Supplementary Information

Covalent targeting of remote cysteine residues to develop CDK12 and 13 inhibitors

Authors: Tinghu Zhang[1,2,10], Nicholas Kwiatkowski[1,2,3,10], Calla M Olson[1,2,10], Sarah E Dixon-Clarke[5], Brian J Abraham[3], Ann K Greifenberg[6], Scott B Ficarro[1,2,7], Jonathan M Elkins[5], Yanke Liang[1,2], Nancy M Hannett[3], Theresa Manz[1,8], Mingfeng Hao[1,2], Bartlomiej Bartkowiak[9], Arno L Greenleaf[9], Jarrod A Marto[1,2,7], Matthias Geyer[6], Alex N Bullock[5], Richard A Young[3,4*], Nathanael S Gray[1,2*]

Affiliations: [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]Structural Genomics Consortium, University of Oxford, Old Road Campus, Oxford OX3 7DQ, UK. [6]Department of Structural Immunology, Institute of Innate Immunity, University of Bonn, 53127 Bonn, and Center of Advanced European Studies and Research, 53175 Bonn, Germany. [7]Blais Proteomics Center, Dana-Farber Cancer Institute, Boston, MA 02115, USA. [8]Pharmaceutical and Medicinal Chemistry, Department of Pharmacy, Saarland University, 66123 Saarbrücken, Germany. [9]Department of Biochemistry, Duke University Medical Center, Durham, NC 27710, USA. [10]These authors contributed equally to this work.

*To whom correspondence may be addressed: Dept. of Biological Chemistry and Molecular Pharmacology, Harvard Medical School, 250 Longwood Ave., Boston, MA 02115. Tel.: 617-582-8590; Fax: 617-582-8615; E-mail: Nathanael_Gray@ dfci.harvard.edu or Whitehead Institute for Biomedical Research, 9 Cambridge Center, Cambridge, MA 02142. Tel.: 617-258-5218; E-mail: young@wi.mit.edu.

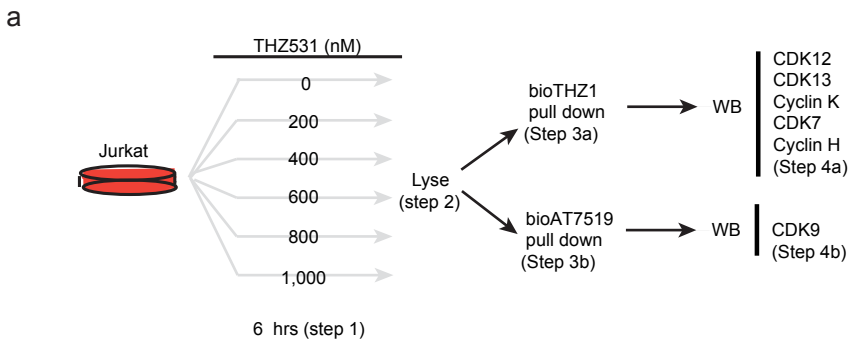
Nature Chemical Biology: doi:10.1038/nchembio.2166

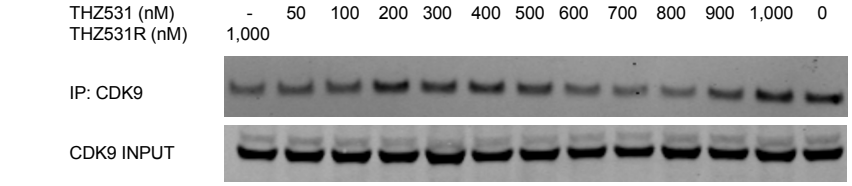
Supplementary Results

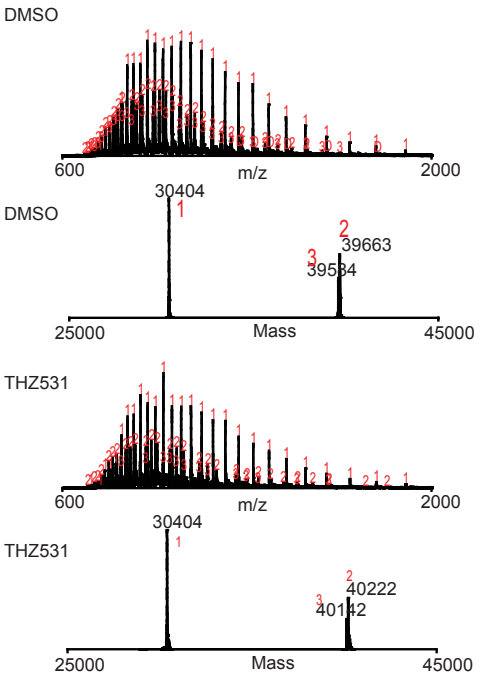
Supplementary Figures

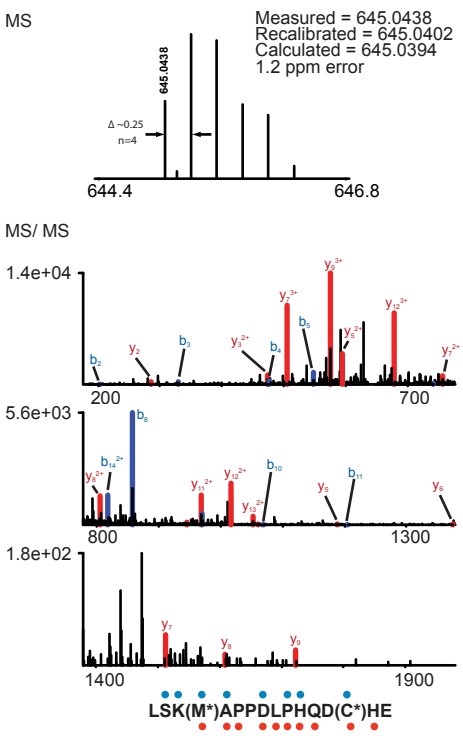
Supplementary Figure 1 | THZ531 targets CDK12 and 13 by covalent modification of distal C-terminal cysteine residues. **a.** Schematic of target engagement experiment. Jurkat cells were treated with increasing doses of THZ531 or DMSO for 6 hrs (Step 1). Cellular lysates were made from cells from each DMSO or THZ531 concentration point (Step 2). Clarified lysates from each treatment condition were then incubated with either 1μM bioTHZ1 (Step 3a), a concentration that binds CDK7-cyclin H, CDK12-cyclin K, and CDK13-cyclin K complexes or 1 μM bioAT7519 (Step 3b) a concentration known to bind CDK9. Subsequent addition of streptavidin -coated beads permits the immunoprecipitation of the indicated protein complexes. Following washing of beads with lysis buffer, the immunoprecipitated proteins were eluted from the beads by boiling in SDS buffer. Western blotting of precipitated proteins for CDK12, CDK13, or cyclin K was used to identify precipitated CDK12-cyclin K or CDK13-cyclin K complexes (Step 4a). Western blotting of precipitated proteins for CDK7 or cyclin H was used to identify precipitated CDK7-cyclin H complexes (Step 4a). And finally, CDK9 western blotting was used to identify precipitated CDK9 complexes (Step 4b). As THZ531 binds to its intended targets covalently, pretreatment of cells with THZ531 parent compound would be expected to block subsequent capture and immunoprecipitation of CDK12, 13, and 7 complexes with bioTHZ1 (or CDK9 with bioAT7519). Therefore, treating cells with THZ531 in dose titration permits us to ascertain at what concentration THZ531 is binding to each of these kinase complexes in cells, giving us a readout of intracellular selectivity. Uncut western blots are in Supplementary Fig. 10. **b.** THZ531 does not bind intracellular CDK9 complexes. THZ531 does not compete with lysate – introduced bioAT7519 for binding to CDK9. **c,** Mass spectra (top, middle bottom) and zero-charge mass spectra (middle top, bottom) of CDK12- cyclin K complex that were treated with DMSO (top, middle top) or THZ531 (middle bottom, bottom) for 1 hr at room temperature. After covalent bond formation, the masses of CDK12 (3) and phosphorylated CDK12 (2) increase by ~558 Da (THZ531). Although cyclin K (1) contains 7 cysteine residues it does not exhibit a mass shift indicating THZ531 does not form a covalent bond with this protein. **d,** Mass spectrum (top) and MS/MS spectrum (bottom) recorded during nanoLC-MS analysis of glu-c digested CDK12-cyclin K complex after treatment with THZ531 for 1 hr at room temperature illustrate detection of precursor (top) and product (bottom) ions of the peptide LSK(M*)APPDLPHQD(C*)HE (CDK12 residues 1026-1041). Ions y2 and y3 indicate C1039 forms a covalent bond with THZ531. Blue and red dots next to the sequence highlight detected ions of type b and y, respectively. (C*), THZ531 labeled cysteine; (M*), oxidized methionine. **e,** Structures of THZ531R and THZ532.

