This paper	N/A
Recombinant DNA		
pX330	N/A	Addgene: 42230; Cong et al., 2013
pUC57-AMP	Laboratory of Richard Young	N/A
Software and Algorithms		
DynaFit software	Kuzmic, 1996	http://www.biokin.com/dynafit/
GraphPad Prism	GraphPad Software, Inc	http://www.graphpad.com/
mzStudio software	Ficarro et al., 2017	http://github.com/BlaisProteomics/mzStudio
Mascot v 2.2	Perkins et al., 1999	http://www.matrixscience.com
Multiplierz scripts	Askenazi et al., 2009	https://github.com/MaxAlex/multiplierz/wiki
R	R Core Team, 2014	https://www.r-project.org/
Columbus image storage and analysis	PerkinElmer Informatics.	http://www.perkinelmer.com/product/image-data-storage-and-analysis-system-columbus
R scripts for RNAseq Analysis	This Paper	https://github.com/linlabcode/olson_YKL5124

CONTACT FOR REAGENT AND RESOURCE SHARING

Further information and requests for resources and reagents should be directed to and will be fulfilled by the Lead Contact, Nathanael Gray (nathanael_gray@dfci.harvard.edu).

EXPERIMENTAL MODEL AND SUBJECT DETAILS

Cell Lines

HAP1 (male) cells (Horizon Discovery) were cultured in IMDM media (Cat # 12440061, Thermo Fisher) supplemented with 10% FBS and 1% penicillin/streptomycin (Cat # 10378016, Thermo Fisher) and Jurkat (male) cells (T. Looke Lab, DFCI) were cultured in RPMI media (Cat# 11875093, Thermo Fisher) and incubated at 37°C with 5% CO_2. All cell lines were routinely tested for mycoplasma infection (Lonza, LT07-318). Cell lines were not authenticated.

Constructs

For CRISPR studies, guide RNAs (sgRNA) targeting CDK7 were cloned into pX330 (Addgene: 42230) (Cong et al., 2013). pUC57-AMP was used as storage vector for CDK7 genome reference sequence.

METHOD DETAILS

Genome Editing

The CRISPR/Cas9 system was used to mutate the endogenous CDK7 WT locus to encode for CDK7 C312S, a THZ1 and YKL-5-124 -refractory mutant of CDK7. Target-specific oligonucleotides were cloned into pX330 (Addgene: 42230), which carries a codon-optimized version of Cas9 and was further modified to express GFP for identifying transfectants. Cells were co-transfected (X-tremeGENE 9 (Roche)) with (1) pX330 expressing Cas9 and CDK7 targeting sgRNA and (2) a pUC57-AMP construct bearing 1500 bp of modified CDK7 reference genome that is centered around the CRISPR targeting site in CDK7. Two days after transfection, cells were sorted using GFP as a marker of transfected cells and cells were re-plated for five days. Cells were then re-plated at low density to facilitate the isolation of individual clones. Individual clones were isolated, expanded, and PCR genotyped using mutant specific PCR primers. Following initial PCR screening, individual clones were Sanger sequenced to confirm the presence of the desired mutation. Western blot confirmed the presence of intact CDK7 kinase. Subsequent experiments were conducted using a CDK7 C312S clone and a WT control clone that was carried through the entirety of the CRISPR protocol, but that was verified by Sanger sequencing to be WT CDK7.

The genomic sequence complementary to the CDK7 -directed guide RNA that was cloned into pX330 and used in the genome editing experiments is: GCCAAGACCAAACTGTCCAG.

The modified genomic sequence that was cloned into pUC57-AMP by Genewiz and used as the repair template for genome editing is:

GCCATGAGGAGTCGACTTAATCTTGGCTTCCATCCCCAAGGTACGTCATGTATATGCAAATATTCCACAGTCCAAAATCTGAAGC
ACTTCTGGTCCCAAGCATTTTAGATAAGATATATTCAGCCTGTATAGGCTATTTAATTATGGAAAAATTAGAAGGGAGTCTTTTTTTCT
TTTCTTTTTTTGATACGGAGCCTTGCTCTGTCGCCCAGGCTGGAGTGCAGTGGTGCGATCTCGGCTCACTGCAACCTCTGCCTCT
GGGTTCAAGCGATTCTCCTACCTCAGCCTCCCGAGTAGTGGGGACTACAGACATGCACCACCACGCCCAGCTTATTTTTGTATTTT
TAGTAGAGATGGAGTTTCGCCACGTTGGCCAAGCTGATCTCGAACTTCTGACCTCATGTGATCCACCCGCCTCAGCCTCCCAAGG
TGCTGGGATTACAGGCATGAGCTACTGTGCCCGGCCAGTCTTTTAATGGAATTATCAGCCAATCTTAGGATCCTAGGTATTTATTTG
TAGTAATTTCCATTTGCAAATACTTGACATTAATTTTGTTTTGTTGGGAATGAGGGCATTCACATAGTATTTTTAATGCTCCCTATAAT
CTTGAAGGTAAGATTTTCCTTAACACATTTCAGATACTCACCTACCTTACTTTTGGTATCTTTTCTTTTAAAAGGCACTGAAAATGAAG
TATTTCAGTAATCGGCCAGGGCCAACACCTGGATGTCAGCTGCCAAGACCAAACTCACCGGTAGAAACCTTAAAGGAGCAATCAA
ATCCAGCTTTAGCAATAAAAAGGAAAAGAACAGAGGCCTTAGAACAAGGTAAGATTCCCACTTTTAAAAGAAATTAAATGAATTTAG
AAAACTCCAAATAGCTCGTGTATGGCTAGCGACTGAACAACAGCAAAGAAATGTAGCTTGCTAGCTATTTAATTTTTATAAATTTAAT
TGTATAAAGTAGAGAGACTGTGCCGTTTTAGGTATGTTTACTTTCATTGTGAATACTTCGGCAGCTTTTGTGTATACCTCAGAATTAT
GGGAACTATGTATGTGTACCTTGTTCCCTTCATAAATTCCTTCTGTAAACTACATGCTAAATGTTTAACAAGTGCTACTGGAAAAA
ATGTGTCAGTGAAATAATGTAGTTGGAATGCAGTGACTGGTTAAAATTGCTATGAGAATGTGAACTTTGTTAAATGTGAACAACTTT
TTTTTTTCTTGTTCTTTTTGCTTCCTAGGAGGATTGCCCAAGAAACTAATTTTTTAAAGAGAACACTGGACAACATTTTACTACTGAGG
GAAATAGCCAAAAAGGCAAATAATGGAAAAATAGTAAACATTAAGTAAATGCTGTAGAAGTGAGTTTGTAAATATTCTACACATGTA
AAATATGTAAAACTATGGGTTATTTTTATTAAATGTATTTTAAAATAAAAATTTAATTCTGGTTTTTCTGATTAGAGTGCAAAAGTGA
GAAAGAATTCTCGAGTACCG

