Published in final edited form as:
Nature. 2014 July 31; 511(7511): 616–620. doi:10.1038/nature13393.

Targeting transcription regulation in cancer with a covalent CDK7 inhibitor

Nicholas Kwiatkowski[1,2,3,*], **Tinghu Zhang**[1,2,*], **Peter B Rahl**[3], **Brian J Abraham**[3], **Jessica Reddy**[3,4], **Scott B Ficarro**[1,2,5], **Anahita Dastur**[6], **Arnaud Amzallag**[6,7], **Sridhar Ramaswamy**[6,7], **Bethany Tesar**[8,9], **Christopher R Jenkins**[10], **Nancy M Hannett**[3], **Douglas McMillin**[8,9], **Takaomi Sanda**[11,12], **Taebo Sim**[13], **Nam Doo Kim**[14], **Thomas Look**[11,15], **Constantine Mitsiades**[8,9], **Andrew P Weng**[10], **Jennifer R Brown**[8,9], **Cyril H. Benes**[6], **Jarrod A Marto**[1,2,5], **Richard A Young**[3,4], and **Nathanael S. Gray**[1,2]

[1]Department of Cancer Biology, Dana-Farber Cancer Institute, Boston, MA 02115, USA

[2]Department of Biological Chemistry and Molecular Pharmacology, Harvard Medical School, Boston, MA 02115, USA

[3]Whitehead Institute for Biomedical Research, 9 Cambridge Center, Cambridge, MA 02142, USA

[4]Department of Biology, Massachusetts Institute of Technology, Cambridge, MA 02139, USA

[5]Blais Proteomics Center, Dana-Farber Cancer Institute, Boston, MA 02115, USA

[6]Department of Medicine Massachusetts General Hospital Cancer Center and Harvard Medical School, Charlestown, MA 02129, USA

[7]Broad Institute of MIT and Harvard, 7 Cambridge Center, Cambridge, MA 02142, USA

[8]Department of Medical Oncology, Dana-Farber Cancer Institute, Harvard Medical School, Boston, MA 02115, USA

[9]Department of Medicine, Brigham and Women's Hospital, Harvard Medical School, Boston, MA 02115, USA

[10]Terry Fox Laboratory, British Columbia Cancer Agency, Vancouver, British Columbia, Canada

Correspondence and requests for materials should be addressed to N.S.G. (nathanael_gray@dfci.harvard.edu) or R.A.Y. (young@wi.mit.edu).
*These authors contributed equally to this work.

Supplementary Information is linked to the online version of the paper at www.nature.com/nature.

Author Contributions: N.S.G., N.K., and T.Z. conceived the project. N.S.G. and T.Z. conceived and directed chemistry effort with input from T.S.[11]. T.Z. performed chemical synthesis and small molecule structure determination. R.A.Y., N.S.G., P.B.R., and N.K. conceived genomics effort. N.K., P.B.R., J.R.B., C.H.B., N.S.G., and R.A.Y. designed biological experimental research with input from J.R., B.J.A., D.M., T.S.[9,10], T.L., A.P.W. and C.M. N.K., P.B.R., J.R., A.D., B.T., C.R.J., and N.M.H performed experimental biological research. S.B.F. designed and performed protein mass spectrometry on THZ1/ CDK7 adducts with input from J.A.M. S.R. and A.A. performed elastic net regression analysis. N.D.K. performed molecular modeling studies with input from T.S.[11] B.J.A. designed and performed genomics data analyses. N.K., N.S.G. and R.A.Y. co-wrote the paper. All authors edited the manuscript. N.K. and T.Z. contributed equally to this work.

Sequencing and expression data were deposited in the Gene Expression Omnibus (http://www.ncbi.nlm.nih.gov/geo) under accession number GSE50625.

The authors declare the following competing interests N.S.G, T.Z., N.K. are inventors on a patent application covering THZ1 which is licensed to a company co-founded by N.S.G and R.A.Y.

[11]Department of Pediatric Oncology, Dana-Farber Cancer Institute, Harvard Medical School, Boston, MA 02215, USA

[12]Cancer Science Institute of Singapore, National University of Singapore, Singapore 117599

[13]Chemical Kinomics Research Center, Korea Institute of Science and Technology, 39-1, Hawolgok-dong, Seongbuk-gu, Seoul, 136-791, Korea/ KU-KIST Graduate School of Converging Science and Technology, 145, Anam-ro, Seoul, Korea

[14]Daegu-Gyeongbuk Medical Innovation Foundation, 2387 dalgubeol-daero, Suseong-gu, Daegu, 706-010, Korea

[18]Division of Hematology/Oncology, Children's Hospital, Boston, MA 02115 USA

Abstract

Tumor oncogenes include transcription factors that co-opt the general transcriptional machinery to sustain the oncogenic state[1], but direct pharmacological inhibition of transcription factors has thus far proven difficult[2]. However, the transcriptional machinery contains various enzymatic co-factors that can be targeted for development of new therapeutic candidates[3], including cyclin-dependent kinases (CDKs)[4]. Here we present the discovery and characterization of the first covalent CDK7 inhibitor, THZ1, which has the unprecedented ability to target a remote cysteine residue located outside of the canonical kinase domain, providing an unanticipated means of achieving selectivity for CDK7. Cancer cell line profiling indicates that a subset of cancer cell lines, including T-ALL, exhibit exceptional sensitivity to THZ1. Genome-wide analysis in Jurkat T-ALL shows that THZ1 disproportionally affects transcription of *RUNX1* and suggests that sensitivity to THZ1 may be due to vulnerability conferred by the *RUNX1* super-enhancer and this transcription factor's key role in the core transcriptional regulatory circuitry of these tumor cells. Pharmacological modulation of CDK7 kinase activity may thus provide an approach to identify and treat tumor types exhibiting extreme dependencies on transcription for maintenance of the oncogenic state.

In an effort to discover new inhibitors of kinases that regulate gene transcription, we performed cell-based screening and kinase selectivity profiling of a library of known and novel ATP-site directed kinase inhibitors (See Supplementary Table 1 for known CDK7 inhibitors). We identified THZ1 (Fig. 1a), a phenylaminopyrimidine bearing a potentially cysteine-reactive acrylamide moiety, as a low nanomolar inhibitor of cell proliferation and biochemical CDK7 activity (Fig. 1b, c). To investigate the functional relevance of the acrylamide moiety we prepared a non-cysteine reactive analog THZ1-R, which displayed diminished activity for CDK7andreduced anti proliferative potency (Fig. 1b, c). KiNativ™ profiling[5], which measures the ability of a compound to block nucleotide-dependent enzymes from biotinylation with a reactive desthiobiotin-ATP probe, established CDK7 as the primary intracellular target of THZ1, but not of THZ1-R (Supplementary Table 2). Kinome-wide profiling identified additional kinase targets of THZ1; however, we confirmed CDK7 as the only target displaying time-dependent inhibition, which is suggestive of covalent binding (Extended Data Fig. 1a–c and Supplementary Table 3).

As no covalent inhibitors of CDKs have been reported, we next focused our studies on the mechanism by which THZ1 could achieve covalent inhibition of CDK7. We first incubated recombinant CDK7/cyclin H/MAT1 trimeric complex with a biotinylated version of THZ1 (bio-THZ1, Fig. 1a) and demonstrated that it indeed covalently modifies CDK7 (Fig. 2a; Extended Data Fig. 1d–g). Mass spectrometry identified the site of covalent modification as C312, a residue located outside the kinase domain (Extended Data Fig. 2a–d). Inspection of the crystal structure reveals that a C-terminal extension of CDK7 bearing C312 traverses the ATP cleft in the kinase domain and would be predicted to position Cys312 directly adjacent to the reactive acrylamide moiety of THZ1 (Fig. 2b). Mutation to serine (C312S), a less nucleophilic amino acid, prevented THZ1 from covalently binding to CDK7 and from inhibiting CDK7 activity in an irreversible fashion (Fig. 2c; Extended Data Fig. 2e). Sequence alignment of the 20-member CDK family suggests that Cys312 is unique to CDK7, however CDK12 and CDK13 also possess accessible cysteines within 4 amino acids of Cys312 (Extended Data Fig. 3a). Indeed, we found that THZ1 can inhibit CDK12 kinase activity at slightly higher concentrations (Extended Data Fig. 3b–f). THZ1 is the first inhibitor demonstrated to target a cysteine located outside of the kinase domain, which provides an unanticipated means of achieving covalent selectivity.

CDK7 kinase activity has been implicated in the regulation of both transcription, where it phosphorylates the C-terminal domain (CTD) of RNAP polymerase II (RNAPII)[6–8] and CDK9[9], and the cell cycle, where it functions as the CDK-activating kinase (CAK) for CDKs1/2/4/6[10–14]. THZ1, but not THZ1-R, completely inhibits the phosphorylation of the established intracellular CDK7 substrate RNAPII CTD at Ser-5 and Ser-7[6,8], with concurrent loss of Ser-2 phosphorylation at 250 nM in Jurkat cells (Fig. 2d). Cellular washout experiments demonstrate that THZ1 indeed acts in an irreversible fashion (Fig. 2e, f; Extended Data Fig. 4a–e). We observed a loss of CAK activity, as evidenced by decreased phosphorylation of the activation loops of CDK1,2 and 9, indicating disruption of both recognized CDK7 signaling pathways in Jurkat cells (Fig. 2d; Extended Data Fig. 4f, g) and Loucy cell lines (Extended Data Fig. 4). Ectopic expression of dox-inducible FLAG-CDK7 C312S, but not FLAG-CDK7 WT, in Hela S3 cells restored RNAPII CTD p-Ser 5/7 to near WT levels at concentrations of THZ1 up to 2.5 μM, establishing C312 as a critical determinant of the cellular pharmacology of the inhibitor (Extended Data Fig. 5a–b). Additionally, FLAG-CDK7 C312S expression restored CDK1/2 T-loop phosphorylation, reduced early induction of cleaved PARP and restored the expression of a subset of genes, including the highly –expressed transcription factors MYC, KLF4, ID1, and GATA2 (Extended Data Fig. 5c–e). The partial rescues of the hyperphosphorylated form of RNAPII (RNAPII0) and RNAPII p-Ser2 CTD phosphorylation combined with the incomplete restoration of gene expression may result, in part, from lower affinity cross-reactivity of THZ1 with CDK12/13, which are bona fide Ser2 kinases[15].

Our evidence that CDK7 inhibition leads to reduction in RNAPII CTD phosphorylation status appears in conflict with evidence that inhibition of CDK7 alone is insufficient to reduce RNAPII CTD phosphorylation in HCT116 cells[9]. It is possible that covalent inhibition and reversible inhibition can engender different effects on kinase structure; we did not find evidence that THZ1 impacts TFIIH or CAK complex stability (Extended Data Fig. 4h). It is also possible that inhibition of CDK12/13 (or another undetected kinase)

contributes to reduced RNAPII CTD phosphorylation, although our evidence that RNAPII CTD phosphorylation levels are restored following expression of CDK7 C312S suggests otherwise.

To better understand the breadth of antiproliferative activity of THZ1, we screened it against a diverse panel of over 1,000 cancer cell lines[16]. THZ1 displayed broad-based activity with IC_{50}s less than 200 nM against 53% of the cell lines tested (Fig. 3a; Supplementary Table 4). Elastic net regression analysis incorporating gene expression, copy number, and sequence variation genomics data[16] across 527 of the cell lines tested were used to identify genomic features common to sensitive cell lines. Gene ontology (GO) term enrichment analysis[17] indicated a strong enrichment of (proto-) oncogenic transcription factors commonly overexpressed in cancer and factors involved in RNAPII-driven transcriptional regulation, suggesting the dominant activity of THZ1 was through modulation of transcription (Fig. 3b; Supplementary Table 5).

In agreement with the net elastic regression analysis, T-ALL cell lines that display characteristic misregulation of T-cell lineage-specific transcription factors, were broadly sensitive to THZ1, but not THZ1-R (Fig. 3c; Extended Data Fig. 6a and Supplementary Table 4). Treatment of T-ALL cell lines with THZ1 caused decreased cellular proliferation and an increase in apoptotic index with concomitant reduction in anti-apoptotic proteins, most notably MCL-1 and XIAP (Extended Data Fig. 6, 7). These strong anti proliferative responses induced at sub-effective doses of THZ1 suggest that these cells may be particularly sensitive to small perturbations in transcription and CDK7 kinase function. Indeed, THZ1 demonstrated efficacy against primary leukemia cells and in a bioluminescent xenografted model using the human T-ALL cell line, KOPTK1, when dosed twice daily (BID) at 10 mg/kg (Fig. 3d, e; Extended Data Fig. 8 and Supplementary Table 6). Importantly, THZ1 was well tolerated at these doses with no observable body weight loss or behavioral changes (Extended Data Fig. 8f), suggesting no overt toxicity to the animals. These results were mirrored in cell culture with non-transformed BJ fibroblast and retinal pigment epithelial (RPE-1) cells responding to relatively high doses of THZ1 by undergoing cell cycle arrest rather than initiating apoptosis or cell death, further suggesting that normal cells might tolerate transcriptional disruption (Extended Data Fig. 9).