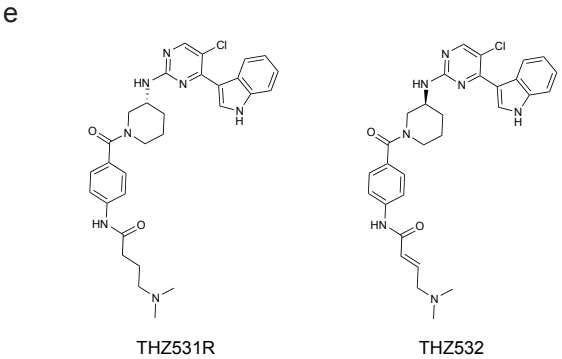
Nature Chemical Biology: doi:10.1038/nchembio.2166


a
THZ531 (nM)
0
200
400
600
800
1,000
Jurkat
6 hrs (step 1)
Lyse
(step 2)
bioTHZ1
pull down
(Step 3a)
bioAT7519
pull down
(Step 3b)
WB
WB
CDK12
CDK13
Cyclin K
CDK7
Cyclin H
(Step 4a)
CDK9
(Step 4b)


b
THZ531 (nM)
THZ531R (nM)
- 50 100 200 300 400 500 600 700 800 900 1,000 0
1,000
IP: CDK9
CDK9 INPUT

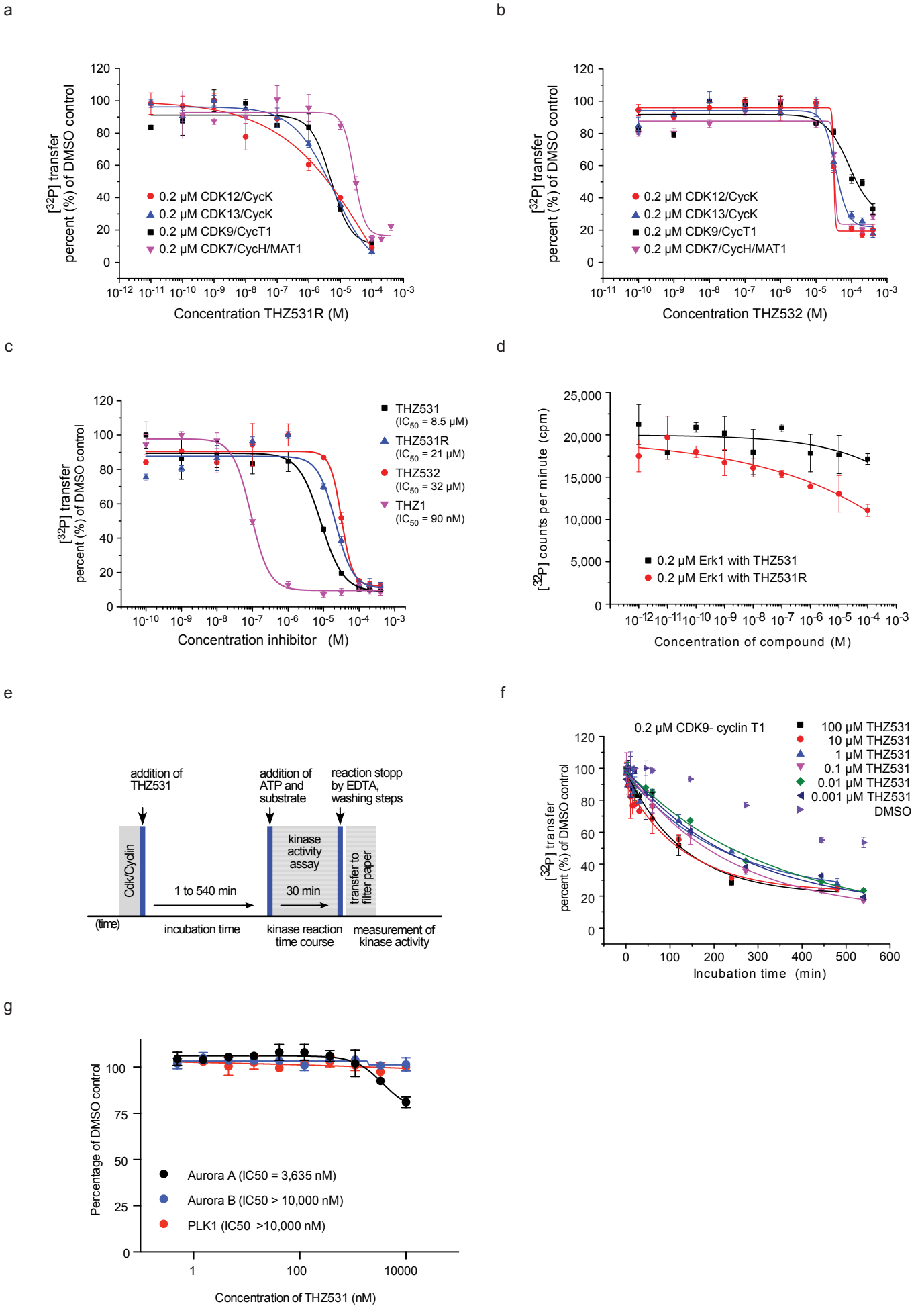

c
DMSO
600
m/z
2000
30404
39663
39584
25000
Mass
45000
THZ531
40222
40142


d
MS
Measured = 645.0438
Recalibrated = 645.0402
Calculated = 645.0394
1.2 ppm error
645.0438
644.4
646.8
MS/ MS
1.4e+04
5.6e+03
1.8e+02
200
700
800
1300
1400
1900
LSK(M*)APPDLPHQD(C*)HE