1. Green highlighting indicates the introduced desired TCA mutation, which codes for serine (C312S), replacing TGT which codes for cysteine (C312, WT).
2. Yellow highlighting indicates wobble mutations introduced to remove Cas9 –targeting sites, to prevent cutting of repair template.
3. Pink highlighting indicates Wobble mutations introduced for PCR-based screening to permit WT vs. mutated allele discrimination.
4. Red highlighting indicates Sal I and EcoRI sites used for pUC57 cloning

Immunoblotting

HAP1 or Jurkat cells were lysed in RIPA buffer (50mM Tris-HCl, 150 mM NaCl, 1% NP-40, 0.5% sodium deoxycholate, and 0.1% SDS, pH 7.4 ± 0.2) with Protease Inhibitor (Roche), Phosstop Phosphatase Inhibitor (Roche), and 2.5 U/ml Universal Nuclease for Cell Lysis (Pierce) by incubating on ice 30 min. The lysates were clarified by spinning at 21,000 *g* for 30 min at 4°C and the concentration of the lysate was determined using BCA protocol (Pierce). Primary antibodies in this study include: CDK7 (Cell Signaling Technology, 2916, 1:1,000), Cyclin H (Bethyl Labs, A301-674A , 1:1000), Cyclin K (Bethyl Labs, A301-939A, 1:1000), CDK2 (Bethyl Labs, A301-812, 1:1,000), Phospho-CDK2 T160(Cell Signaling Technology, 2561, 1:1000) CDK1 (Bethyl Labs, A303-663, 1:1,000), Phospho-CDK1 T161 (Cell Signaling Technology, 9114, 1:1000)Tubulin (Cell Signaling Technology, 3873 1:5,000), PARP (Cell Signaling Technology, 9542, 1:2,000), Phospho-CTD Ser2 (Millipore, 04-1571, 1:2,000), Phospho-CTD Ser5 (Millipore, 04-1572, 1:5,000), Phospho-CTD Ser7 (Millipore, 04-1570-I, 1:1000), Total Pol II (Santa Cruz Biotechnology, sc-899, 1:200). Secondary antibodies were infra-red labeled antibodies (LI-COR, used at 1:10,000) and blots were imaged on an Odyssey CL_Ximager.

Commercial *In Vitro* Kinase Assays

Adapta Eu kinase assays were conducted for CDK7/CycH/MNAT1 and CDK9/CycT1 at Life Technologies using Km ATP concentrations. Z'LYTE kinase assays was conducted for CDK2/CycA at Life Technologies using Km ATP concentrations.

In Vitro Kinase Assays

Recombinant kinases GST-CDK12 (714-1063)/GST-Cyclin K (1-267) and GST-CDK13 (694-1093)/GST-Cyclin K (1-267) were co-expressed with Cdk-Activating Kinase (CAK) from *S. cerevisiae* in *Sf9* insect cells. The tripartite complex GST-CDK7 (2-346)/Cyclin H (1-323)/MAT1 (1-309) was co-expressed from a single vector in baculo-virus infected *Sf9* insect cells using the MultiBac^{Turbo} system (Bieniossek et al., 2012).

Cells expressing the respective kinase complexes were harvested by centrifugation and resuspended in lysis buffer (50 mM HEPES pH 7.6, 500 mM NaCl, 10% glycerol, 5 mM β-mercaptoethanol). Cells were lysed by sonication. All kinases were purified via GST-affinity purification using pre-packed GSTrap 4FF columns (GE Healthcare). The lysate was cleared by centrifugation in a Beckman Optima L-80 XP Ultracentrifuge with a Ti45 rotor (45,000 r.p.m. for 45 min at 4°C) and applied to GSTrap 4FF columns (GE Healthcare) equilibrated with lysis buffer using an ÄKTA PrimePlus chromatography system (GE Healthcare). Following extensive washing with 10 column volumes (CV) of lysis buffer and 5 CV of wash buffer (50 mM HEPES pH 7.6, 1000 mM NaCl, 10% glycerol, 5 mM β-mercaptoethanol), the protein was eluted in elution buffer (50 mM HEPES pH 7.6, 500 mM NaCl, 10% glycerol and 5 mM β-mercaptoethanol, 10 mM glutathione). GST-tags were removed by TEV-protease cleavage overnight at 4°C. Proteins were further purified by size exclusion chromatography on a Superdex 200 PG column (GE Healthcare) equilibrated in 20 mM HEPES pH 7.6, 150 mM NaCl, 1 mM TCEP). Proteins were eluted from the column with the same buffer used for equilibration in 2 mL fractions. Fractions were analyzed by Coomassie staining of SDS-PAGE and fractions containing stoichiometric complexes were pooled and concentrated by ultrafiltration in Amicon filters (Millipore). Kinases were aliquoted in single use aliquots, snap frozen, and stored at -80°C until used.