CDK7 is a component of the general transcription factor IIH (TFIIH) complex[18–20], so we next investigated how THZ1 treatment affects genome-wide gene expression. We chose Jurkat T-ALL for these studies because it a well-studied T-ALL cell line model with a defined core transcriptional regulatory circuitry, consisting of key transcription factors, which is also found in human T-ALL primagrafts[21]. Treatment with 250 nM THZ1, but not THZ1-R, led to progressive reduction in global steady-state mRNA levels over time, with 75% and 96% of mRNAs showing greater than 2-fold reduction by 6 and 12 hrs, respectively (Fig. 4a, Extended Data Fig. 10a and Supplementary Table 7). Consistent with global downregulation of mRNA transcripts, 250 nM THZ1 reduced RNAPII occupancy genome-wide at both promoters and gene bodies (Fig. 4b). By comparison, Flavopiridol reduced RNAPII density across only gene bodies (Fig. 4b). This is consistent with the model that CDK7 regulates RNAPII initiation and pausing while CDK9 regulates pause release leading to processive elongation[8,9,18–20,22–24].

Although 250 nM THZ1 inhibits global transcription, we found that some cancer cell lines, particularly T-ALL, are sensitive to considerably lower concentrations of THZ1. We postulated that the expression of certain genes might be especially sensitive to low doses of THZ1 and therefore have a key role in driving the cellular response. Indeed, we found that transcripts for only a subset of genes were substantially affected by treatment with 50 nM THZ1, with that for RUNX1 among the most profoundly affected (Fig. 4c). There are at least two reasons that low dose THZ1 treatment might cause a preferential loss of RUNX1 expression. Tumor cell oncogenes can acquire super-enhancers, which drive high-level expression yet can be especially sensitive to perturbation[25–28]. Super-enhancer analysis in Jurkat cells revealed that *RUNX1* contains an exceptionally large super-enhancer domain containing a previously described hematopoietic cell –specific enhancer (Fig. 4d; Extended Data Fig. 10b–d and Supplementary Table 8)[29]. In addition, RUNX1 forms a core regulatory circuitry with two additional transcription factors that play prominent roles in leukemia biology, TAL1 and GATA3 (Fig. 4e)[21]. These factors auto regulate their own gene expression while simultaneously regulating many other genes that comprise the active gene expression program of Jurkat cells. Treatment with 50 nM THZ 1l ed to significant reduction in both the transcript and protein levels of *RUNX1*, *TAL1*, and *GATA3* (Extended Data Fig. 10e and f). Loss of the *RUNX1* driven transcriptional program is likely key to the response to low dose THZ1 treatment, as gene set enrichment analysis revealed that the Jurkat transcripts downregulated by 50 nM THZ1 were enriched in transcripts similarly downregulated following *RUNX1* depletion using shRNA (Fig. 4f).

Here we have reported the discovery and characterization of the first covalent inhibitor of CDK7, THZ1. THZ1 employs a unique mechanism, combining ATP-site and allosteric covalent binding, as means of attaining potency and selectivity for CDK7. This mechanistic insight should be useful for designing next generation inhibitors of CDKs, where high sequence and shape homology in the ATP pocket has posed a formidable challenge to achieving selectivity with conventional ATP-competitive inhibitors. THZ1 displayed exquisite anti proliferative activity for T-ALL cell lines and other blood cancers, where oncogenic transcription factors feature prominently in the disease state.

In Jurkat cells, low dose THZ1 had a profound effect on a small subset of genes, including the key regulator *RUNX1*, thus contributing to subsequent loss of the greater gene expression program and cell death. Identification of additional cancer cell lines whose gene expression programs display vulnerability to THZ1 or other transcriptional inhibitors should delineate additional cancers that are exquisitely susceptible to perturbation of transcription.

Methods Summary

T-ALL culture conditions

Jurkat, Loucy, KOPTK1, and DND-41 cell lines were grown in RPMI-1640 supplemented with 10% fetal bovine serum and 1% glutamine. All cell lines were cultured at 37°C in a humidified chamber in the presence of 5% CO_2, unless otherwise noted.

Inhibitor treatment experiments

Time-course experiments such as those described in Extended Data Fig. 5a were conducted to determine the minimal time required for full inactivation of CDK7. Cells were treated with THZ1, THZ1-R, or DMSO for 0–6 hrs to assess the effect of time on the THZ1 – mediated inhibition of RNAPII CTD phosphorylation. For subsequent experiments cells were treated with compounds for 4 hrs as determined by time-course experiment described above, unless otherwise noted. For inhibitor washout experiments (Fig. 2e, f; Extended Data Fig. 5) cells were treated with THZ1, THZ1-R, or DMSO for 4 hrs. Medium containing inhibitors was subsequently removed to effectively 'washout' the compound and the cells were allowed to grow in the absence of inhibitor. For each experiment, lysates were probed for RNAPII CTD phosphorylation and other specified proteins.

High-throughput cell line panel viability assay

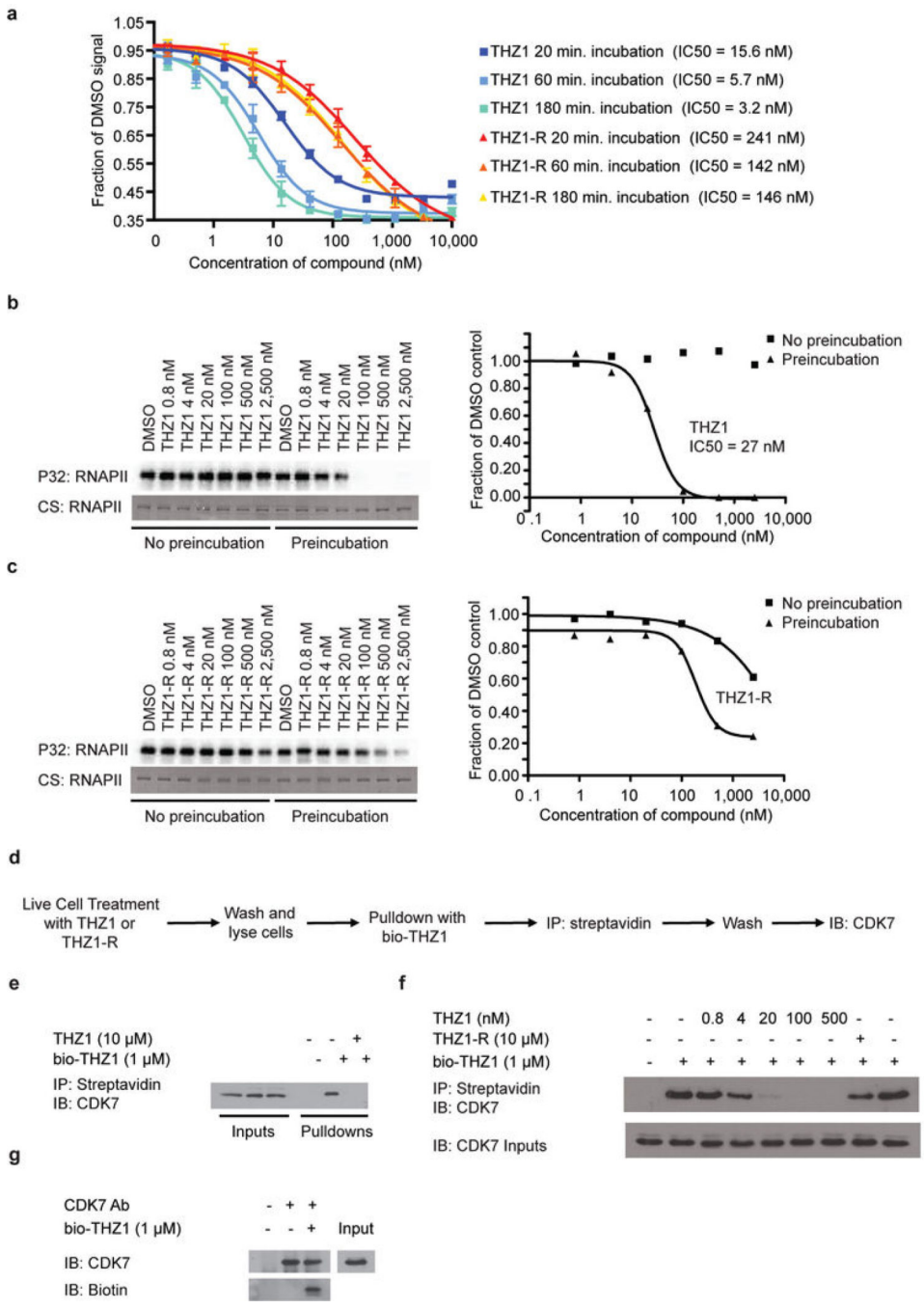
Cells were seeded in 384-well microplates at ~15% confluency in medium with 5% FBS and penicillin/streptavidin. Cells were treated with THZ1 or DMSO for 72 hrs and cell viability was determined using resazurin.

RNA Extraction and Synthetic RNA Spike-In

Total RNA and sample preparation was performed as previously described[30]. Briefly, following inhibitor treatment cell number was determined, total RNA was isolated, and ERCC RNA Spike-In Mix (Ambion, cat# 4456740) was added to total RNA relative to cell number.

Expanded protocols and synthetic chemistry schemes can be found in Supplementary Information.

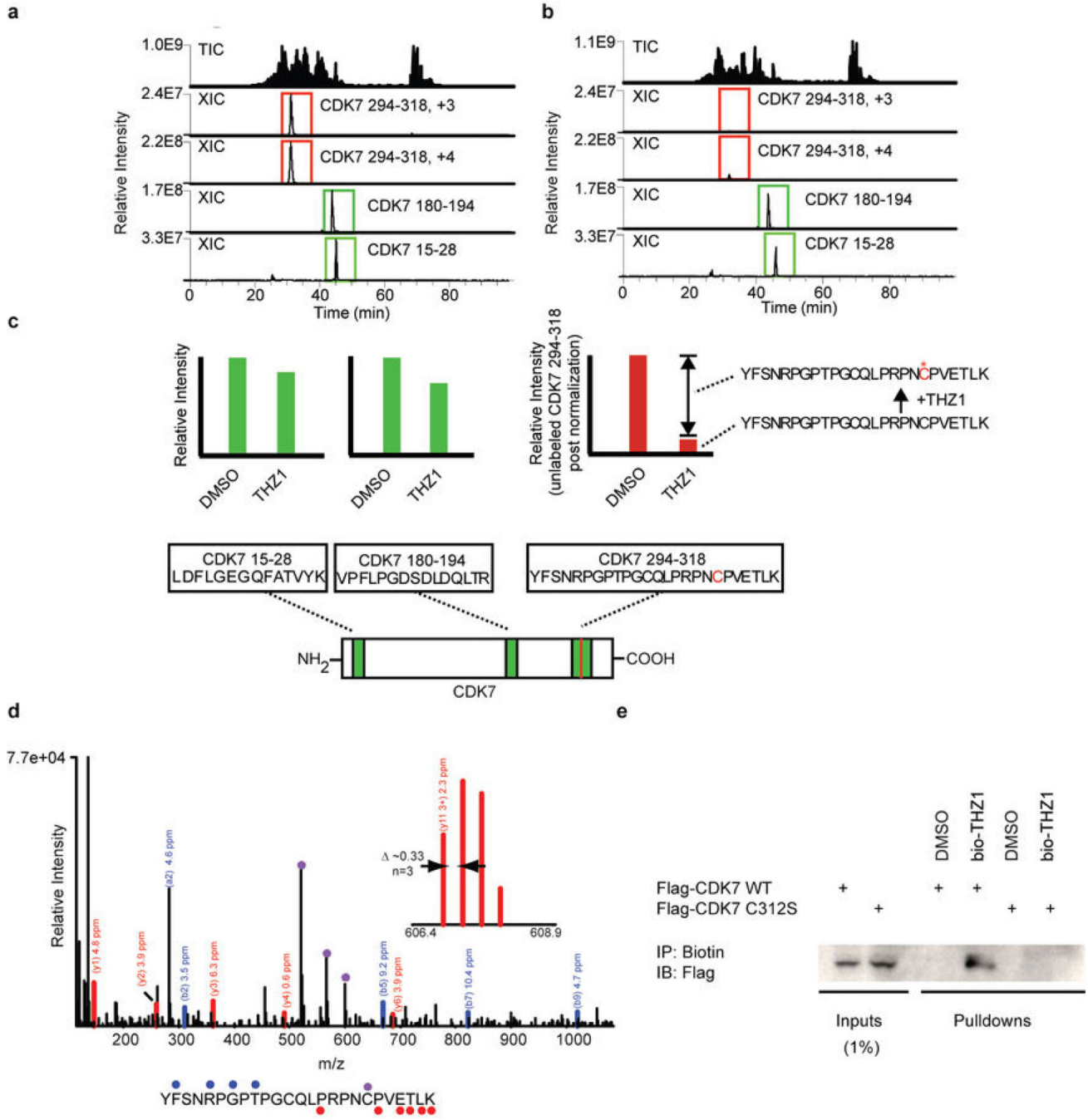
Extended Data



Extended Data Figure 1. THZ1 demonstrates time-dependent inhibition of CDK7 *in vitro* and covalent binding of intracellular CDK7

a, THZ1 but not THZ1-R shows time-dependent inhibition. LanthaScreen® Eu Kinase Binding assay was conducted at Life Technologies in a time-dependent manner (20, 60, and 180 min.) showing that THZ1 but not THZ1-R shows time-dependent inhibition of CDK7. **b and c,** Pre-incubation of THZ1 increases CDK7 inhibitory activity *in vitro*. Recombinant CAK complex was incubated with THZ1 **(b)** or THZ1-R **(c)** in dose response format with or