e
THZ531R
THZ532

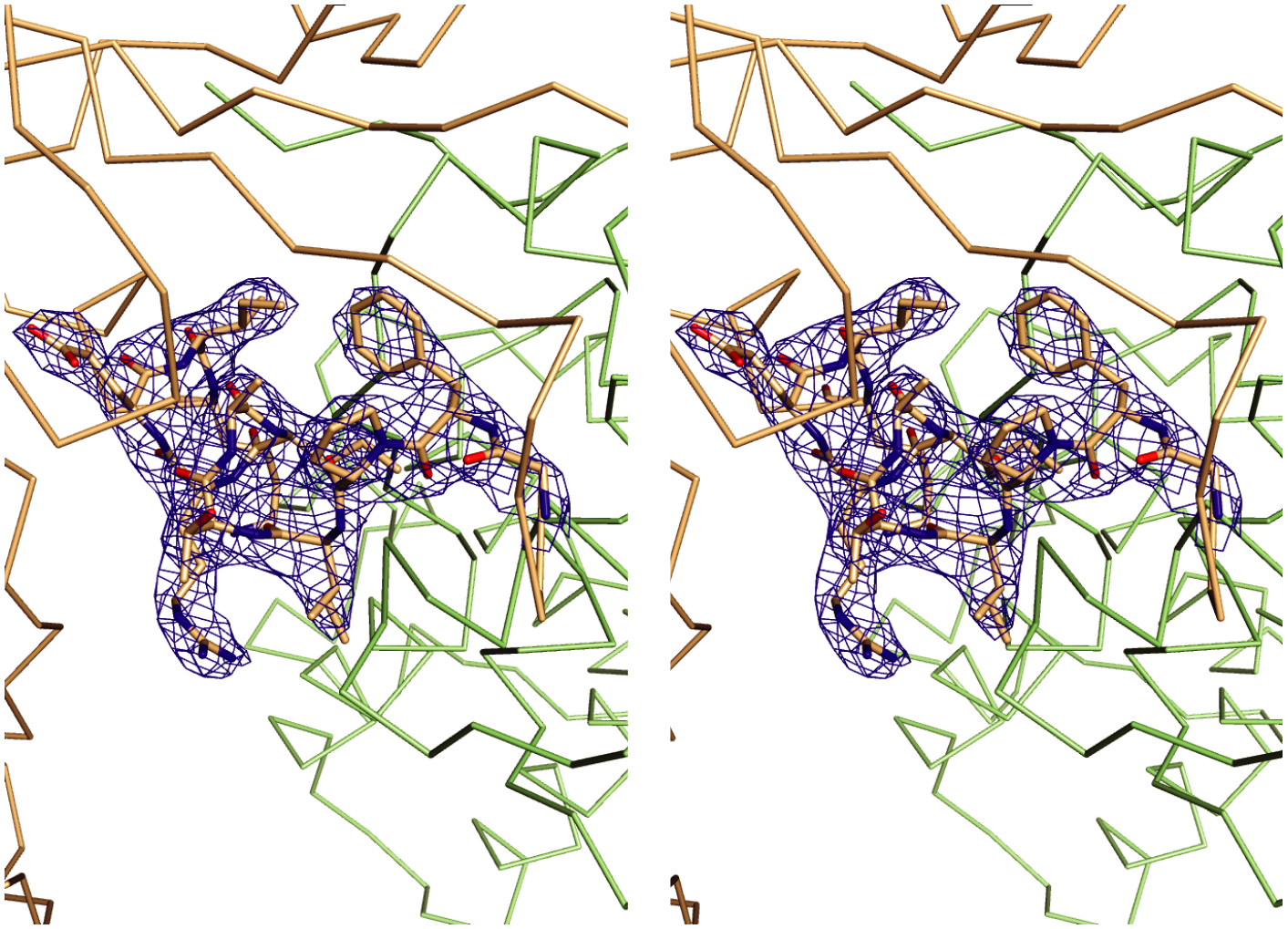
Supplementary Figure 2 | **Time series of transcription kinase inhibition at different pre-incubation times and varying inhibitor concentrations.** **a,** Kinase inhibition assay using THZ531R, the reversible analog of THZ531, resulted in similarly high IC_{50} values against CDK12, CDK13, CDK9, and CDK7. Measurements were made in triplicate and data represent the mean ± S.D. **b,** Kinase inhibition assays using THZ532, the inactive enantiomer of THZ531, produced similarly high IC_{50} values against CDK12, CDK13, CDK9, and CDK7. Measurements were made in triplicate and data represent the mean values ± S.D. **c,** THZ531 exhibits reduced activity against CDK7. Measurements were made in triplicate and data represent the mean values ± S.D. **d,** THZ531 and THZ531R do not inhibit Erk1. *In vitro* kinase activity assays using recombinant protein were applied to analyze the effect of the THZ531 compound on a member of the MAP kinase family. Measurements were made in triplicate and data represent the mean values ± S.D. **e,** Assay schematic: To a concentration of 0.2 μM CDK- cyclin complex different concentrations of THZ531 were added, ranging from 0.001 μM to 100 μM. Pre-incubation times of 1 min to 9 hrs were followed before the kinase reaction was started by addition radioactively labeled ATP and substrate peptide. A kinase reaction time course of 30 min was applied before the reaction was stopped and the kinase activity measured. The incubation time and the kinase activity time course were performed at 30°C at 350 r.p.m. Measurements were made in triplicate and data represent the mean values ± S.D. **f,** *In vitro* kinase activity assay of 0.2 μM CDK9- cyclin T1 after different preincubation times with varying concentrations of THZ531. Increasing concentrations of THZ531 do not result in significantly reduced kinase activity at longer pre-incubation times, supporting that the CDK9–THZ531 interaction is reversible. As control, the decrease of kinase activity in the absence of THZ531 was measured to monitor the loss of enzymatic activity over time. The counts per minute of the kinase activity measurements were normalized to the relative [^{32}P] transfer. Measurements were made in triplicate and data represent the mean values ± S.D. **g,** THZ531 has lower affinity for PLK1, Aurora A and Aurora B, compared to CDK12/13. *In vitro* kinase assays were performed by Life Technologies in duplicate at an ATP concentration = K_m for each kinase. Data represent the mean of values ± S.D.

Nature Chemical Biology: doi:10.1038/nchembio.2166


a
[32P] transfer percent (%) of DMSO control
0.2 μM CDK12/CycK
0.2 μM CDK13/CycK
0.2 μM CDK9/CycT1
0.2 μM CDK7/CycH/MAT1
Concentration THZ531R (M)
b
[32P] transfer percent (%) of DMSO control
0.2 μM CDK12/CycK
0.2 μM CDK13/CycK
0.2 μM CDK9/CycT1
0.2 μM CDK7/CycH/MAT1
Concentration THZ532 (M)
c
[32P] transfer percent (%) of DMSO control
THZ531 (IC50 = 8.5 μM)
THZ531R (IC50 = 21 μM)
THZ532 (IC50 = 32 μM)
THZ1 (IC50 = 90 nM)
Concentration inhibitor (M)
d
[32P] counts per minute (cpm)
0.2 μM Erk1 with THZ531
0.2 μM Erk1 with THZ531R
Concentration of compound (M)
e
addition of THZ531
addition of ATP and substrate
reaction stopp by EDTA, washing steps
Cdk/Cyclin
kinase activity assay
transfer to filter paper
1 to 540 min
30 min
(time)
incubation time
kinase reaction time course
measurement of kinase activity
f
0.2 μM CDK9- cyclin T1
100 μM THZ531
10 μM THZ531
1 μM THZ531
0.1 μM THZ531
0.01 μM THZ531
0.001 μM THZ531
DMSO
[32P] transfer percent (%) of DMSO control
Incubation time (min)
g
Percentage of DMSO control
Aurora A (IC50 = 3,635 nM)
Aurora B (IC50 > 10,000 nM)
PLK1 (IC50 >10,000 nM)
Concentration of THZ531 (nM)

Supplementary Figure 3 | Electron density for THZ531 and the PITAIRE helix of CDK12. Stereo view showing electron density (2Fo-Fc contoured at 1.0σ) of the CDK12 PITAIRE helix (αC).