Radioactive kinase activity measurements were performed at a concentration of 0.2 μM CDK/Cyclin complex. Kinase was incubated with varying concentrations of inhibitory compounds in presence of 1 mM ATP containing 0.45 μCi ^{32}P /μL (Perkin Elmer) for 5 minutes at 30°C prior to starting the kinase reaction by addition of substrate. Reactions were incubated for 15 minutes at 30°C and stopped by adding EDTA to a final concentration of 50 mM. Reaction mixture was spotted onto filter sheets of Amersham Protran nitrocellulose membrane (GE Healthcare). Filters were washed three times for 5 minutes with 0.75% (v/v) phosphoric acid. Radioactivity was counted in a Beckman Liquid Scintillation Counter (Beckman-Coulter) for 1 minute. Data were obtained from three independent experiments and normalized to DMSO control for comparison.

Radioactive *In Vitro* Kinase Assays from Cellular CDK7

For kinase assays Hela (female) cells with stable expression of wildtype or C312S flag-tagged CDK7 were first treated with DMSO, YKL-5-124, or YKL-5-167 for 6 hours. Cells were then harvested by lysis in 50 mM TrisHCl pH 8.0, 150 mM NaCl, 1% NP-40, 5 mM EDTA, Protease Inhibitor (Roche), and Phosstop Phosphatase Inhibitor (Roche). CDK7-FLAG was immunprecipitated from lysates using M2-agarose resin (Sigma, A2220). Precipitated proteins were washed with lysis buffer 4 times, followed by 2 washes with kinase buffer (40 mM Hepes pH 7.5, 150 mM NaCl, 10 mM MgCl2, 5% glycerol) and subjected to *in vitro* kinase assays at 30°C for 45 minutes using 1 μg of the large subunit of RNAPII (RPB1) as substrate and 25 μM ATP and 10 μCi of 32P ATP. Reactions were run on SDS-PAGE gel, fixed and stained with coomassie reagent, and exposed to phospho-imager screen before being scanned on a TYPHOON imager (GE Healthcare Life Sciences).

Pulldown Assays

HAP1 or Jurkat cells were treated with DMSO, YKL-5-124, YKL-5-167, or THZ1 for time and concentration indicated. Cells were trypsinized (HAP1 only), pelleted, washed with 1× PBS and resuspended in 50 mM Tris-HCl pH 8.0, 150 mM NaCl, 1% Nonidet P-40, 5 mM EDTA, 1 mM DTT and protease/phosphatase cocktails. Lysates were clarified and input samples were taken before 1 μM biotinylated-THZ1 probe was added and samples were rotated overnight at 4°C. Lysates with probe were then incubated with streptavidin resin for 3 h at 4°C. Beads were washed with lysis buffer five times and resuspended in 2× LDS load dye. Samples were boiled at 95°C for 5 min before they were loaded onto SDS-PAGE gels and subjected to immunoblotting.

Covalent Inhibition Kinetics Using Mobility Shift Assay

k_{inact}/K_i measurements were determined using kinetic CDK7 enzyme activity data obtained from a peptide substrate mobility shift assay (Blackwell et al., 2009). Recombinant CDK7/CycH/Mat1 and fluorescent peptide substrate were obtained from Carna Biosciences (CDK7/CycH/Mat1, Cat. 04-108; CTD3 peptide, Cat. 04-108MS). Reactions contained 3-5 nM CDK7 enzyme, 1 μM CTD3 peptide and 1 mM ATP in the assay buffer (20 mM HEPES pH7.5, 0.01% TritonX-100, 5 mM $MgCl_2$, 0.5 mM DTT). Phosphorylated product and non-phosphorylated substrate were separated and detected using a LabChip EZ Reader® (PerkinElmer) using the following settings: pressure 1.8 psi and voltage differentail -2400V to -500V with separation buffer consisting of 100 mM Hepes pH 7.3, 0.015% Brij-35, 1 mM disodium EDTA, 0.1% coating reagent 3, 5% DMSO and 1X coating reagent 8, similar to previous reports (Tan et al., 2017) . Nonlinear least-squares regression was performed using the DynaFit software

package (Kuzmic, 1996) using with default settings for a 2-step inactivation model. Figures were generated with GraphPad Prism software (version 6.07).