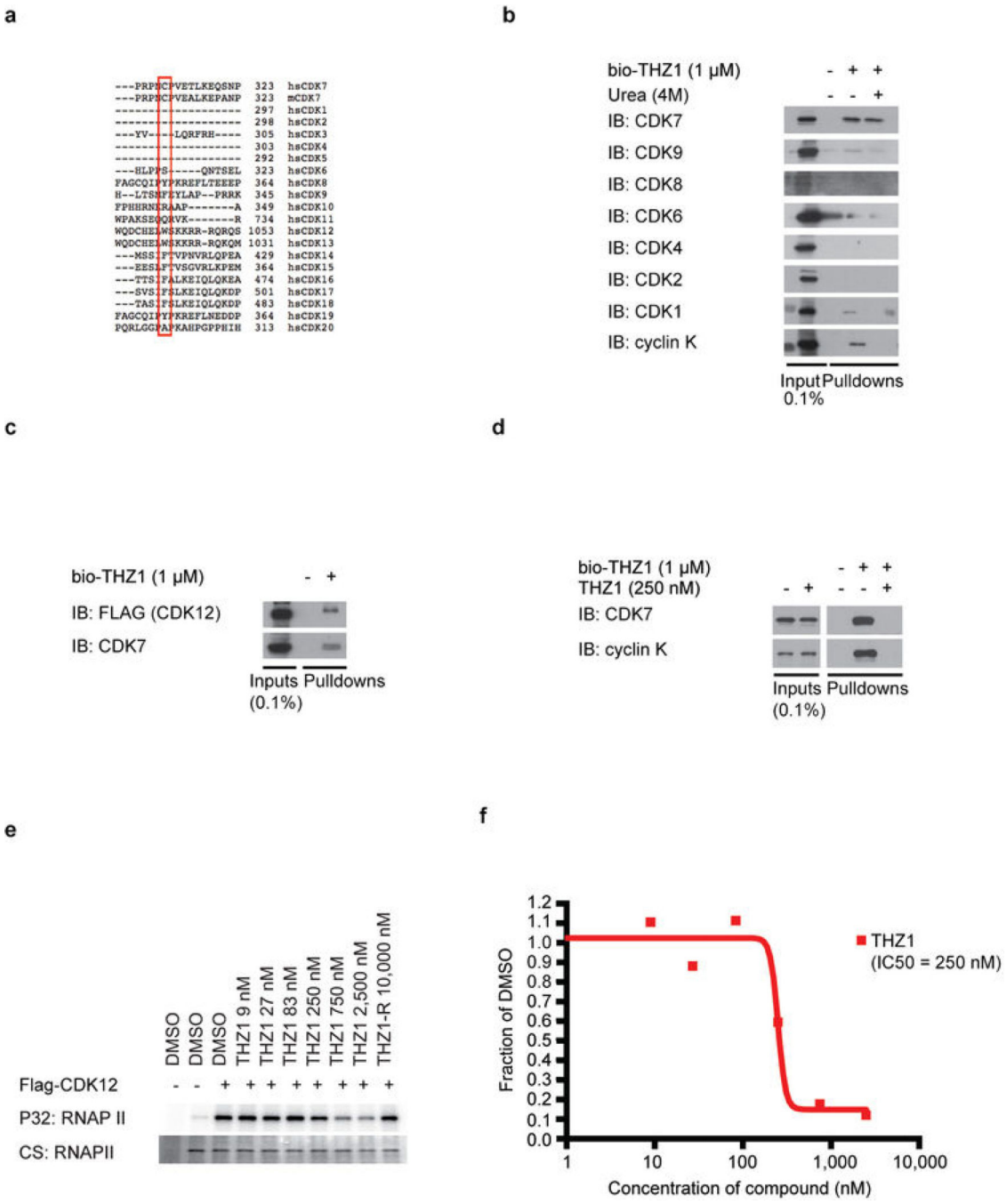
without pre-incubation prior to ATP (25 μM) addition. Kinase reaction was then allowed to proceed for 45 minutes at 30°C. **d,** Workflow of bio-THZ1 pull down competition experiment. **e,** bio-THZ1 pulls down CDK7 from cellular lysates. Loucy cellular lysates were incubated with bio-THZ1 (1 μM) with or without THZ1 (10 μM) and streptavidin-precipitated proteins were probed for CDK7. IB = immunoblot. **f,** Free intracellular THZ1 competes in a dose-dependent manner for bio-THZ1 binding to CDK7. Loucy cells were treated with increasing concentrations of THZ1 or with 10 μM THZ1-R for 4 hrs. Cellular lysates were incubated with bio-THZ1 and processed as indicated in **a**. **g,** bio-THZ1 labels CDK7 in lysates. Loucy cellular lysates were incubated with bio-THZ1 at 4°C for 12 hrs followed by immunoprecipitation of CDK7 at 4°C for 3 hrs. Precipitated proteins were washed and probed with streptavidin-HRP.



Extended Data Figure 2. THZ1 covalently binds CDK7 C312

a and **b,** Total ion chromatograms (TIC) and extracted ion chromatograms (XIC) for CDK7 peptides recorded during analysis of CAK complexes treated with DMSO (**a**) or THZ1 (**b**). **c,** Efficiency of labeling was estimated to be approximately 85% 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. **d,** Orbit rap HCD MS/MS spectrum of a quadruply charged CDK7 derived peptide (residues 294–318) labeled

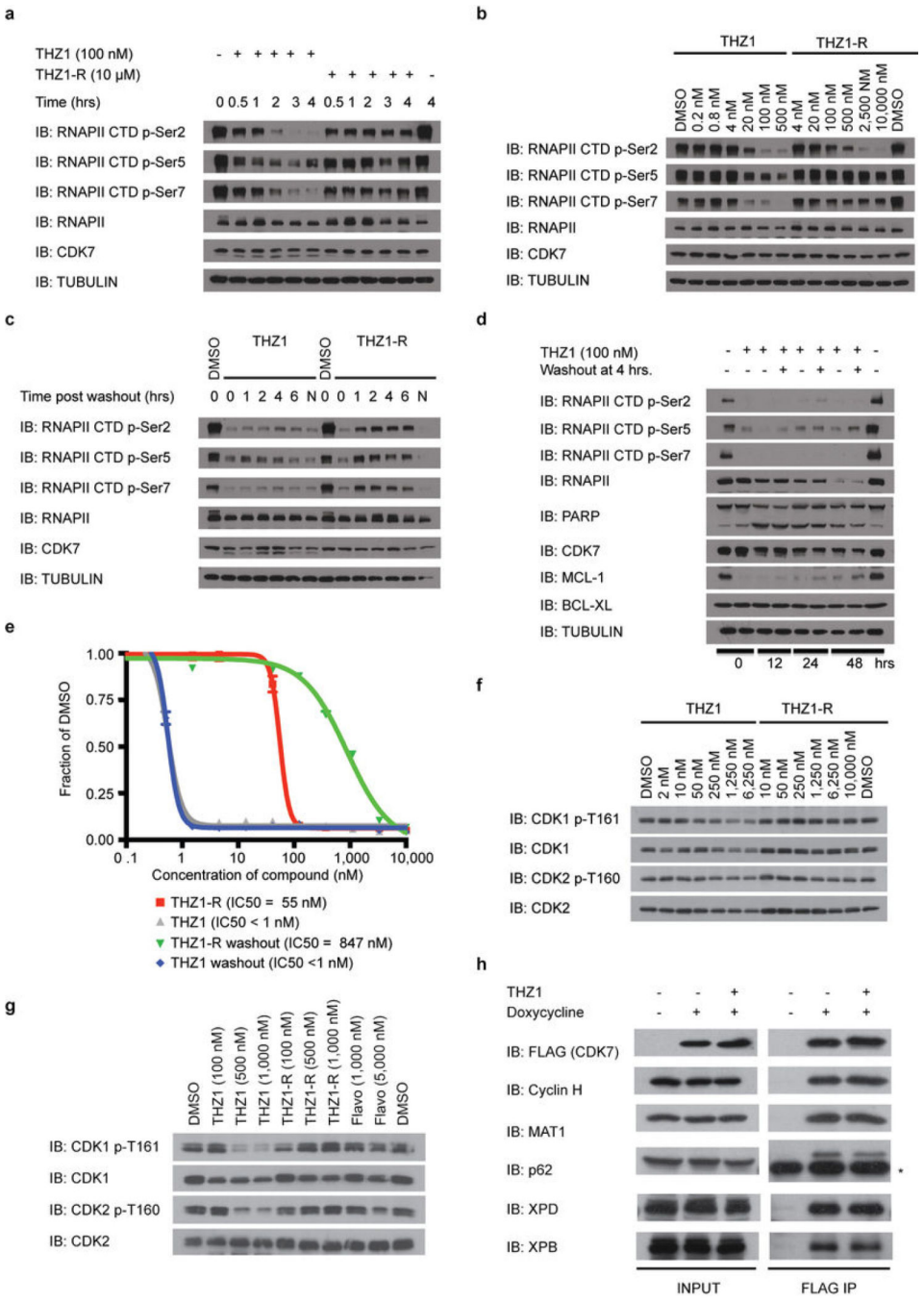
by THZ1 at C312. Fragment ions containing the peptide C-terminus (y-type) or N-terminus (b-type), along with the associated mass errors are shown in red and blue, respectively. Fragment ions marked by (*) contain the inhibitor and have the expected heavy isotope contribution from chlorine. The site of labeling was determined to be C312 (as opposed to C305) based on fragment ions observed in additional MS/MS spectra (for example $y11^{3+}$ observed with < 3 ppm mass error by fragmentation of the +6 charged precursor; see inset mass spectrum). **e,** C312S mutation eliminates THZ1 covalent binding. Cellular lysates from HCT116 cells expressing either FLAG-CDK7 WT or C312S were incubated with bio-THZ1 for 12 hrs at 4°C and then room temperature for 3 hrs to facilitate covalent binding. Precipitated proteins were then probed for the presence of FLAG-tagged CDK7.