Nature Chemical Biology: doi:10.1038/nchembio.2166

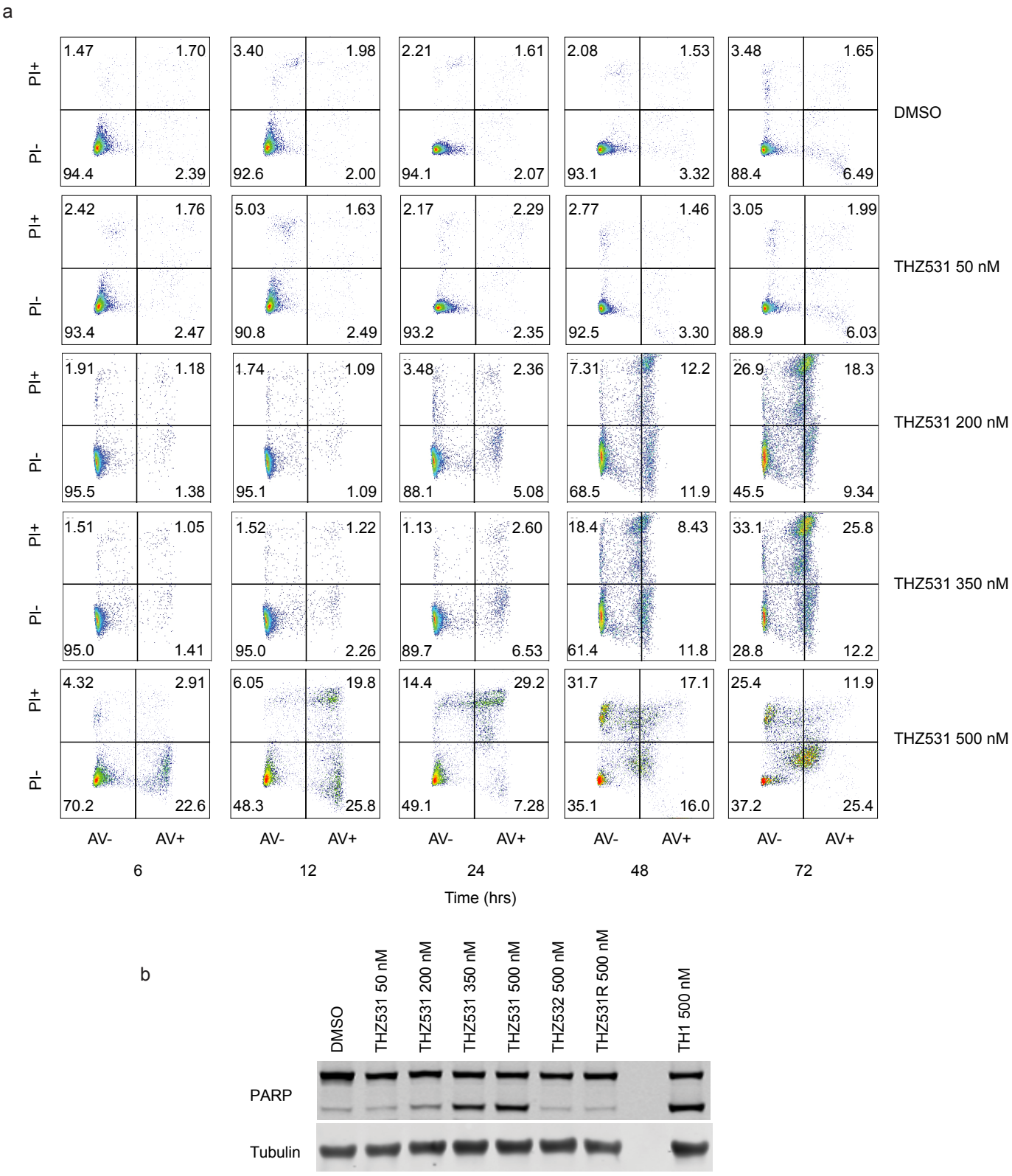


Nature Chemical Biology: doi:10.1038/nchembio.2166

Supplementary Figure 4 | THZ531 induces apoptosis to Jurkat cells

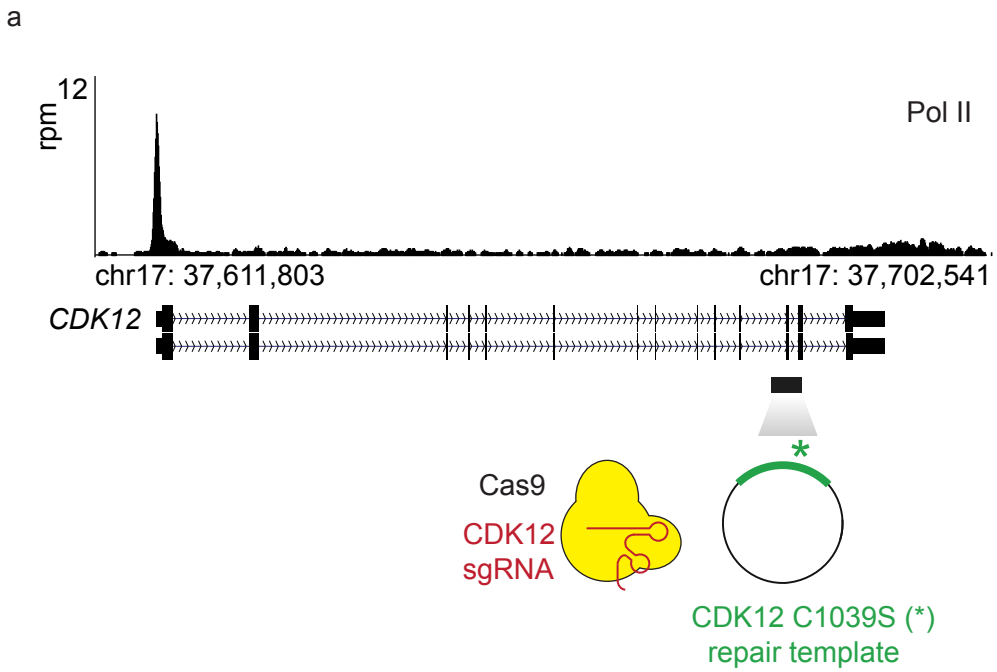
a. Representative Annexin V and propidium iodide stainings for Jurkat cells incubated with THZ531. Jurkat cells were treated with 50, 200, 350, and 500 nM THZ531 for the indicated times. Cells were stained with Annexin V and propidium iodide. Experiments were performed in biological triplicate. **b.** THZ531 induces PARP cleavage. Jurkat cells were treated with 50, 200, 350, and 500 nM THZ531 for 24 hours. Lysates were probed with PARP and tubulin antibodies. Uncut western blots are in Supplementary Fig. 10.

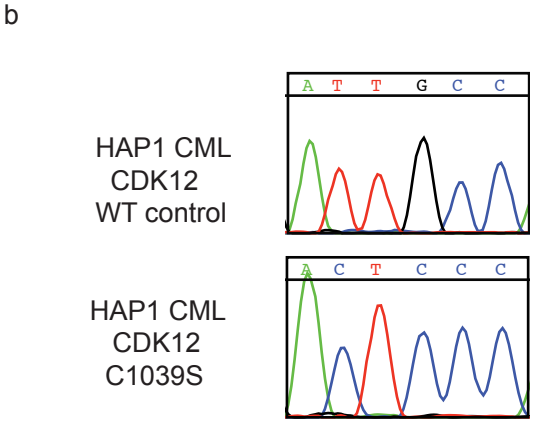
Nature Chemical Biology: doi:10.1038/nchembio.2166