Mass Spectrometry Analysis

nanoLC/MS

Recombinant CAK complex (CDK7/Cyclin H/MNAT1, Thermo Fisher Scientific) was incubated with YKL-05-124 or DMSO for 4 hours at 37°C, and reactions were resolved by SDS-PAGE. Bands corresponding to CDK7 were excised, digested in gel with trypsin according to standard protocols, and analyzed by nanoLC-MS (Ficarro et al., 2009). Extracted peptides were loaded via autosampler injection (NanoAcquity Sample Manager) onto a precolumn (4 cm POROS 10R2, Applied Biosystems, Framingham, MA) and eluted with an HPLC gradient (NanoAcquity Binary Solvent Manager, Waters; 2-35% B in 60 minutes; A=0.1 M acetic acid in water, B=0.1M acetic acid in acetonitrile). Peptides were resolved on a self-packed analytical column (50 cm Monitor C18, Column Engineering) and introduced to a Q-Exactive HF mass spectrometer (ThermoFisher Scientific) at a flow rate of ~30 nL/min (ESI spray voltage = 2.2 kV). The instrument was operated in data dependent mode such that the 10 most abundant precursors were subjected to MS/MS (HCD, NCE=30%, resolution 15K). mzStudio software (Ficarro et al., 2017) was used to perform peak integrations for the triply and quadruply charged precursors of the doubly carbamidomethylated C312 containing peptide. Two unlabeled, non-cysteine containing CDK7 peptides were used for normalization.

ZipChip CE-MS

Recombinant CAK complex was treated with YKL-05-124 as above, denatured with Rapigest (0.1% final concentration; Waters, Milford, MA), reduced (10 mM dithiothreitol), alkylated (22.5 mM iodoacetamide), and digested with trypsin (Promega) overnight at 37°C. After cleaving Rapigest according to the manufacturer's instructions, peptides were desalted using C18, dried by vacuum centrifugation, and reconstituted in 1% formic acid/50% acetonitrile with 100 mM ammonium acetate. Peptides were then analyzed by CE-MS using a ZipChip CE system and autosampler (908 Devices) interfaced to an Orbitrap Lumos mass spectrometer (ThermoFisher Scientific). Peptide solution was loaded for 30 seconds and separation performed at 500 V/cm on an HR chip for 10 minutes with a background electrolyte consisting of 1% formic acid in 50% acetonitrile. Pressure assist was utilized and started at 1 minute. The mass spectrometer was operated in data dependent mode and subjected the 5 most abundant ions in each MS scan (60k resolution, 1E6 target, lock mass enabled) to MS/MS (15k resolution, 2E5 target, 100 ms max inject time). Dynamic exclusion was enabled with a repeat count of 1 and an exclusion time of 6 seconds. MS/MS data was extracted to .mgf using multiplierz scripts (Alexander et al., 2017; Askenazi et al., 2009) and searched against a custom database containing the sequences of CDK7, Cyclin H, and MAT1 using Mascot version 2.2 (Perkins et al., 1999). Search parameters specified fixed carbamidomethylation of cysteine, and variable oxidation (methionine) and variable YKL-05-124 modification (cysteine). Precursor mass tolerance was set to 10 ppm and product ion tolerance was 25 mmu. To confirm the site of labeling, the mass spectrometer performed cycles of one MS scan (m/z 400-2000, electron multiplier detection, target = 5E4) followed by targeted ETD-MS/MS scans corresponding to the quadruply and quintuply charged precursor ions of the YKL-05-124 modified C312 containing CDK7 peptide (image current detection, 15k resolution, 5E5 target, 250 ms max inject time).

Kinativ Sample Preparation and Analysis

Jurkat cells were treated with DMSO, YKL-5-124 (1μM), or YKL-5-167 (1μM) for 6 hours. Cells were washed with PBS 1X and frozen in liquid nitrogen. Samples were processed by Kinativ (ActivX Biosciences Inc.) as previously described (Patricelli et al., 2011). Briefly, cells were lysed using sonication in lysis buffer (50 mM HEPES, pH 7.5, 150 mM NaCl, 0.1% Triton-X-100, phosphatase inhibitors (Cocktail II AG Scientific #P-1518)) before being cleared by centrifugation. The supernantant is labeled with 20 μM desthiobiotin-ATP acyl-nucleotide probe for 10 minutes to allow covalent engagement of the probe to proteins with available active sites to identify target engagement. Lysates are then denatured (6 M urea, 10 mM DTT, 65°C, 15 min), alkylated(40 mM iodoacetamide, 37°C, 30 min), and filtered into 10 mM ammonium bicarbonate, 2 M urea, 5 mM methionine before digestion with trypsin for 1 hour at 37°C and subjecting peptides to pulldown using high-capacity streptavidin resin (Thermo Scientific). Peptides were was in 50% CH_3CN/water mixture containing 0.1% TFA at room temperature then subjected to mass spectrometric analysis.

Proliferation Assays

Proliferation assays were performed in triplicate by plating Jurkat cells at 2 x 10^4 cells/mL in 50μLin a 384-well plate. Cells were treated with 100nL of YKL-5-124 or THZ1 for final concentrations ranging from 20μM to 600pM for 72 h. Anti-proliferative effects of compounds were assessed by adding 25 μL of Cell Titer Glo (Promega) and luminescence was measured on an Envision plate reader (PerkinElmer). IC_{50} values were determined using Graphpad Prism nonlinear regression curve fit.

Cell Cycle and Growth Rate Analysis

HAP1 WT and C312S mutant cells were plated at a density of ranging 500 cells per well in 384-well Cell Carrier plates (Perkin Elmer) using a Multidrop Combi Reagent Dispenser (Thermo Fisher Scientific) and allowed to adhere to for 24 hours prior to drug treatment. Cells were treated with a dilution series of the indicated drugs using a D300 Digital Dispenser (Hewlett-Packard).