Extended Data Figure 3. THZ1 inhibits CDK12 but at higher concentrations compared to CDK7

a, Protein sequence alignment of the C-terminal regions of all human (hs) CDKs and mouse (m) CDK7 using Uniprot default settings. Note that the canonical cell cycle CDKs 1,2,4 as well as 5 do not have C-terminal domains that extent to the equivalent position of CDK7 C312 and therefore do not display aligned sequence in this region. **b,** bio-THZ1 covalently pulls down CDK7 from cellular lysates. Jurkat cellular lysates were incubated with bio-THZ1 (1 μM) at 4°C for 12 hrs and 2 hrs at room temperature. Precipitated proteins were washed with or without urea (4M), here used as a denaturing agent, and probed for the

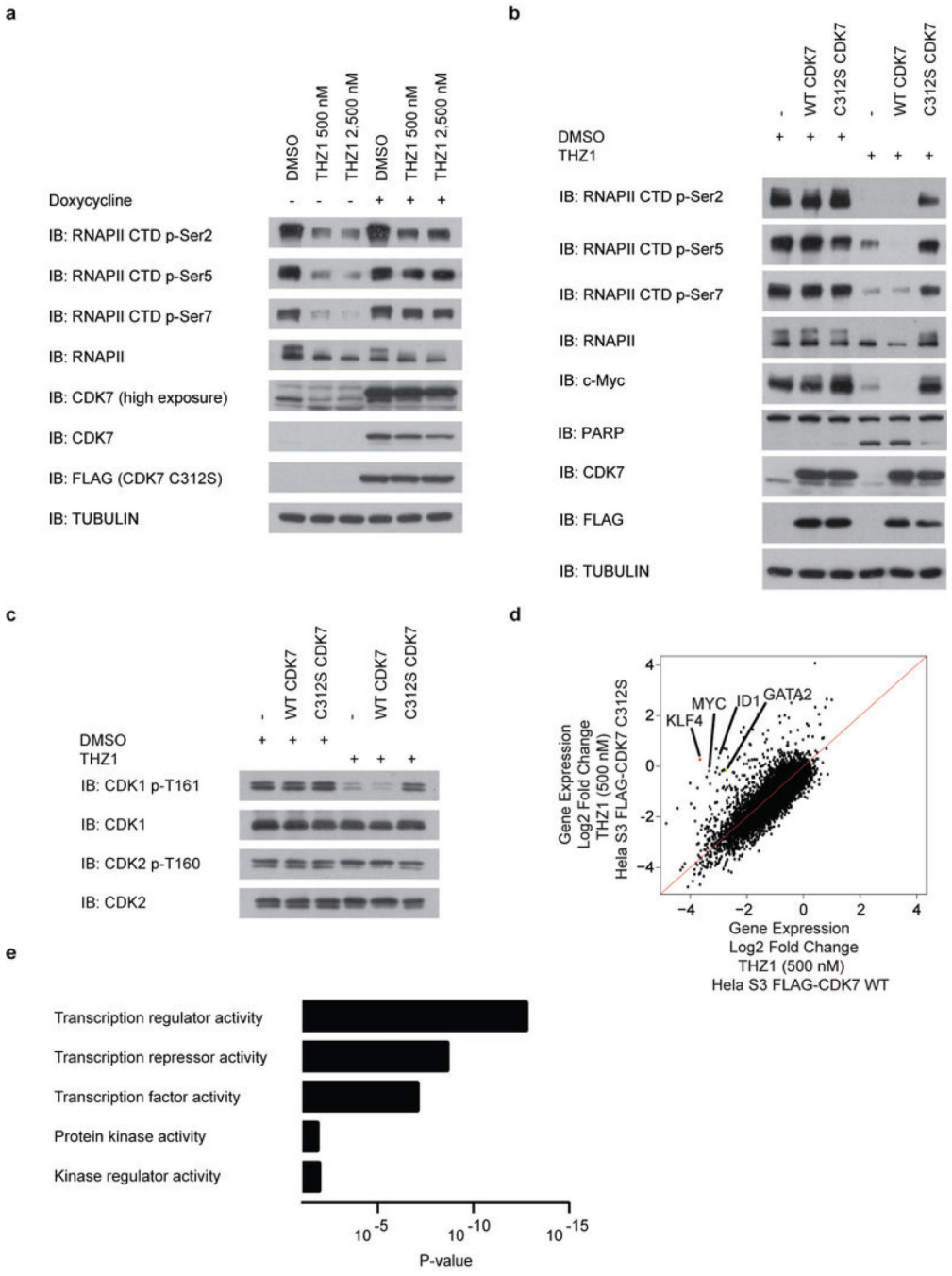
indicated CDKs. **c,** bio-THZ1 pulls down FLAGCDK12 from lysates. Lysates from 293A cells stably expressing FLAG-tagged WT CDK12 were incubated with bio-THZ1 (1 μM) at 4°C for 12 hrs and 2 hrs at rt. Immunoprecipitated proteins were probed with FLAG antibody to recognize CDK12 or with CDK7 antibody. **d,** bio-THZ1 pulls down cyclin K from cellular lysates. Jurkat cellular lysates were incubated with bio-THZ1 (1 μM) at 4°C for 12 hrs and 2 hrs at rt. Precipitated proteins were probed for the indicated proteins. **e,** THZ1 inhibits CDK12 in an *in vitro* kinase assay. 293A cells stably expressing FLAG-tagged WT CDK12 were treated with THZ1 or THZ1-R for 4 hrs. Exogenous CDK12 was immunoprecipitated from cellular lysates using FLAG antibody. Precipitated proteins were washed and subjected to *in vitro* kinase assays at 30°C for 30 minutes using the large subunit of RNAPII (RPB1) as substrate and 25 μM ATP. CS = coomassie stain. **f,** Quantitation of *in vitro* kinase assay conducted in (d).



Extended Data Figure 4. THZ1 irreversibly inhibits RNAPII CTD and CAK phosphorylation

a, THZ1 exhibits time-dependent inactivation of intracellular CDK7. Loucy cells were treated with THZ1 or THZ1-R for 0 to 4 hrs. At each time point cells were harvested, lysed, and the cellular lysates were probed with antibodies against the specified proteins. **b,** THZ1 inhibits RNAPII CTD phosphorylation. Loucy cells were treated with THZ1 or THZ1-R for 4 hrs. Cellular lysates were then probed with antibodies recognizing the Ser-2, Ser-5, and Ser-7 CTD RNAPII phosphoepitopes. **c,** Loucy cells were treated with THZ1 or THZ1-R for 4 hrs followed by washout of inhibitor-containing medium. Cells were allowed to grow in

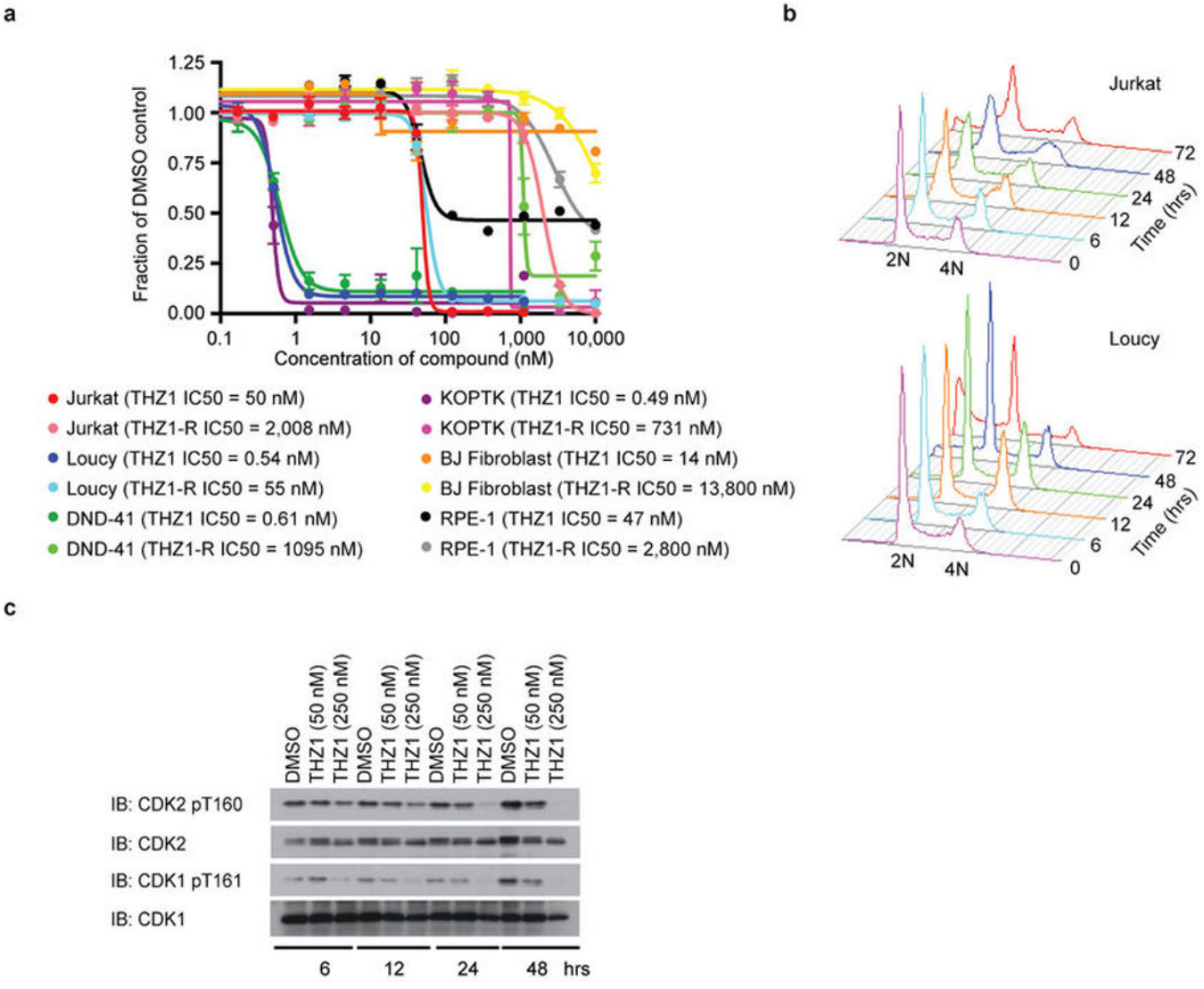
medium without inhibitor for 0 to 6 hrs. At each time point cells were lysed and the cellular lysates were probed with antibodies against the specified proteins. 'N' signifies cells where medium was never washed out. **d,** Apoptotic signaling is maintained despite washout of THZ1. Loucy cells were treated with THZ1 or THZ1-R for 4 hrs followed by washout of inhibitor-containing medium, at which point cells were allowed to grow in medium with or without inhibitor for 0 to 48 hrs. At each time point cells were lysed and the cellular lysates were probed with antibodies against the specified proteins. **e,** Anti proliferative effects of THZ1 are impervious to inhibitor washout. Loucy cells were treated with THZ1 or THZ1-R in dose response format for 72 hrs. Anti proliferative effects were determined using cell titer glo analysis. **f,** THZ1 reduces the T-loop phosphorylation status of CDK1 and CDK2 in Jurkat cells over a 3 hour exposure. Asynchronous cells were treated with increasing concentrations of THZ1 or THZ1-R for 3 hrs. Cellular lysates were then probed with antibodies against the indicated proteins or phosphoproteins. **g,** THZ1, but not THZ1-R, completely inhibits T-loop phosphorylation of CDK1 and CDK2 following treatment over one cell cycle. Loucy cells were treated with THZ1, THZ1-R, Flavopiridol, or DMSO vehicle at the indicated concentrations for 24 and 14 hrs, respectively (roughly one cell cycle). Cell lysates were harvested and probed with antibodies against the specified proteins or phosphoproteins. **h**, Hela S3 cells stably expressing FLAG-WT CDK7 were treated with THZ1 (1 μM) or DMSO vehicle for 5 hrs with and without the presence of doxycycline. Proteins were immunoprecipitated using FLAG antibody. Precipitated proteins were probed using the indicated antibodies. * indicates heavy-chain from IgG antibody.