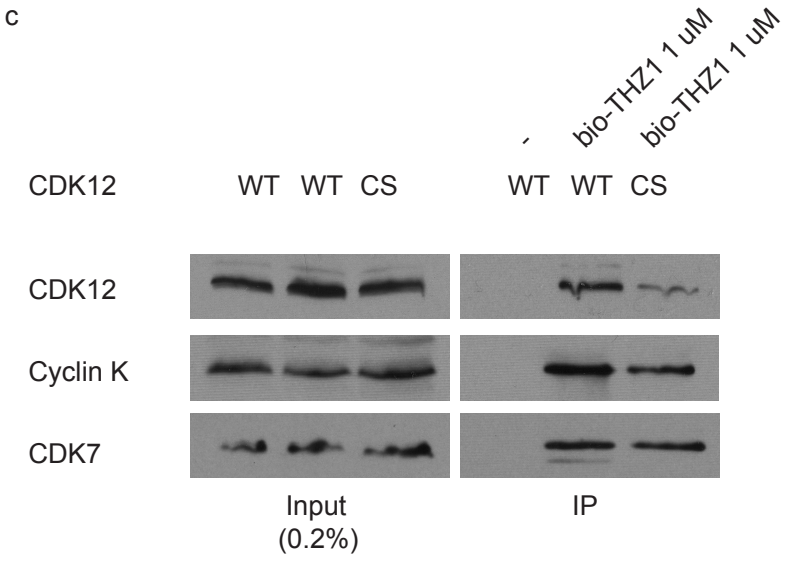

a
PI+
PI-
1.47
1.70
94.4
2.39
3.40
1.98
92.6
2.00
2.21
1.61
94.1
2.07
2.08
1.53
93.1
3.32
3.48
1.65
88.4
6.49
DMSO
2.42
1.76
93.4
2.47
5.03
1.63
90.8
2.49
2.17
2.29
93.2
2.35
2.77
1.46
92.5
3.30
3.05
1.99
88.9
6.03
THZ531 50 nM
1.91
1.18
95.5
1.38
1.74
1.09
95.1
1.09
3.48
2.36
88.1
5.08
7.31
12.2
68.5
11.9
26.9
18.3
45.5
9.34
THZ531 200 nM
1.51
1.05
95.0
1.41
1.52
1.22
95.0
2.26
1.13
2.60
89.7
6.53
18.4
8.43
61.4
11.8
33.1
25.8
28.8
12.2
THZ531 350 nM
4.32
2.91
70.2
22.6
6.05
19.8
48.3
25.8
14.4
29.2
49.1
7.28
31.7
17.1
35.1
16.0
25.4
11.9
37.2
25.4
THZ531 500 nM
AV-
AV+
6
12
24
48
72
Time (hrs)
b
DMSO
THZ531 50 nM
THZ531 200 nM
THZ531 350 nM
THZ531 500 nM
THZ532 500 nM
THZ531R 500 nM
TH1 500 nM
PARP
Tubulin

Supplementary Figure 5 | Mutation of Cys-1039 to serine reduces CDK12 covalent affinity and rescues THZ531 –induced proliferation defects. a, Gene track of CDK12 with schematic of CRISPR technique to mutate CDK12 allele. **b,** CRISPR technique mutates C1039 to serine (C1039S). Genomic DNA from CDK12 WT control and CDK12 C1039S HAP1 cells were Sanger sequenced. TGC (Cys) was successfully mutated to TCC (Ser). Other silent mutations were added to remove NGG CRISPR targeting sequence and to permit initial PCR screening of mutated alleles. **c,** CDK12 C1039S mutation prevents CDK12 pulldown with bioTHZ1. 25 million cells of WT control and C1039S HAP1 clones were lysed and probed with 1 μM bioTHZ1 at 4 degrees overnight. Interacting proteins were precipitated with streptavidin beads and probed with indicated antibodies. Uncut westerns blots are in Supplementary Fig. 10.

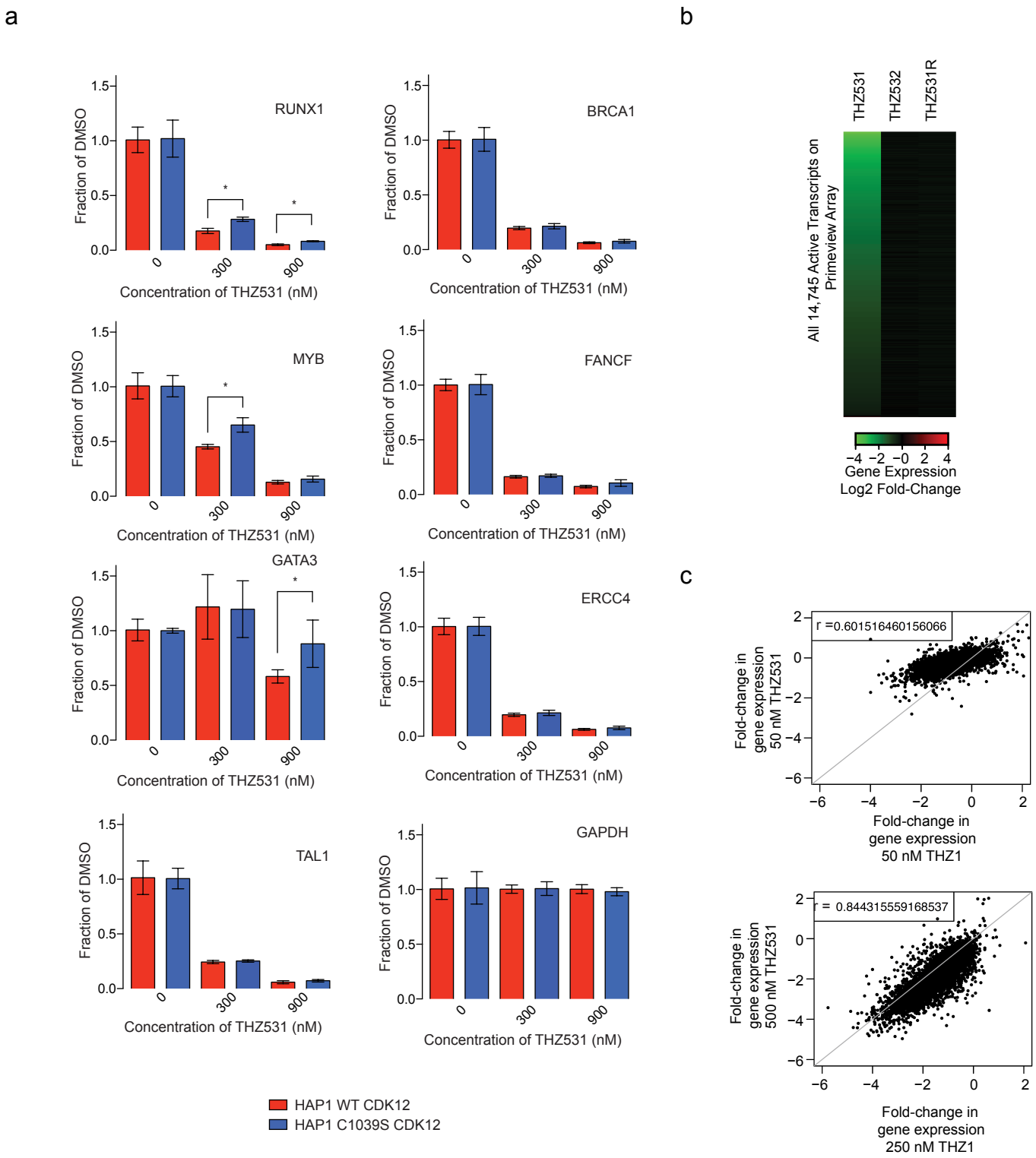
Nature Chemical Biology: doi:10.1038/nchembio.2166


a
rpm
12
Pol II
chr17: 37,611,803
chr17: 37,702,541
CDK12
Cas9
CDK12
sgRNA
*
CDK12 C1039S (*)
repair template