Cells were stained and fixed for analysis at the time of drug delivery and after 72 hours of incubation. Cells were pulsed for one hour with EdU (Lumiprobe) and stained with 1:2000 LIVE/DEAD Far Red Dead Cell Stain (LDR) (Thermo Fisher Scientific). Cells were then fixed with 3.7% formaldehyde (Sigma Aldrich) for 30 minutes and permeabilized with 0.5% Triton X-100 in PBS. The EdU was labeled with cy3-azide (Lumiprobe) for 30 min. The cells were then blocked for one hour with Odyssey blocking buffer (LI-COR), and stained overnight at 4°C with 2 μg/ml Hoechst 33342 (Sigma Aldrich) and a 1:1000 dilution of anti-phospho-histone H3 (pHH3)Alexa 488 (Ser10, clone D2C8) conjugated antibody (Cell Signaling Technologies). Fixed cells were imaged with a 10x objective using an Operetta microscope and analyzed using the Columbus image data storage and analysis system (Perkin Elmer). Nuclei were segmented using Columbus software (Perkin Elmer) based on their Hoechst signal. DNA content was defined by the total Hoechst intensity within the nuclear mask to identify cells in the G1 and G2 phases of the cell cycle. The average LDR, EdU and phospho-histone H3 intensities within the nuclear masks were determined. The LDR signal was used to classify cells as live or dead, the EdU and pHH3 signals to identify S and M phase cells respectively. Cells with intermediate DNA content and no EdU signal were classified as S phase dropout cells. Live cell counts were normalized to DMSO-treated controls on the same plates to yield normalized growth rate inhibition (GR) values as described previously (Hafner et al., 2016). Experiments were performed as technical duplicates in biological triplicate.

RNA Sequencing

1×10^6 HAP1 Cells were treated for 24 h with 500nM or 2μM YKL-5-124, 100nM THZ1, 250nM THZ531, or the combination of 250nM THZ531 and 500nM or 2μM YKL-5-124. Library preparation was performed using TruSeq Stranded mRNA Library Prep Kit.

RNA-seq Analysis

All RNA-seq analysis was performed using the human reference genome build HG19 with HG19/GRCh37 Refseq gene annotations. Data were aligned using Hisat2 with default parameters to the UCSC Refseq transcriptome as provided by the Illumina iGenomes project (https://support.illumina.com/sequencing/sequencing_software/igenome.html). Raw and processed data are deposited into the GEO: GSE124607.

DESeq Analysis

Differential Expression analysis was performed using the "DESeq2" package in R. Counts were first generated from Hisat2 outputted single-end bam files in R via "Rsamtools". These counts were then filtered based on minimum count number (>10). DESeq datasets were created using these filtered counts to compare counts between the treatment and DMSO.

Jaccard Heatmap Analysis

Heatmaps were generated based on gene set similarity scores to show the amount of overlap in genes downregulated by each treatment. Jaccard scores were calculated based on union and intersect of genes downregulated in the two treatments being compared. A table consisting of these scores was imported into python. The "seaborn" and "matplotlib.pyplot" packages were used to generate the heatmap.

Gene Set Enrichment Analysis

Gene set enrichment analysis was performed using the GSEA computational platform developed at the Broad Institute. This analysis required expression data, phenotype data, and a gene set database. Expression data was derived from FPKM data which was used to create a Gene Cluster Text file (.gct). This file is comprised of FPKM averages across the treatment and the control. The phenotype data was input in the form of a Categorical Class file (.cls). This file featured information regarding the number of classes and samples in our analysis as well as classification for each of the classes. Finally, we utilized the C2 all curated gene set directly from the GSEA website for the gene set database in the form of a Gene Matrix Transposed file (.gmt). All other parameters were kept at defaults with the exception of the metric for ranking genes which changed from "Signal2Noise" to "Ratio_of_Classes".

Radar Plot Generation

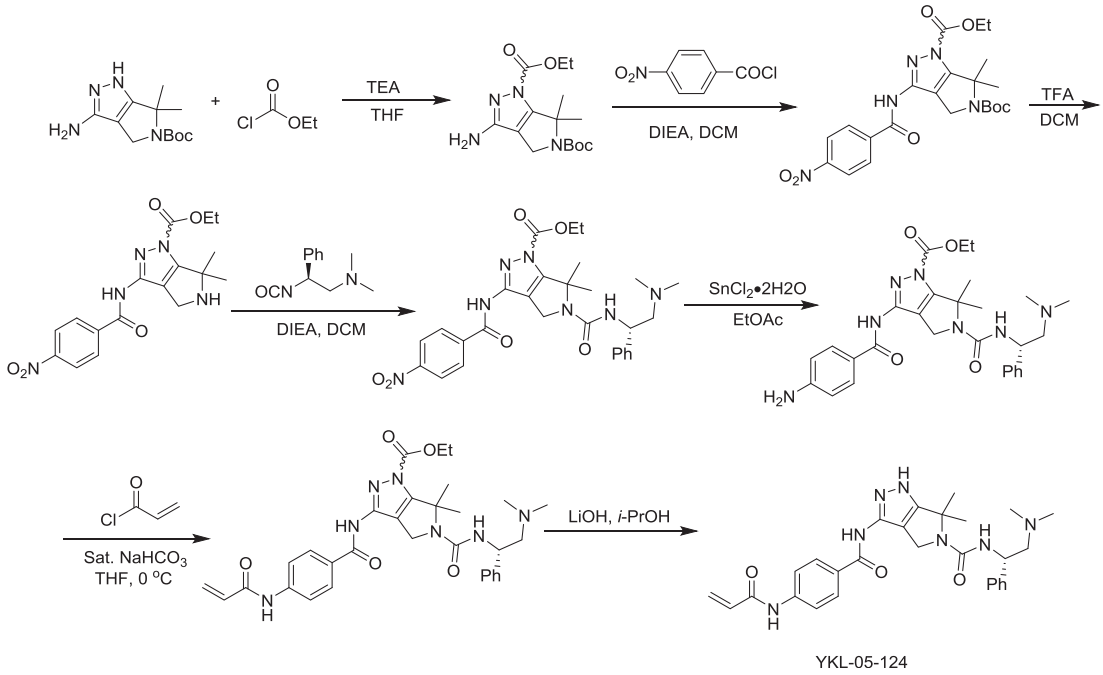
Radar plots were created via the "fmsb" package in R to examine differences in gene expression changes across different treatments. This was performed both on a specific cell cycle related gene set derived from GSEA as well as specific gene examples. Log_2 fold change values were used to determine magnitudes of gene expression changes between the treatment and control conditions. Calculation of these values was done in R for both the wildtype and YKL-5-124 refractory cell lines. Following compilation of the log_2 fold change values into a separate data frame, the "fmsb" package was used to create radar plots with overlapping plots for both cell lines.