Extended Data Figure 5. Mutation of CDK7 Cys-312 to serine rescues Ser-5/7 and partially Ser2RNAPII CTD phosphorylation

a, Expression of C312S rescues Ser-5/7 and partially rescues Ser-2 RNAPII CTD phosphorylation. Hela S3 cells stably carrying a doxycycline-inducible FLAG-C312S CDK7 construct were treated with THZ1 or DMSO for 5 hrs with and without the presence of doxycycline. Cellular lysates were then probed for the indicated proteins. **b,** Phenotypic rescue is specific to C312S mutation as rescue is not achieved with overexpression of FLAG-WT CDK7. Hela S3 cells stably carrying doxycycline-inducible FLAG-WT and

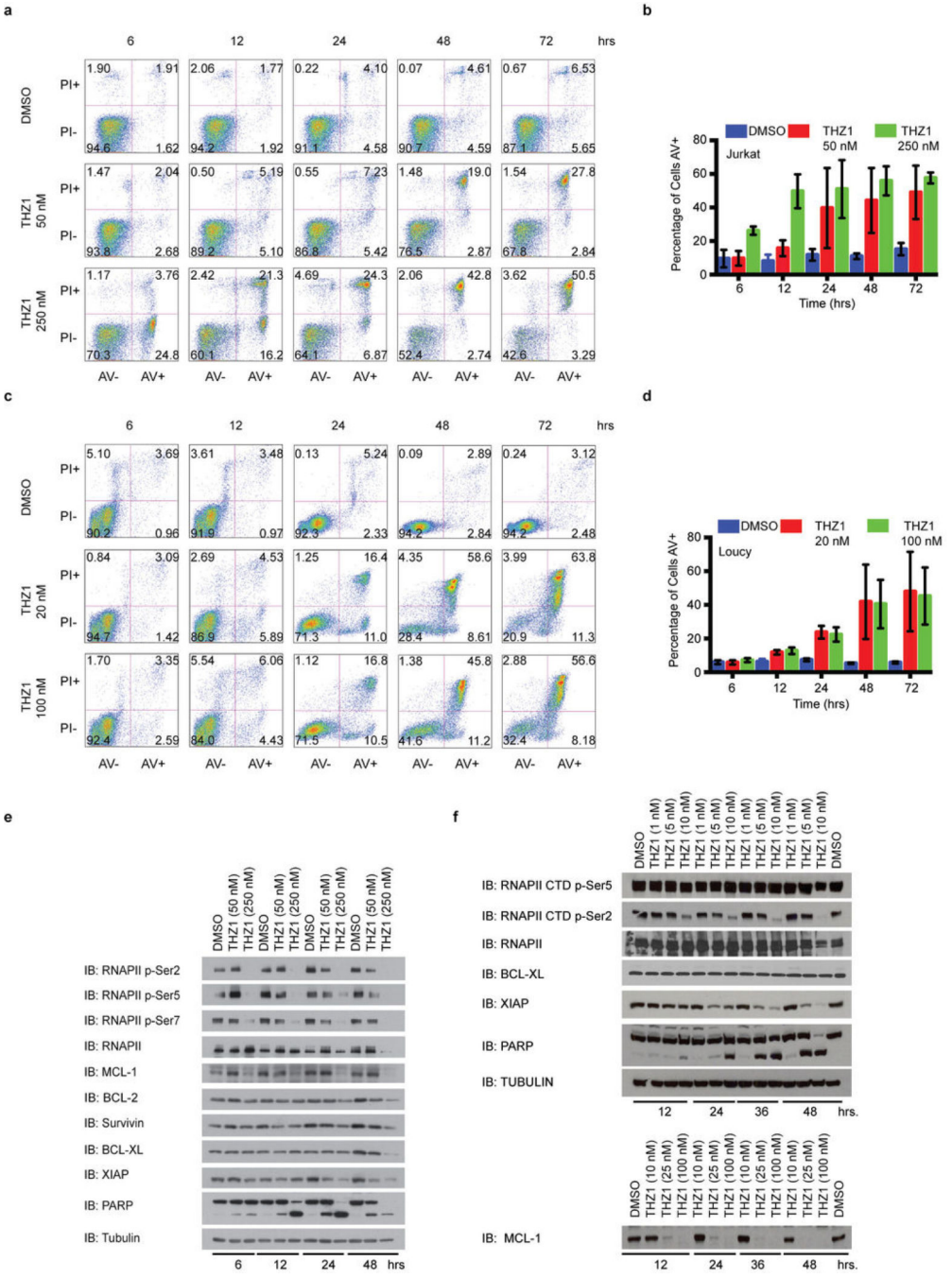
C312S CDK7 constructs (or empty vector) were treated with THZ1 or DMSO for 5 hrs in the presence of doxycycline. **c,** Expression of C312S largely restores CDK1/2 T-loop phosphorylation. Hela S3 cells stably carrying a doxycycline-inducible FLAG-C312S CDK7 construct were treated with THZ1 or DMSO for 5 hrs with and without the presence of doxycycline. Cellular lysates were then probed for the indicated proteins or phosphoproteins. **d,** Overexpression of FLAG-CDK7 C312 rescues the expression of a subset of transcripts in Hela S3 cells. Log2 fold change in gene expression in Hela S3 cells expressing FLAG-CDK7 WT (x axis) and FLAG-CDK7 C312S (y axis) following a 4 hr treatment with 500 nM THZ1. **e,** Gene ontology molecular function analysis of transcripts increased by 1 log2 order or more following expression of FLAG-CDK7 C312S compared to FLAG-CDK7 WT in the presence of 500 nM THZ1.



Extended Data Figure 6. THZ1 potently disrupts T-ALL proliferation

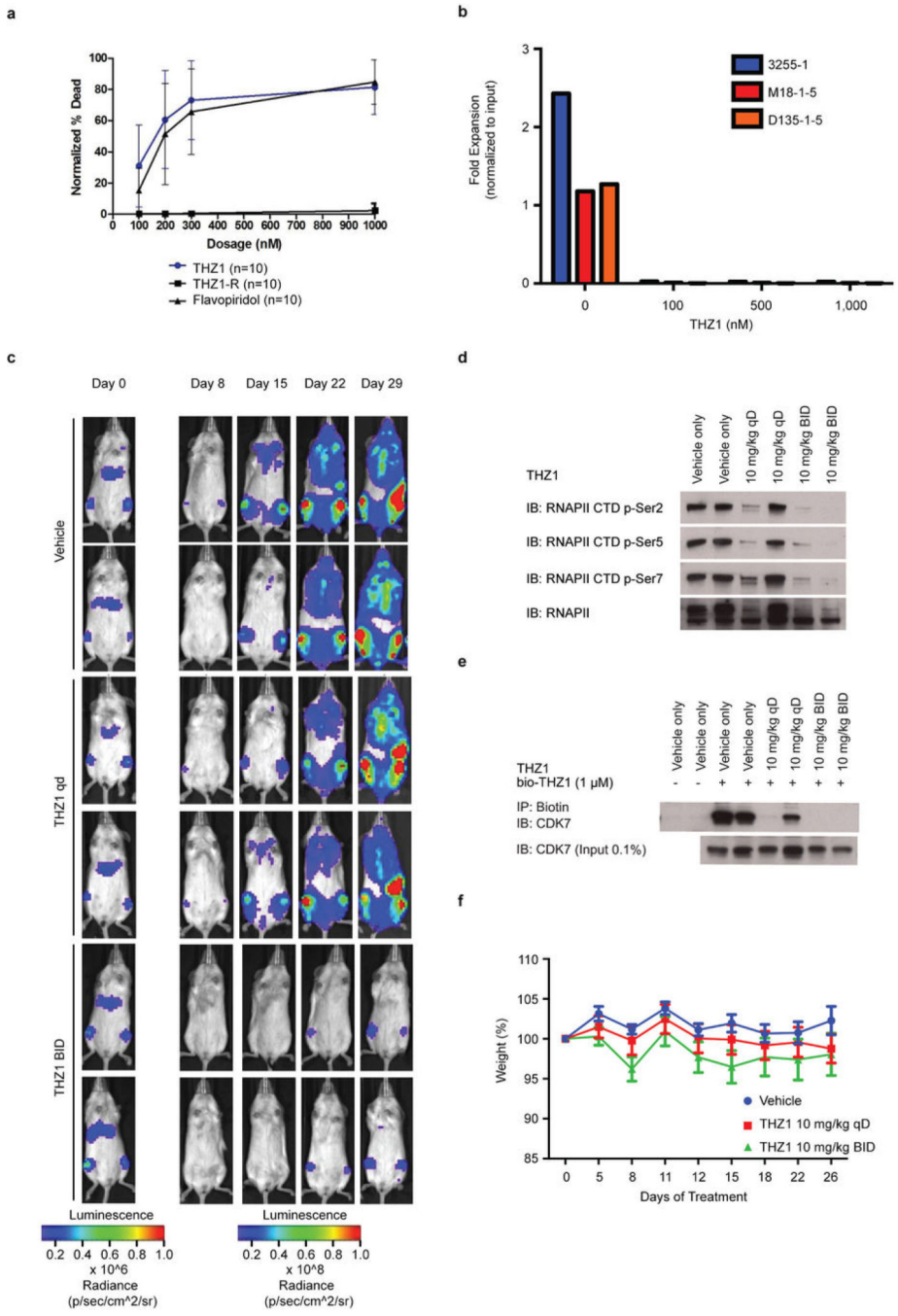
a, THZ1, but not THZ1-R, exhibits strong antiproliferative effect against T-ALL cell lines. Cells were treated with THZ1, THZ1-R, or DMSO vehicle for 72 hrs and assessed for

antiproliferative effect by Cell Titer Glo analysis. Error bars are +/− SD. **b,** THZ1 causes cell cycle arrest. Jurkat (top) and Loucy (bottom) T-ALL cells were treated with THZ1 for the indicated time periods. Cell cycle progression was assessed using FACS cell cycle analysis. 2N = G1, 4N = G2. **c,** Treatment with THZ1 decreases CDK1/2 T loop phosphorylation. Jurkat cells were incubated with THZ1 for the indicated duration of time and lysates were probed for the specified proteins.



Extended Data Figure 7. Treatment with THZ1 induces apoptosis in T-ALL cells

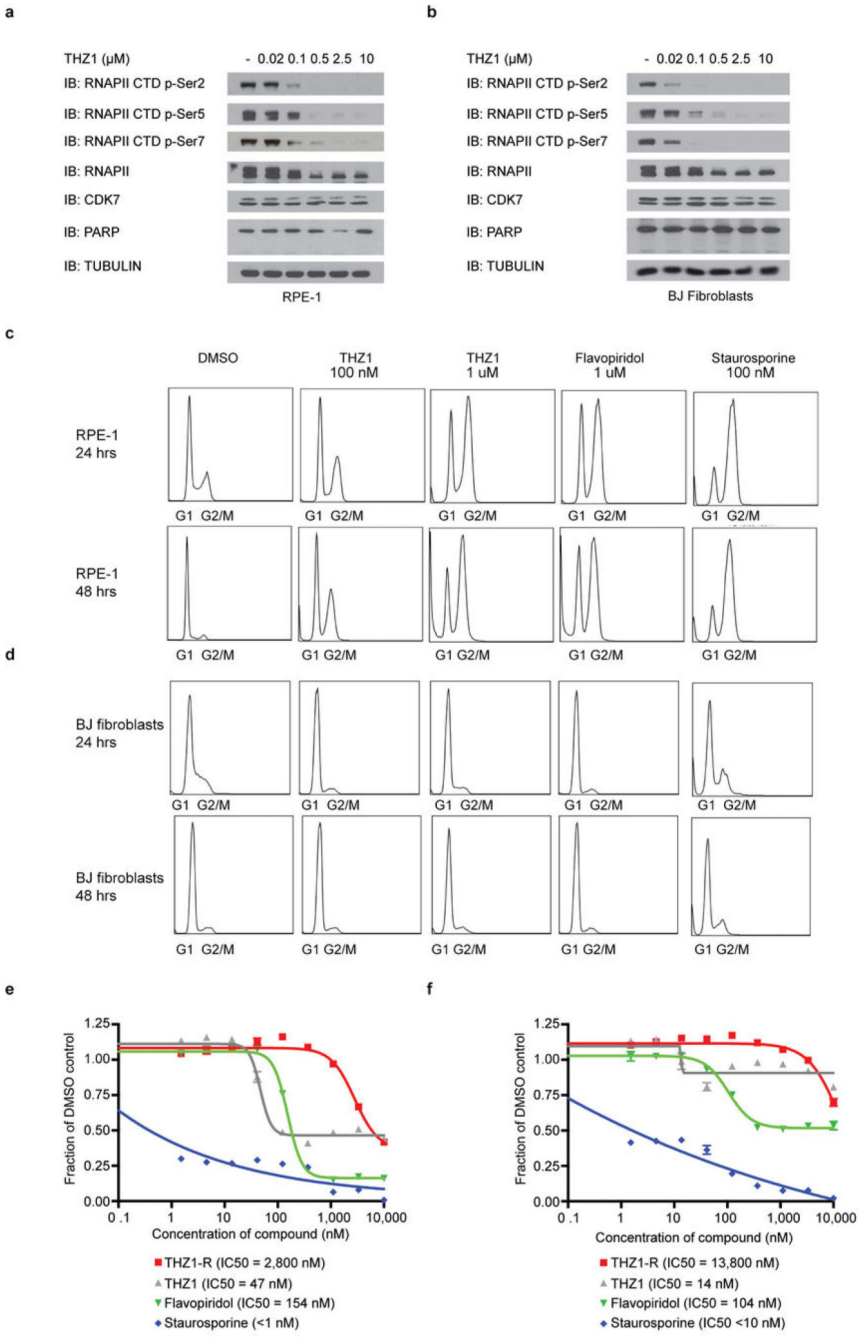
a, Representative Annexin V and propidium iodide stainings for Jurkat cells incubated with THZ1 for the indicated amount of time and harvested to determine the percentage of apoptotic and/ or dead cells by Annexin V and propidium iodide staining, respectively. The percentage of cells in each cell population is shown in the four quadrants. **b,** Treatment with THZ1 induces apoptosis. Quantitation of Annexin V and propidium iodide staining data from **a.** Experiments were performed in biological triplicates and error bars are +/− SD. **c,** Representative Annexin V and propidium iodide stainings for Loucy cells incubated with THZ1 for the indicated amount of time and harvested to determine the percentage of apoptotic and/ or dead cells by Annexin V and propidium iodide staining, respectively. The percentage of cells in each cell population is shown in the four quadrants. **d,** Treatment with THZ1 induces apoptosis. Quantitation of Annexin V and propidium iodide staining data from **c.** Experiments were performed in biological triplicates and error bars are +/− SD. **e** and **f,** Sustained treatment with THZ1 induces apoptosis coincident with loss of RNAPII CTD phosphorylation and reduction in antiapoptotic proteins. Jurkat (**e**) and Loucy (**f**) cells were incubated with THZ1 for the indicated duration of time and lysates were probed for the specified proteins. Apoptosis was monitored by PARP cleavage.