b
HAP1 CML
CDK12
WT control
A T T G C C
HAP1 CML
CDK12
C1039S
A C T C C C

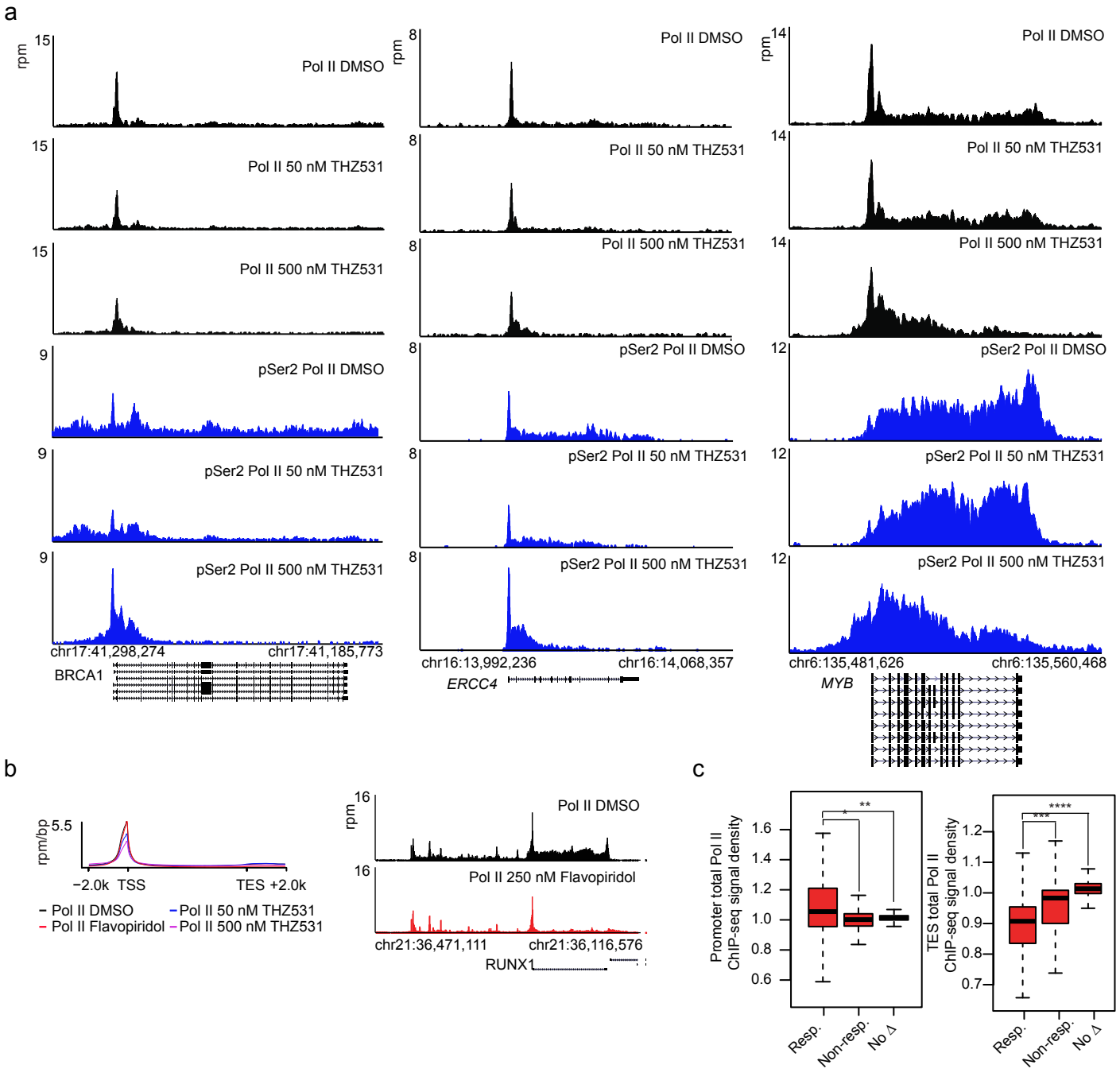

c
-
bio-THZ1 1 uM
bio-THZ1 1 uM
CDK12
WT WT CS
WT WT CS
CDK12
Cyclin K
CDK7
Input
(0.2%)
IP

Supplementary Figure 6 | **THZ531 inhibits gene expression** **a,** Expression of CDK12 C1039S partially restores T-ALL transcription factor gene expression. RT qPCR of T-ALL transcription factors and DDR gene transcripts. RT qPCRs were performed in biological triplicate and error bars are +/- SD. **b,** THZ531R and THZ532 do not affect steady-state mRNA levels. Jurkat cells were treated with 500 nM THZ531, THZ531R, or THZ532 for 6 hrs. Heatmaps display the log2 fold-change in gene expression vs. DMSO for the 14,745 transcripts expressed in DMSO. **c,** THZ1 and THZ531 display similar yet distinct effects on the expression of 14,745 expressed genes (in DMSO). Log2 fold-change in gene expression for 50 nM THZ1 vs. THZ531 (left) and 250 nM THZ1 vs. 500 nM THZ531 (right). Pearson coefficient r = 0.60 and 0.84 respectively.


a
RUNX1
BRCA1
MYB
FANCF
GATA3
ERCC4
TAL1
GAPDH
Fraction of DMSO
Concentration of THZ531 (nM)
HAP1 WT CDK12
HAP1 C1039S CDK12
b
THZ531
THZ532
THZ531R
All 14,745 Active Transcripts on Primeview Array
Gene Expression Log2 Fold-Change
c
r =0.601516460156066
Fold-change in gene expression 50 nM THZ531
Fold-change in gene expression 50 nM THZ1
r = 0.844315559168537
Fold-change in gene expression 500 nM THZ531
Fold-change in gene expression 250 nM THZ1

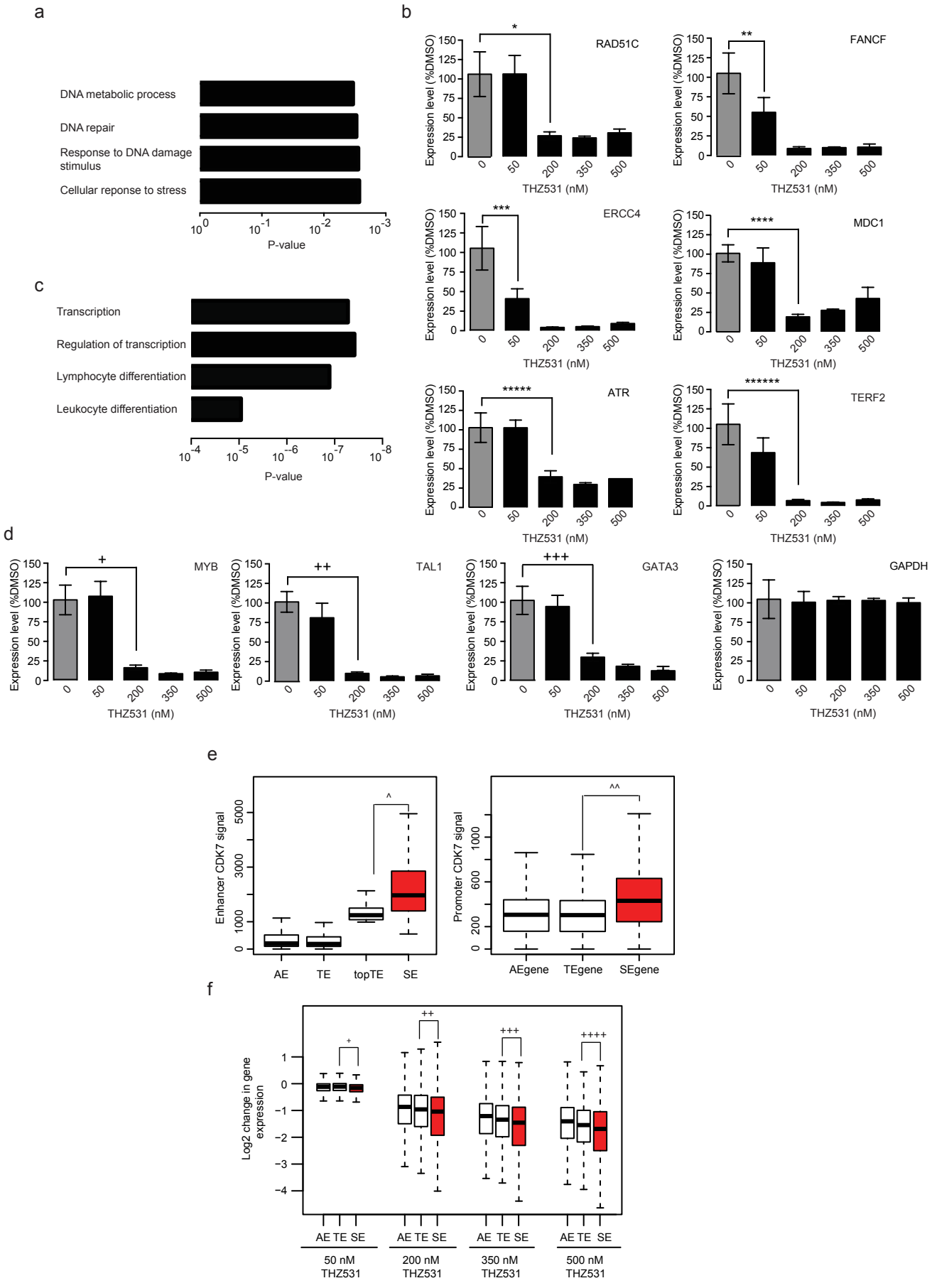
Supplementary Figure 7 | **THZ531 inhibits transcriptional elongation. a,** Gene tracks whose expression is sensitive to THZ531. **b,** THZ531 and Flavopiridol show different effects on Pol II distributions. Metagene representation of global Pol II occupancy across gene bodies. Average ChIP-seq signal in 13906 genes expressed in 6h DMSO conditions in units of rpm/bp (left). Gene tracks of Pol II ChIP-seq at *RUNX1* gene locus following 250 nM Flavopiridol treatment for 6 hrs. (right). **c**, Flavopiridol significantly increases promoter–bound Pol II ChIP-seq signal and decreases elongating Pol II ChIP-seq signal at Flavopiridol–responsive genes relative to non-responsive genes. Box plots of Pol II ChIP-seq signal density at 5' transcriptional start sites (TSS) and 3' termination sites (TSS) at 2001 Flavopiridol –responsive genes (Resp.), 2001 non-responsive (Non-resp.), and 2001 genes whose expression doesn't change (no Δ). *p-value = 9.37e-54, **p-value = 2.60e-85, ***p-value = 2.27e-97, ****p-value = 1.34e-296. Responsive genes are defined as those having > log2 fold-change in gene expression.