Single Gene Expression Analysis

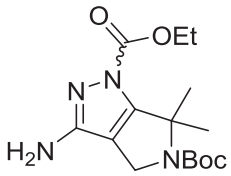
Specific genes from the GSEA derived cell cycle gene set were further analyzed for changes in expression across different dosages of YKL-5-124 treatment in both the wildtype and mutant cell lines. To do this, FPKM averages were calculated across the triplicate for the DMSO control and the two dosages of YKL-5-124. These averages were combined with standard deviation values in a single data frame. Using "ggplot" in R, we generated a bar plot for expression of these specific genes across different dosages in both cell lines. Significance levels in changes of expression were calculated via the paired Student's t-test.

Chemical Synthesis

Synthesis of (R)-3-(4-acrylamidobenzamido)-N-(2-(dimethylamino)-1-phenylethyl)-6,6-dimethyl-4,6-dihydropyrrolo [3,4-c]pyrazole-5(1H)-carboxamide (YKL-05-124)

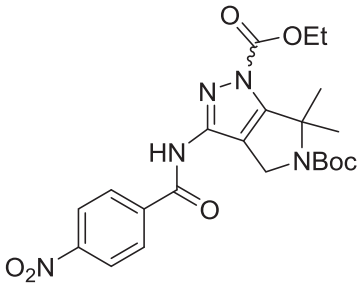


5-(tert-butyl) 1-ethyl 3-amino-6,6-dimethyl-4,6-dihydropyrrolo[3,4-c]pyrazole-1,5-dicarboxylate (1- and 2-protected ethyl Formate was Obtained as an Inseparable Mixture, Which was Used Together Until The Last Step)



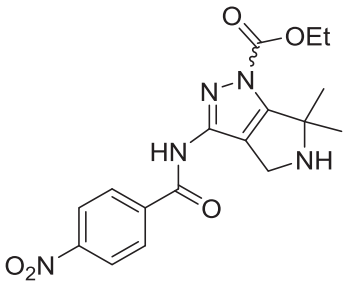
To a solution of *tert*-butyl 3-amino-6,6-dimethyl-4,6-dihydropyrrolo[3,4-c]pyrazole-5(1*H*)-carboxylate (1.0 g, 3.95 mmol) and TEA (0.6 g, 0.82 mL, 5.93 mmol) in THF (10 mL) was added ethyl chloroformate (0.43 g, 0.38 mL, 3.95 mmol) dropwise at 0°C. The mixture was stirred at 0°C for 1 h. The solvent was removed and the residue was partitioned with EtOAc and sat. $NaHCO_3$. The organic layer was washed with water, brine and dried with Na_2SO_4. The solvent was then removed under vacuum to provide 5-(*tert*-butyl) 1-ethyl 3-amino-6,6-dimethyl-4,6-dihydropyrrolo[3,4-*c*]pyrazole-1,5-dicarboxylate as an off white solid (1.28 g, 100 %). LC/MS (ESI) m/z = 325 $(M + H)^+$.

5-(tert-butyl) 1-ethyl 6,6-dimethyl-3-(4-nitrobenzamido)-4,6-dihydropyrrolo[3,4-c]pyrazole-1,5-dicarboxylate



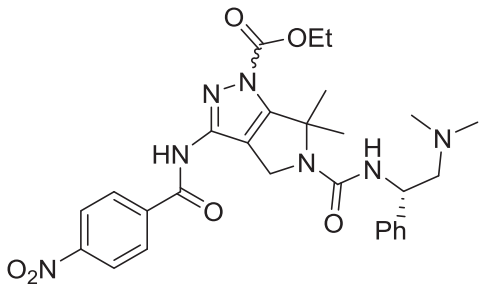
To a solution of 5-(*tert*-butyl) 1-ethyl 3-amino-6,6-dimethyl-4,6-dihydropyrrolo[3,4-*c*]pyrazole-1,5-dicarboxylate (162 mg, 0.5 mmol) and DIEA (98 μL, 1.0 mmol) in DCM (5 mL) was added 4-nitrobenzoyl chloride (110 mg, 0.6 mmol) at 0°C. The reaction was stirred at room temperature for 2 h and concentrated. The crude was purified by flash column chromatography on silica gel (EtOAc/hexanes, 0-70%) to provide 5-(*tert*-butyl) 1-ethyl 6,6-dimethyl-3-(4-nitrobenzamido)-4,6-dihydropyrrolo[3,4-*c*]pyrazole-1,5-dicarboxylate (177 mg, 75%). LC/MS (ESI) m/z = 474 $(M + H)^+$.