Extended Data Figure 8. THZ1 demonstrates potent killing of primary chronic lymphocytic leukemia (CLL) cells and anti-proliferative activity against primary TALL cells and *in vivo* against a human T-ALL xenograft

a, CLL cells were cultured *in vitro* for 24 hrs in the presence of multiple doses of the specified compound. Results shown are mean normalized % death based on Annexin V / PI single and double positive cells (+/− SD) normalized to baseline death in the DMSO control wells. 10 patient samples were exposed to THZ1, THZ1-R and Flavopiridol (THZ1 vs. THZ1-R p = 1.5E-38; THZ1 vs. Flavopiridol p = 0.05). P-values were generated using an analysis of variance model. **b**, Patient-derived xenografts (patient ID# 3255-1, M18-1-5,

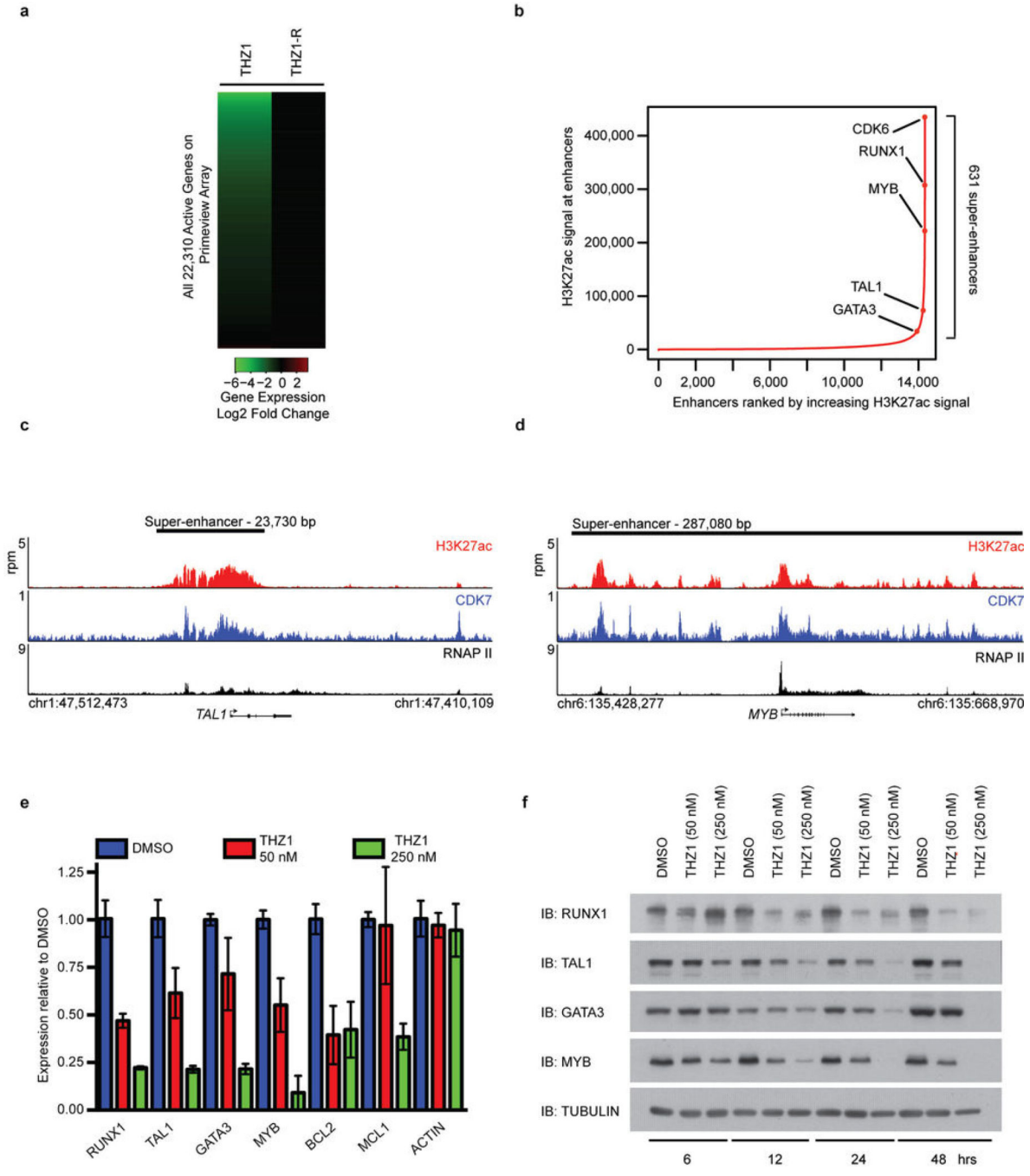
D135-1-5; n=3) were treated with THZ1 for 3 hrs followed by compound washout. An aliquot of input cells was then counted by flow cytometry using a known quantity of flow cytometry calibration beads (data not shown; Molecular Probes). The remaining cells were plated onto MS5-DL1feeder cells in the presence of serum-free media (supplemented with 0.75 μM SR1, 10 ng/ml IL7, 10 ng/ml IL2). 72 hrs later, cultures were harvested by vigorous pipetting with Trypsin, filtered through nylon mesh to deplete feeders, and counted by flow cytometry using a known quantity of flow cytometry calibration beads and with gating to discriminate between T-ALL cells and carryover feeders. The final cell number was normalized to the input cell number to calculate fold expansion. This experiment was performed once per patient- derived sample. **c,** Bioluminescent images of two representative mice treated with either vehicle control, 10 mg/kg THZ1 qD (once daily), or 10 mg/kg/day THZ1 BID (twice daily) for indicated number of days. **d,** Spleen tissue from mice treated with THZ1 show decreased RNAPII CTD phosphorylation. Mice were treated with THZ1 10 mg/kg qD or BID or vehicle control. The animals were sacrificed and spleen tissues were isolated. Lysates prepared from homogenized spleen tissue were probed for RNAPII CTD phosphoepitopes. **e,** THZ1 binds directly to CDK7 in mouse tissues. Mice were treated with THZ1 10 mg/kg qD or BID or vehicle control. The animals were sacrificed and spleen tissues were isolated. Lysates prepared from homogenized spleen tissue were incubated with bio-THZ1 for 12 hrs at 4 °C and 2 hrs at rt to induce covalent bond formation. Proteins pulled down were then probed for the presence of CDK7. **f,** Bodyweights of mice treated with either vehicle control, 10 mg/kg THZ1 qD (once daily), or 10 mg/kg/day THZ1 BID (twice daily) over the duration of the drug treatment.



Extended Data Figure 9. THZ1 inhibits RNAPII CTD phosphorylation and causes cell cycle arrest in non-transformed cell lines

a and **b,** THZ1 inhibits RNAPII CTD phosphorylation. RPE-1 **(a)**, and BJ fibroblasts **(b)** were treated with THZ1 or THZ1-R for 4 hrs. Cellular lysates were then probed with antibodies against the indicated proteins. **c** and **d**, THZ1 causes cell cycle arrest in non-transformed cells. RPE-1 **(c)** and BJ fibroblasts **(d)** cells were treated with THZ1, Flavopiridol, Staurosporine, or DMSO vehicle for the indicated time periods. Cell cycle progression was analyzed following permeabilization and staining with propidium iodide. **e**

and **f,** THZ1 inhibits proliferation of non-transformed cell lines. RPE-1 **(e)** and BJ fibroblasts **(f)** cells were treated with THZ1, THZ1-R, Flavopiridol, or Staurosporine for 72 hrs and antiproliferative effect was determined by Cell Titer Glo. Error bars are +/− SD.



Extended Data Figure 10. High dose THZ1 reduces global steady-state mRNA levels, but low dose THZ1 preferentially downregulates components of the TAL1/RUNX1/GATA3 transcriptional circuit

a, THZ1, but not THZ1-R, causes global downregulation of steady-state mRNA levels. Jurkat cells were treated with THZ1 (250 nM) or THZ1-R (250 nM) for 4 hrs. Total RNA

was isolated and ERCC spike-in controls were added relative to cell number and analyzed using Affymetrix Prime view microarrays. Heat maps displaying the Log 2 fold change in gene expression vs. DMSO for 22,310 genes expressed in DMSO conditions at 6h in THZ1 or THZ1-R. **b,** Total H3K27Ac ChIP-seq signal (length * density) in enhancer regions for all stitched enhancers in Jurkat. Enhancers are ranked by increasing H3K27Ac ChIP-seq signal. **c** and **d,** Gene tracks of H3K27Ac (top), CDK7 (middle), and RNAPII (bottom) ChIP-seq occupancy at the TSS, gene body, and enhancer regions of *TAL1* (**c**) and *MYB* (**d**). **e,** THZ1 downregulates mRNA transcripts of the TAL1/RUNX1/GATA3 transcriptional circuitry. RT-qPCR expression analysis in Jurkat cells of transcripts identified as downregulated following THZ1 treatment relative to DMSO. All experiments shown were performed in biological triplicate. Each individual biological sample was qPCR-amplified in technical triplicate. Error bars are +/− SD. Taqman universal expression probes and normalized to ACTB. **f,** THZ1 treatment reduces the protein levels of TAL1/RUNX1/GATA3 transcriptional circuitry. Jurkat cells treated with THZ1 for the indicated time points were probed for the specified proteins.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

We thank members of the Gray and Young laboratories for helpful discussions. D. Orlando, L. Lawton, and L. Anders for advice. C. Thoreen and D. Sabatini as well as S. Cheng and G. Morin for reagents. We thank K. Jones and N. Kohl in the Lurie Family Imaging Center at DFCI for performing mouse studies and K. Jones and C. Christensen for prepping mouse tissues. We thank S. Riddle at Life Technologies for performing Lanthascreen kinase assays. This work was supported by the National Institutes of Health (R01 CA130876-04 and U54 HG006097-02 N.S.G.; CA178860-01 and P01 NS047572-10 J.A.M.), and the American Cancer Society Postdoctoral Fellowship 120272-PF-11-042-01-DMC (P.B.R.).

References

1. O'Neil J, Look AT. Mechanisms of transcription factor deregulation in lymphoid cell transformation. Oncogene. 2007; 26:6838–6849. [PubMed: 17934490]
2. Berg T. Inhibition of transcription factors with small organic molecules. Current opinion in chemical biology. 2008; 12:464–471. [PubMed: 18706517]
3. Lu Q, et al. Perspectives on the discovery of small-molecule modulators for epigenetic processes. Journal of biomolecular screening. 2012; 17:555–571. [PubMed: 22392809]
4. Sanso M, Fisher RP. Pause, Play, Repeat: CDKs Push RNAP II's Buttons. Transcription. 2013; 4
5. Patricelli MP, et al. Functional interrogation of the kinome using nucleotide acyl phosphates. Biochemistry. 2007; 46:350–358. [PubMed: 17209545]
6. Akhtar MS, et al. TFIIH kinase places bivalent marks on the carboxy-terminal domain of RNA polymerase II. Molecular cell. 2009; 34:387–393. [PubMed: 19450536]
7. Drapkin R, Le Roy G, Cho H, Akoulitchev S, Reinberg D. Human cyclin-dependent kinase-activating kinase exists in three distinct complexes. Proceedings of the National Academy of Sciences of the United States of America. 1996; 93:6488–6493. [PubMed: 8692842]
8. Glover-Cutter K, et al. TFIIH-associated Cdk7 kinase functions in phosphorylation of C-terminal domain Ser7 residues, promoter-proximal pausing, and termination by RNA polymerase II. Molecular and cellular biology. 2009; 29:5455–5464. [PubMed: 19667075]
9. Larochelle S, et al. Cyclin-dependent kinase control of the initiation-to-elongation switch of RNA polymerase II. Nature structural & molecular biology. 2012; 19:1108–1115.