Nature Chemical Biology: doi:10.1038/nchembio.2166


a
rpm
Pol II DMSO
Pol II 50 nM THZ531
Pol II 500 nM THZ531
pSer2 Pol II DMSO
pSer2 Pol II 50 nM THZ531
pSer2 Pol II 500 nM THZ531
chr17:41,298,274
chr17:41,185,773
BRCA1
chr16:13,992,236
chr16:14,068,357
ERCC4
chr6:135,481,626
chr6:135,560,468
MYB
b
rpm/bp
−2.0k TSS
TES +2.0k
Pol II DMSO
Pol II 50 nM THZ531
Pol II Flavopiridol
Pol II 500 nM THZ531
Pol II 250 nM Flavopiridol
chr21:36,471,111
chr21:36,116,576
RUNX1
c
Promoter total Pol II ChIP-seq signal density
TES total Pol II ChIP-seq signal density
Resp.
Non-resp.
No Δ

Supplementary Figure 8 | THZ531 downregulates DDR and transcription factor gene expression **a**, The top 2% of genes downregulated with 50 nM show enrichment for genes encoding factors that regulate DDR. David gene ontology analysis, p-values supplied by David program. **b,** RT-qPCR of additional DDR genes transcript expression following THZ531 treatment. *p-value=2.48e-05, **p-value=1.93e-03, ***p-value=2.05e-04, ****p-value=6.71e-10, *****p-value=6.13e-06, ******p-value=1.04e-06. **c,** The top 2% of genes downregulated with 200 nM show enrichment for genes encoding factors that regulate transcription. David gene ontology analysis, p-values supplied by David program. **d,** RT-qPCR of additional T-ALL transcription factor gene transcripts following THZ531 treatment. +p-value=1.29e-07, ++p-value=1.11e-09, +++p-value=6.04e-07. **e,** Super –enhancers and promoters of their associated genes contain more CDK7 ChIP-seq signal compared to typical enhancers and their associated gene promoters. Boxplots demonstrating CDK7 ChIP-seq enhancer (left) and promoter (right) signal at all enhancers (AE), typical enhancers (TE), the top 818 TEs (Top TEs), and the 818 super –enhancers (SE). ^p-value = 4.68e-51, ^^p-value = 4.58e-142. **f**, Super–enhancer –associated gene expression is more sensitive to THZ531. Boxplots showing the fold-change in gene expression for those genes associated with AEs, TEs, and SEs. +p-value = 3.78e-06, ++p-value = 8.90e-06, +++p-value = 3.16e-06, ++++p-value = 1.79e-06, calculated with the two-tailed Student's t test. All RT-qPCR experiments were performed in biological triplicate and error bars are +/- SD. GAPDH gene expression was used as internal control for all RT-qPCR experiments. P-values were determined with a two-tailed Student's T test.

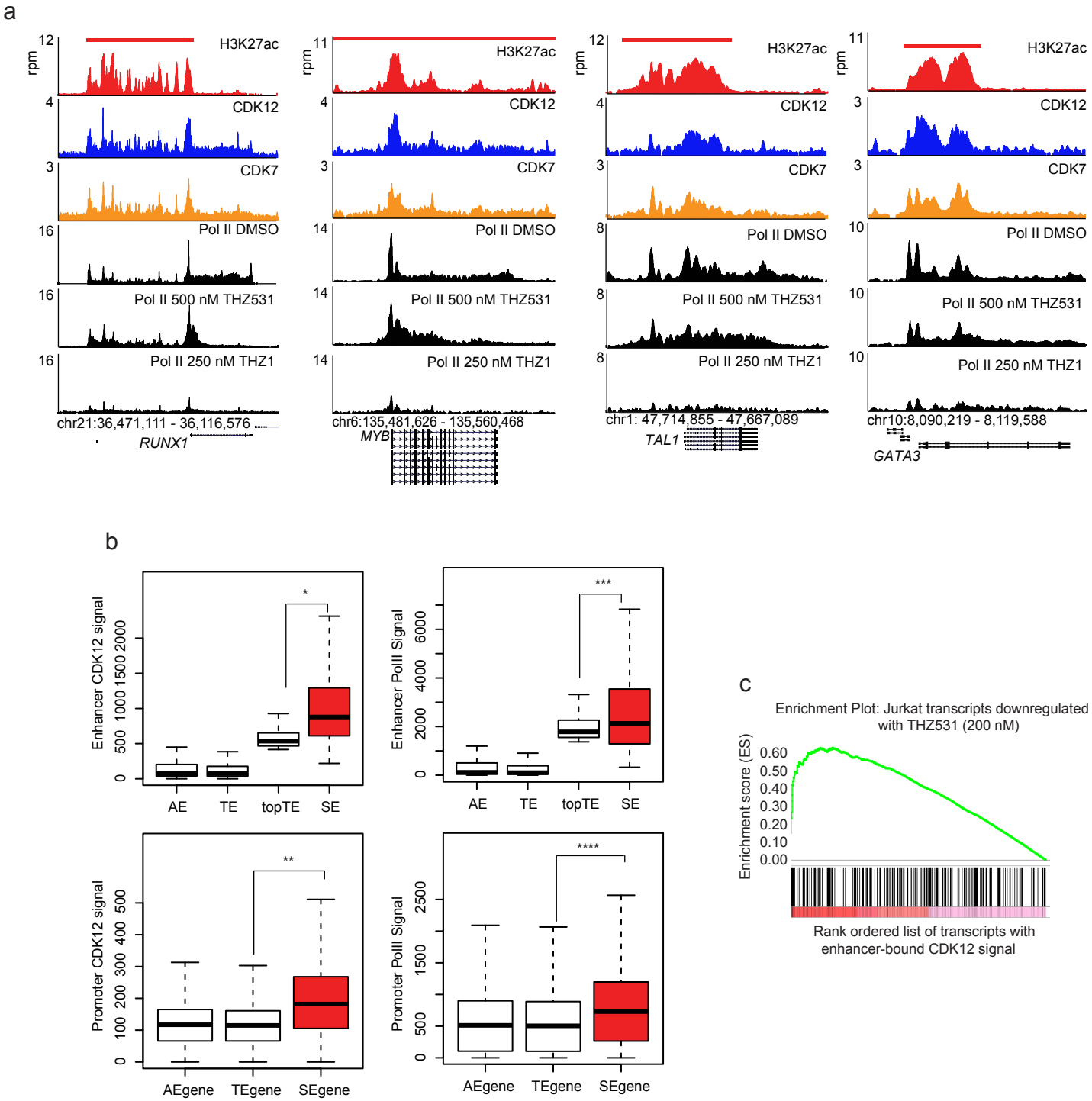
Nature Chemical Biology: doi:10.1038/nchembio.2166