1-Ethyl 6,6-dimethyl-3-(4-nitrobenzamido)-5,6-dihydropyrrolo[3,4-c]pyrazole-1(4H)-carboxylate



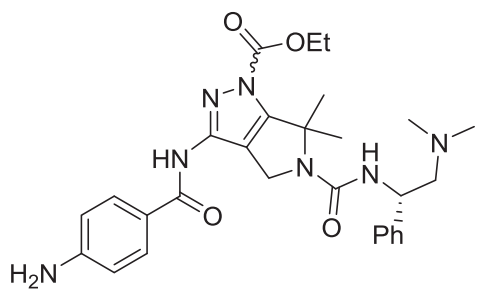
To a solution of 5-(*tert*-butyl) 1-ethyl 6,6-dimethyl-3-(4-nitrobenzamido)-4,6-dihydropyrrolo[3,4-*c*]pyrazole-1,5-dicarboxylate (177 mg, 0.23 mmol) in DCM (1 mL) was added TFA (1 mL). The reaction was stirred at room temperature for 1 h and concentrated to provide 1-ethyl 6,6-dimethyl-3-(4-nitrobenzamido)-5,6-dihydropyrrolo[3,4-*c*]pyrazole-1(4*H*)-carboxylate as a TFA salt, which was used directly in next step. LC/MS (ESI) m/z = 374 $(M + H)^+$.

1-Ethyl (S)-5-((2-(dimethylamino)-1-phenylethyl)carbamoyl)-6,6-dimethyl-3-(4-nitrobenzamido)-5,6-dihydropyrrolo [3,4-c]pyrazole-1(4H)-carboxylate



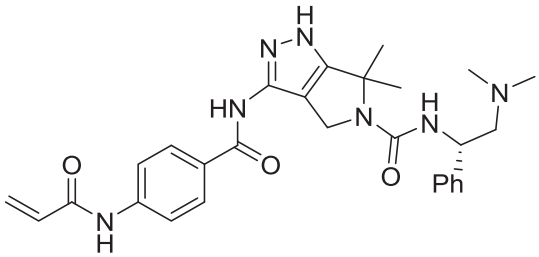
To a mixture of ethyl 6,6-dimethyl-3-(4-nitrobenzamido)-5,6-dihydropyrrolo[3,4-*c*]pyrazole-1(4*H*)-carboxylate (TFA salt, 0.23 mmol) in DCM (2 mL) was added DIEA (148 mg, 0.2 mL, 1.15 mmol) at 0°C, followed by (*S*)-2-isocyanato-*N*,*N*-dimethyl-2-phenylethan-1-amine (62 mg, 0.27 mmol). The solution was stirred at 0°C for 1 h and diluted with $CHCl_3$/*i*-PrOH (v/v = 4:1) and washed with sat. $NaHCO_3$, brine and dried with Na_2SO_4. The solvent was then removed and the crude was purified by flash column chromatography on silica gel to provide 1-ethyl (*S*)-5-((2-(dimethylamino)-1-phenylethyl)carbamoyl)-6,6-dimethyl-3-(4-nitrobenzamido)-5,6-dihydropyrrolo[3,4-*c*]pyrazole-1(4*H*)-carboxylate (113 mg, 85% over 2 steps). LC/MS (ESI) m/z = 564 $(M + H)^+$.

1-Ethyl (S)-3-(4-aminobenzamido)-5-((2-(dimethylamino)-1-phenylethyl)carbamoyl)-6,6-dimethyl-5,6-dihydropyrrolo [3,4-c]pyrazole-1(4H)-carboxylate



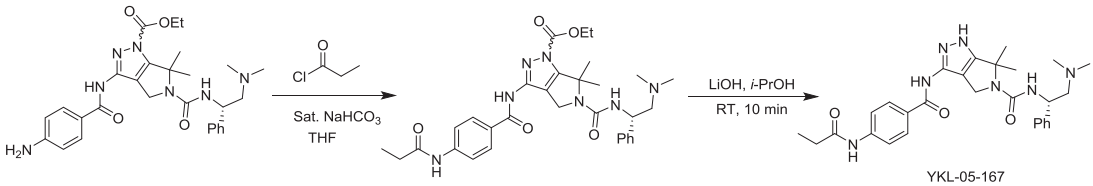
To a solution of 1-ethyl (S)-5-((2-(dimethylamino)-1-phenylethyl)carbamoyl)-6,6-dimethyl-3-(4-nitrobenzamido)-5,6-dihydropyrrolo[3,4-*c*]pyrazole-1(4*H*)-carboxylate (113 mg, 0.20 mmol) in EtOAc (2 mL) was added $SnCl_2{\cdot}2H_2O$ (226 mg, 1.0 mmol). The mixture was stirred at 70°C for 2 h and diluted with $CHCl_3$/*i*-PrOH (v/v = 4:1) and washed with sat. $NaHCO_3$, brine and dried with Na_2SO_4. The crude was purified by flash column chromatography on silica gel to provide 1-ethyl (*S*)-3-(4-aminobenzamido)-5-((2-(dimethylamino)-1-phenylethyl)carbamoyl)-6,6-dimethyl-5,6-dihydropyrrolo[3,4-*c*]pyrazole-1(4*H*)-carboxylate (193 mg, 88%). LC/MS (ESI) *m/z* = 534 $(M + H)^+$.