10. Fisher RP, Morgan DO. A novel cyclin associates with MO15/CDK7 to form the CDK-activating kinase. Cell. 1994; 78:713–724. [PubMed: 8069918]
11. Larochelle S, et al. Requirements for Cdk7 in the assembly of Cdk1/cyclin B and activation of Cdk2 revealed by chemical genetics in human cells. Molecular cell. 2007; 25:839–850. [PubMed: 17386261]
12. Makela TP, et al. A cyclin associated with the CDK-activating kinase MO15. Nature. 1994; 371:254–257. [PubMed: 8078587]
13. Schachter MM, et al. A Cdk7-Cdk4 T-loop phosphorylation cascade promotes G1 progression. Molecular cell. 2013; 50:250–260. [PubMed: 23622515]
14. Solomon MJ, Lee T, Kirschner MW. Role of phosphorylation in p34cdc2 activation: identification of an activating kinase. Molecular biology of the cell. 1992; 3:13–27. [PubMed: 1532335]
15. Bartkowiak B, et al. CDK12 is a transcription elongation-associated CTD kinase, the metazoan ortholog of yeast Ctk1. Genes & development. 2010; 24:2303–2316. [PubMed: 20952539]
16. Garnett MJ, et al. Systematic identification of genomic markers of drug sensitivity in cancer cells. Nature. 2012; 483:570–575. [PubMed: 22460902]
17. Huang da W, Sherman BT, Lempicki RA. Bioinformatics enrichment tools: paths toward the comprehensive functional analysis of large gene lists. Nucleic acids research. 2009; 37:1–13. [PubMed: 19033363]
18. Feaver WJ, Svejstrup JQ, Henry NL, Kornberg RD. Relationship of CDK-activating kinase and RNA polymerase II CTD kinase TFIIH/TFIIK. Cell. 1994; 79:1103–1109. [PubMed: 8001136]
19. Serizawa H, et al. Association of Cdk-activating kinase subunits with transcription factor TFIIH. Nature. 1995; 374:280–282. [PubMed: 7885450]
20. Shiekhattar R, et al. Cdk-activating kinase complex is a component of human transcription factor TFIIH. Nature. 1995; 374:283–287. [PubMed: 7533895]
21. Sanda T, et al. Core transcriptional regulatory circuit controlled by the TAL1 complex in human T cell acute lymphoblastic leukemia. Cancer cell. 2012; 22:209–221. [PubMed: 22897851]
22. Rahl PB, et al. c-Myc regulates transcriptional pause release. Cell. 2010; 141:432–445. [PubMed: 20434984]
23. Watanabe Y, et al. Modulation of TFIIH-associated kinase activity by complex formation and its relationship with CTD phosphorylation of RNA polymerase II. Genes to cells : devoted to molecular & cellular mechanisms. 2000; 5:407–423. [PubMed: 10886368]
24. Yamada T, et al. P-TEFb-mediated phosphorylation of hSpt5 C-terminal repeats is critical for processive transcription elongation. Molecular cell. 2006; 21:227–237. [PubMed: 16427012]
25. Chapuy B, et al. Discovery and characterization of super-enhancer-associated dependencies in diffuse large B cell lymphoma. Cancer cell. 2013; 24:777–790. [PubMed: 24332044]
26. Hnisz D, et al. Super-enhancers in the control of cell identity and disease. Cell. 2013; 155:934–947. [PubMed: 24119843]
27. Loven J, et al. Selective inhibition of tumor oncogenes by disruption of super-enhancers. Cell. 2013; 153:320–334. [PubMed: 23582323]
28. Shi J, et al. Role of SWI/SNF in acute leukemia maintenance and enhancer-mediated Myc regulation. Genes & development. 2013; 27:2648–2662. [PubMed: 24285714]
29. Nottingham WT, et al. Runx1-mediated hematopoietic stem-cell emergence is controlled by a Gata/Ets/SCL-regulated enhancer. Blood. 2007; 110:4188–4197. [PubMed: 17823307]
30. Loven J, et al. Revisiting global gene expression analysis. Cell. 2012; 151:476–482. [PubMed: 23101621]

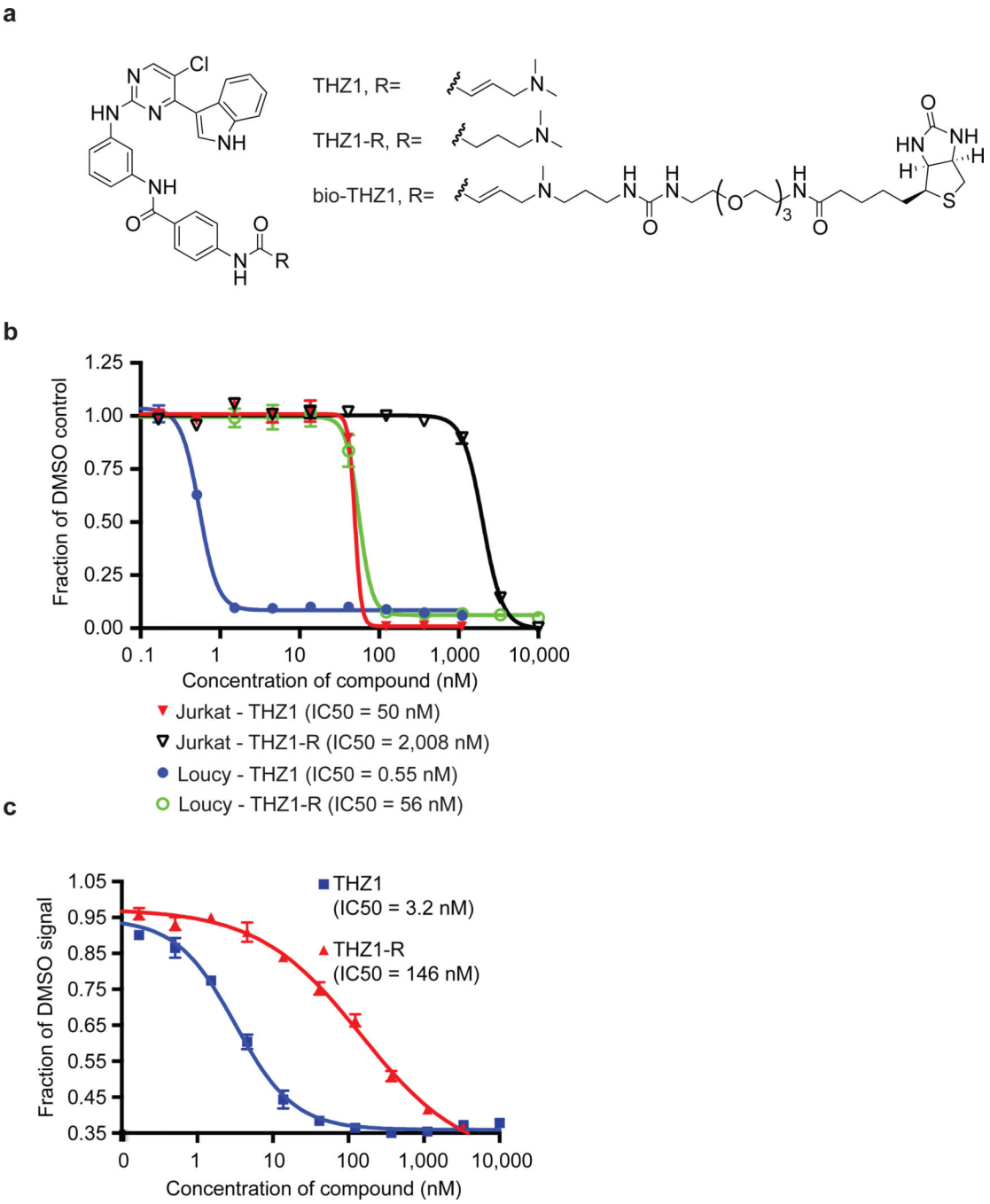


Figure 1. Cell-based screening and kinome profiling identifies phenylamino-pyrimidines as a potential CDK7 scaffold

a, Compound structures of THZ1, THZ1-R, and bio-THZ1. **b,** THZ1 potently inhibits proliferation of Jurkat and Loucy T-ALL cell lines. Cell lines were treated with THZ1 or THZ1-R for 72 hrs. Experiments were performed in biological triplicates and error bars are +/− SD. **c,** THZ1 and THZ1-R have different binding affinities for CDK7. LanthaScreen® Eu Kinase Binding assay was conducted at Life Technologies in a time-dependent manner. Here K_D values are shown following 180-minute incubation with compounds. Experiments were performed in biological triplicates and error bars are +/− SD.

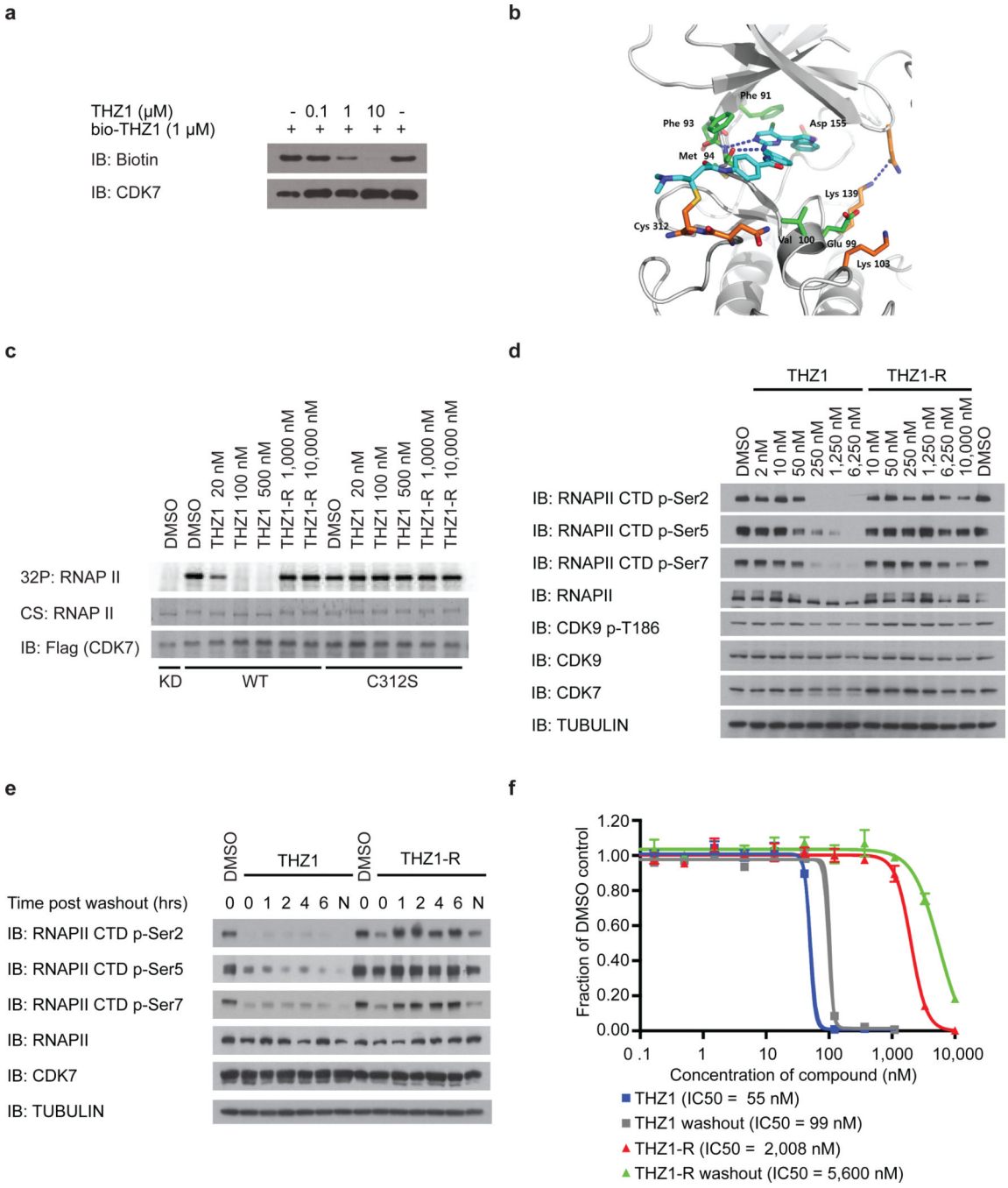


Figure 2. THZ1 irreversibly inhibits RNAPII CTD phosphorylation by covalently targeting a unique cysteine located outside the kinase domain of CDK7

a, bio-THZ1 binds irreversibly to CDK7. Recombinant CAK complex was incubated with bio-THZ1 +/− THZ1 at 37°C for 4 hrs and biotinylated proteins were resolved by SDS-page. **b,** Docking model of THZ1 in the ATP-binding pocket of CDK7 (PDB code 1UA2). CDK7 depicted with grey ribbons and THZ1 in turquois. Key residues are indicated. C312 has been modeled into the crystal structure. **c,** Mutation of C312 to serine (C312S) rescues wild-type kinase activity in the presence of THZ1. HCT116 cells stably expressing FLAG-tagged

CDK7 proteins were treated with THZ1 or THZ1-R for 4 hrs. Exogenous CDK7 proteins were immunoprecipitated with FLAG antibody and subjected to *in vitro* kinase assays. CS = coomassie stain. **d,** THZ1 inhibits RNAPII CTD phosphorylation. Jurkat cells were treated with THZ1 or THZ1-R for 4 hrs. and proteins of interest resolved by SDS page. **e,** THZ1, but not THZ1-R, shows irreversible inactivation of CDK7. Jurkat cells were treated with THZ1 or THZ1-R for 4 hrs followed by washout of inhibitor-containing medium. Cells were then allowed to grow in medium without inhibitor for 0 to 6 hrs. 'N' signifies cells where medium was never washed out (*ie* – 10 hr incubation with compounds). **f,** Anti proliferative effects of THZ1 are impervious to inhibitor washout. Jurkat cells were treated with THZ1 or THZ1-R in dose response format for 72 hrs. Experiments were performed in biological triplicates and error bars are +/− SD.

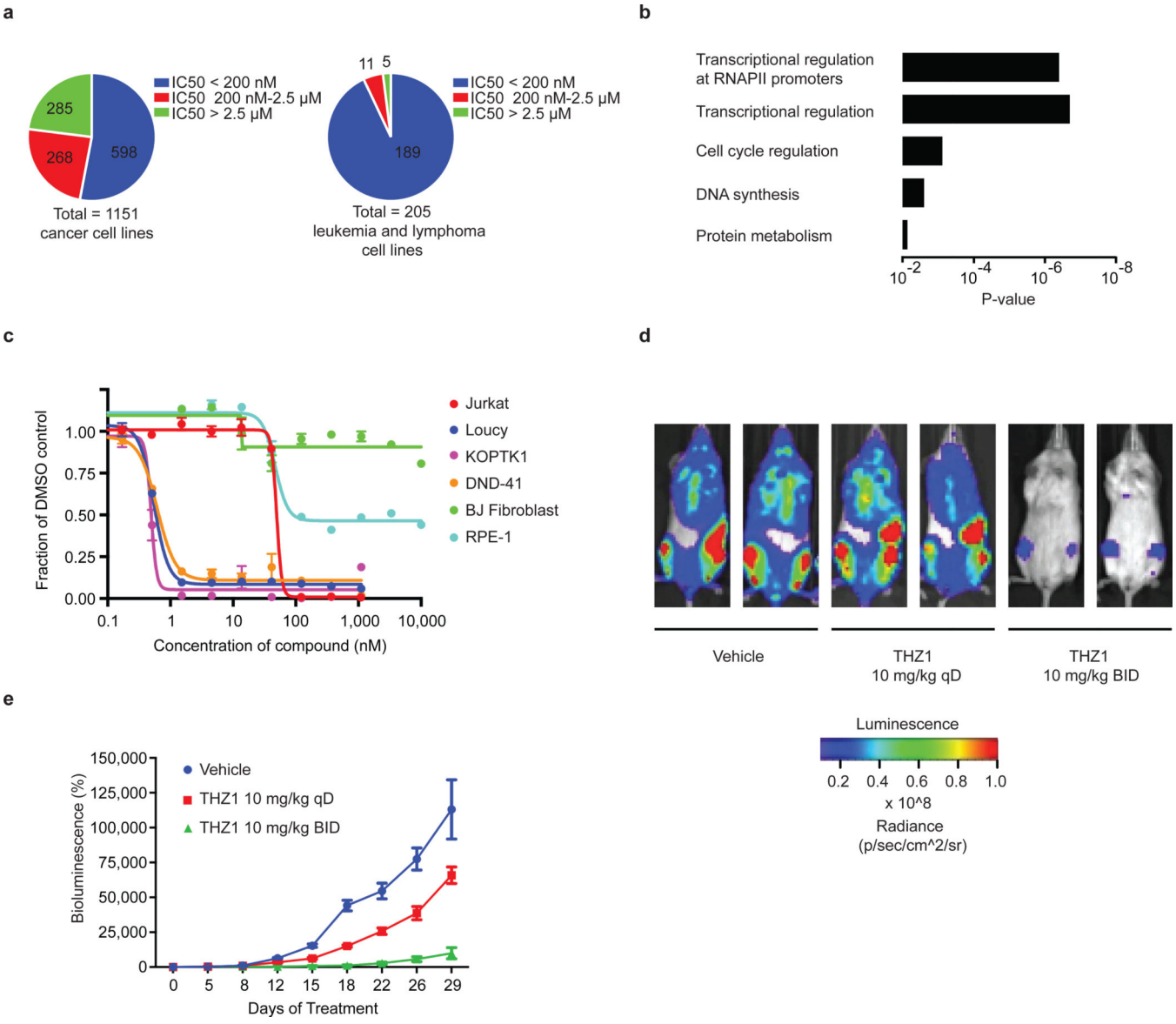


Figure 3. THZ1 strongly reduces the proliferation and cell viability of T-ALL cell lines

a, THZ1 exhibits strong anti proliferative effect across a broad range of cancer cell lines from various cancer types. Cells were treated with THZ1 or DMSO vehicle for 72 hrs and assessed for anti proliferative effect using resazurin. **b,** Overexpression of transcriptional regulators, including (proto) oncogenic transcription factors, is a strong predictor of cell line sensitivity to THZ1. GO terms associated with overexpressed factors found in THZ1 – sensitive cell lines. **c,** THZ1 exhibits strong anti proliferative affect against T-ALL cell lines. BJ fibroblasts and RPE-1 cells are shown as normal cell lines. Cells were treated with THZ1 or DMSO vehicle for 72 hrs. Experiments were performed in biological triplicates and error bars are +/− SD. **d,** THZ1 reduces the proliferation of KOPTK1 T-ALL cells in a human xenograft mouse model. Bioluminescent images of two representative mice treated with either vehicle control, 10 mg/kg THZ1 qD (once daily), or 10 mg/kg/day THZ1 BID (twice daily) for 29 days. **e,** Relative bioluminescence (BLI) of mice treated with vehicle, 10 mg/kg

THZ1 qD (once daily), or 10 mg/kg/day THZ1 BID (twice daily) during the 29 days of treatment. n=10 for all groups. Bioluminescence is shown relative to day 0 and is plotted as average ± SEM. Analysis of the BLI data by repeated measures (RM) two-way ANOVA reveals the anti-proliferative effect of treatment with THZ1 BID is highly statistically significantly different ($p<0.0001$) as compared to the other treatments.

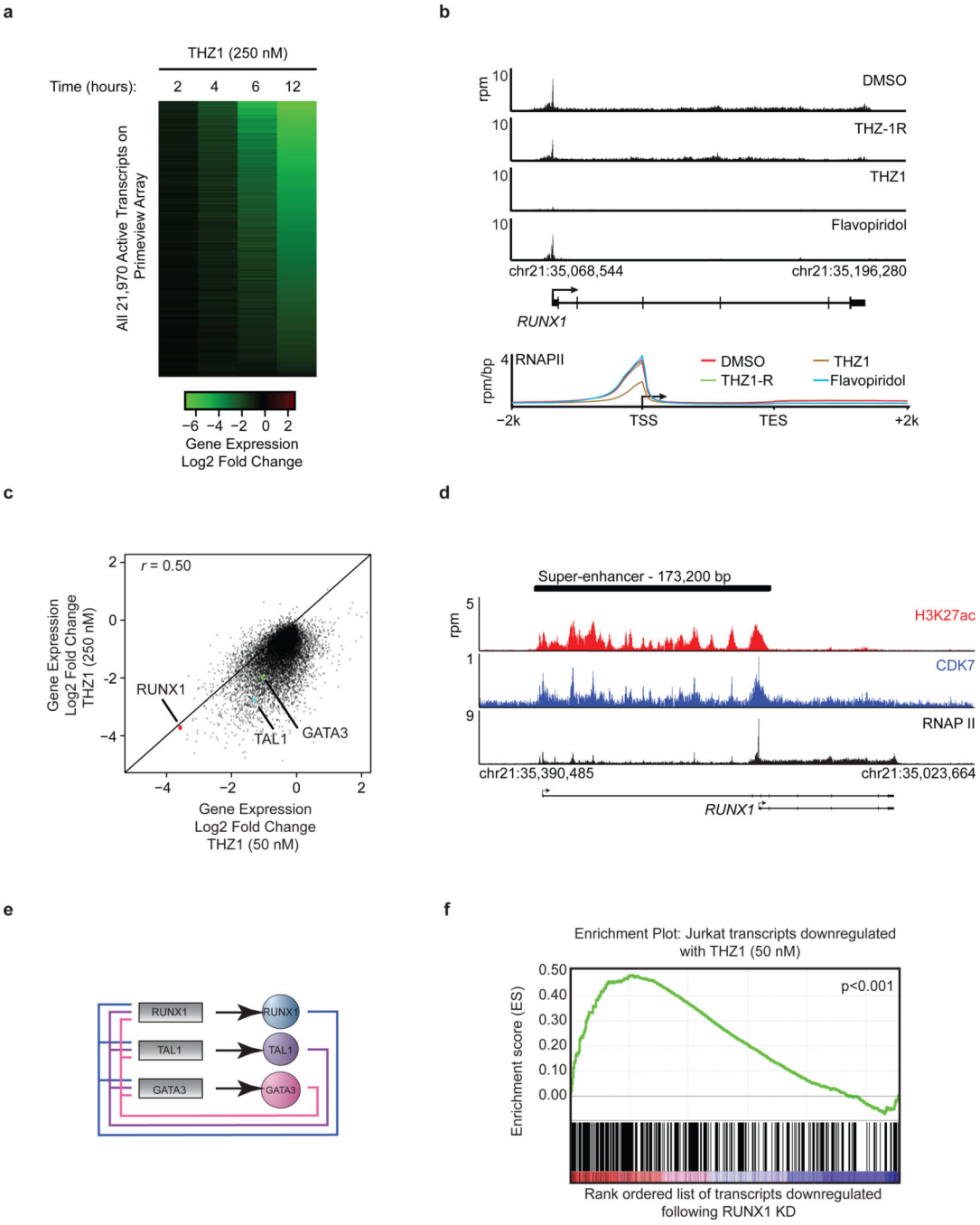


Figure 4. THZ1 preferentially downregulates Jurkat core transcriptional circuitry

a, THZ1 treatment globally downregulates steady-state mRNA levels in a time-dependent manner. Jurkat cells were treated with THZ1 (250 nM) for the indicated amounts of time. Heatmaps display the Log2 fold change in gene expression vs. DMSO for the 21,970 transcripts expressed at 12h in DMSO. **b,** THZ1 reduces RNAPII occupancy across promoters and gene bodies. Meta gene representation of global RNAPII occupancy at promoters and gene bodies (top). Average background subtracted ChIP-seq signal in 22,310 genes expressed in 6 h DMSO conditions in units of rpm/bp. Gene tracks of RNAPII ChIP-

seq occupancy at *RUNX1* following the indicated treatments (bottom). Signal of ChIP-seq occupancy is in units of reads per million (rpm). All treatments were 6 hrs with 250 nM of THZ1, THZ1-R, or Flavopiridol. **c,** THZ1 treatment delineates a subset of transcripts equally sensitive to low dose (50 nM) THZ1. Log2 fold change in gene expression for 50 nM (x axis) and 250 nM THZ1 (y axis) following a 4 hr treatment. Pearson coefficient r = 0.50. **d,** Gene tracks of H3K27Ac (top), CDK7 (middle), and RNAPII (bottom) ChIP-seq occupancy at the TSS, gene body, and a previously described enhancer region in the first intron of *RUNX1*[29]. Total ChIP-seq signal is in units of rpm. **e,** Positive interconnected autoregulatory loop formed by RUNX1, TAL1, and GATA3. Genes are represented by rectangles, and proteins are represented by ovals[21]. **f,** Transcripts down-regulated by low dose THZ1 are enriched for transcripts downregulated following *RUNX1* knockdown. Gene set enrichment analysis of top 500 transcripts downregulated following a 4-hour treatment with THZ1 (50 nM) in comparison to transcripts following a *RUNX1* knockdown[21]. GSEA-supplied p-value < 0.001.