(S)-3-(4-acrylamidobenzamido)-N-(2-(dimethylamino)-1-phenylethyl)-6,6-dimethyl-4,6-dihydropyrrolo[3,4-c] pyrazole-5(1H)-carboxamide (YKL-05-124)



To a mixture of 1-ethyl (*S*)-3-(4-aminobenzamido)-5-((2-(dimethylamino)-1-phenylethyl)carbamoyl)-6,6-dimethyl-5,6-dihydropyrrolo[3,4-*c*]pyrazole-1(4*H*)-carboxylate (20 mg, 0.032 mmol) in THF (1 mL) and sat. $NaHCO_3$ (1 mL) was added acryloyl choride (5.8 mg, 0.064 mmol) dropwise at 0°C. The mixture was stirred at 0°C for 10 min and diluted with $CHCl_3$/*i*-PrOH (v/v 4:1) and washed with sat. $NaHCO_3$ and brine, dried with Na_2SO_4 and then concentrated. The crude was then dissolved in 1.0 mL of *i*-PrOH. The solution was then treated with LiOH (1 M, 1 mL) and was stirred at room temperature for 10 min. The mixture was diluted with $CHCl_3$/*i*-PrOH (v/v 4:1) and washed with sat. $NaHCO_3$ and brine, dried (Na_2SO_4), and concentrated. The crude was purified by HPLC (MeOH/H_2O, 0-100%) to provide YKL-05-124 as TFA salt (8.1 mg, 40%). LC/MS (ESI) *m/z* = 516 $(M + H)^+$. ^1H NMR (500 MHz, DMSO-d_6) δ 10.89 (s, 1H), 10.72 (s, 1H), 9.49 (s, 1H), 8.03 (d, *J* = 8.8 Hz, 2H), 7.85 (d, *J* = 8.8 Hz, 2H), 7.46 (d, *J* = 7.4 Hz, 2H), 7.40 (dd, *J* = 8.5, 6.8 Hz, 2H), 7.30 (td, *J* = 7.0, 1.4 Hz, 1H), 6.79 (d, *J* = 9.1 Hz, 1H), 6.57 (dd, *J* = 16.9, 10.2 Hz, 1H), 6.31 (dd, *J* = 17.0, 1.9 Hz, 1H), 5.81 (dd, *J* = 10.1, 2.0 Hz, 1H), 5.37 (ddd, *J* = 12.4, 9.0, 3.8 Hz, 1H), 4.82 (d, *J* = 11.9 Hz, 1H), 4.56 (d, *J* = 11.9 Hz, 1H), 3.57 (td, *J* = 12.5, 2.8 Hz, 1H), 2.87 (d, *J* = 4.8 Hz, 3H), 2.83 (d, *J* = 4.8 Hz, 3H), 1.68 (s, 3H), 1.60 (s, 3H).

Synthesis of (S)-N-(2-(dimethylamino)-1-phenylethyl)-6,6-dimethyl-3-(4-propionamidobenzamido)-4,6-dihydropyrrolo [3,4-c]pyrazole-5(1H)-carboxamide (YKL-05-167)



To a mixture of 1-ethyl (*S*)-3-(4-aminobenzamido)-5-((2-(dimethylamino)-1-phenylethyl)carbamoyl)-6,6-dimethyl-5,6-dihydropyrrolo[3,4-*c*]pyrazole-1(4*H*)-carboxylate (20 mg, 0.032 mmol) in THF (1 mL) and sat. $NaHCO_3$ (1 mL) was added propionyl chloride (5.9 mg, 0.064 mmol) dropwise at 0°C. The mixture was stirred at 0°C for 10 min and diluted with $CHCl_3$/i-PrOH (v/v 4:1) and washed with sat. $NaHCO_3$ and brine, dried (Na_2SO_4), and concentrated.

The crude was then dissolved in 1.0 mL of *i*-PrOH. The solution was then treated with LiOH (1 M, 1 mL) and was stirred at room temperature for 10 min. The mixture was diluted with $CHCl_3$/*i*-PrOH (v/v 4:1) and washed with sat. $NaHCO_3$ and brine, dried (Na_2SO_4), and concentrated. The crude was purified by HPLC (MeOH/H_2O, 0-100%) to provide YKL-05-167 as TFA salt (9.0 mg, 45%). LC/MS (ESI) *m/z* = 518 $(M + H)^+$. H NMR (600 MHz, DMSO-d_6) δ 12.38 (s, 1H), 10.71 (s, 1H), 10.12 (s, 1H), 7.95 (m, 2H), 7.69 (m, 2H), 7.35 (d, *J* = 7.0 Hz, 2H), 7.28 (t, *J* = 7.6 Hz, 2H), 7.18 (d, *J* = 7.6 Hz, 1H), 6.25 (m, 1H), 4.87 (m, 1H), 4.52 (m, 2H), 2.68 (m, 1H), 2.41 (m, 1H), 2.34 (q, *J* = 7.6 Hz, 2H), 2.20 (m, 6H), 1.62 (m, 3H), 1.55 (s, 3H), 1.08 (t, *J* = 7.6 Hz, 1H).

QUANTIFICATION AND STATISTICAL ANALYSIS

The method of determining error bars is indicated in the corresponding figure legend with the replicate number also indicated. Statistical tests for RNA-seq data is outlined in the STAR Methods section under the relevant analysis. Data met the assumptions for all tests used.

DATA AND SOFTWARE AVAILABILITY

The accession number for the raw RNA-seq data reported in this paper is GEO: GSE124607. Original data can be found at https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE124607.