a
DNA metabolic process
DNA repair
Response to DNA damage stimulus
Cellular reponse to stress
P-value
b
RAD51C
FANCF
ERCC4
MDC1
ATR
TERF2
Expression level (%DMSO)
THZ531 (nM)
c
Transcription
Regulation of transcription
Lymphocyte differentiation
Leukocyte differentiation
P-value
d
MYB
TAL1
GATA3
GAPDH
e
Enhancer CDK7 signal
Promoter CDK7 signal
AE
TE
topTE
SE
AEgene
TEgene
SEgene
f
Log2 change in gene expression
50 nM THZ531
200 nM THZ531
350 nM THZ531
500 nM THZ531

Supplementary Figure 9 | Super-enhancers contain exceptional amounts of CDK12.

a, *S*uper-enhancer genes contain large amounts of CDK7 (yellow), CDK12 signal (blue), and H3K27Ac (red). Pol II (black) elongation is impacted following treatment with both 250 nM THZ1 and 500 nM THZ531. The red bar indicates the genomic coordinates of a super –enhancer. **b,** SEs (top) and their associated gene promoters (bottom) contain more CDK12 and Pol II ChIP-seq signal compared to TEs and their associated gene promoters. Boxplots demonstrating the ChIP-seq signal for CDK12 (top left) and Pol II (top right) at AEs, TEs, 818 Top TEs , and 818 SEs. Boxplots demonstrating the ChIP-seq signal for CDK12 (bottom left) and Pol II (bottom right) at gene promoters associated with AEs, TEs, and SE. *p-value = 2.63e-124, **p-value = 3.16e-46, ***p-value = 4.58e-12, ****p-value = 9.39e-18. **f**, Transcripts down-regulated by 200 nM THZ531 are enriched for transcripts whose associated enhancers contain the highest levels of CDK12 ChIP-seq signal. Gene set enrichment analysis of top 500 transcripts downregulated following a 6-hour treatment with THZ531 (200 nM) in comparison to CDK12 ChIP-seq signal at enhancers for these transcripts GSEA-supplied p-value < 0.001.

Nature Chemical Biology: doi:10.1038/nchembio.2166


a
H3K27ac
CDK12
CDK7
Pol II DMSO
Pol II 500 nM THZ531
Pol II 250 nM THZ1
rpm
chr21:36,471,111 - 36,116,576
RUNX1
chr6:135,481,626 - 135,560,468
MYB
chr1: 47,714,855 - 47,667,089
TAL1
chr10:8,090,219 - 8,119,588
GATA3
b
Enhancer CDK12 signal
Enhancer PolII Signal
Promoter CDK12 signal
Promoter PolII Signal
AE
TE
topTE
SE
AEgene
TEgene
SEgene
c
Enrichment Plot: Jurkat transcripts downregulated with THZ531 (200 nM)
Enrichment score (ES)
Rank ordered list of transcripts with enhancer-bound CDK12 signal

Supplementary Figure 10 | Uncut western blots

Uncut western blots corresponding to cropped western blots in main and supplementary figures.

Nature Chemical Biology: doi:10.1038/nchembio.2166

Fig. 1c

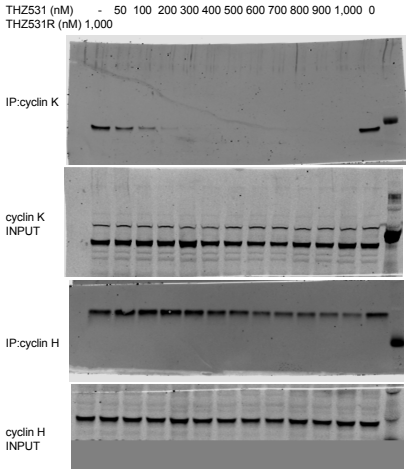


Fig. 1d

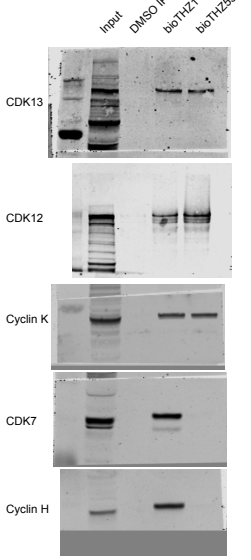


Fig. 3f

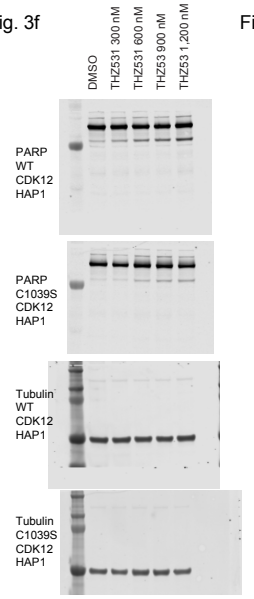


Fig. 5a

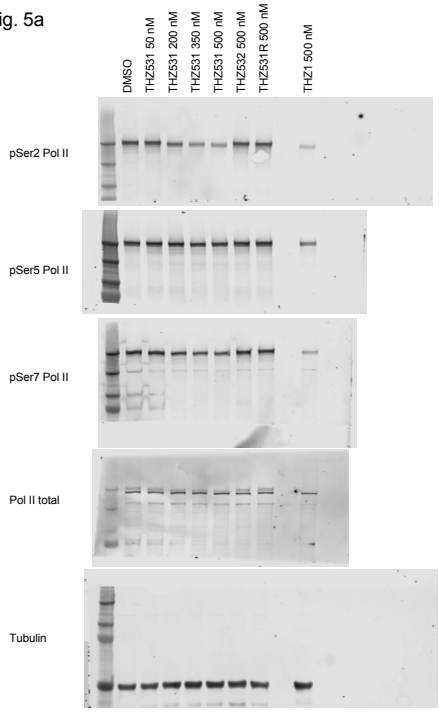
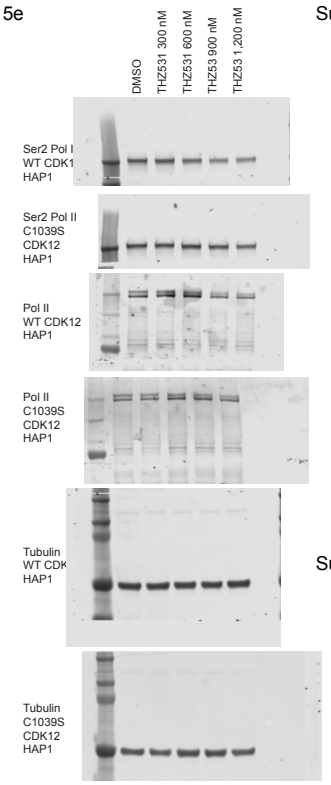
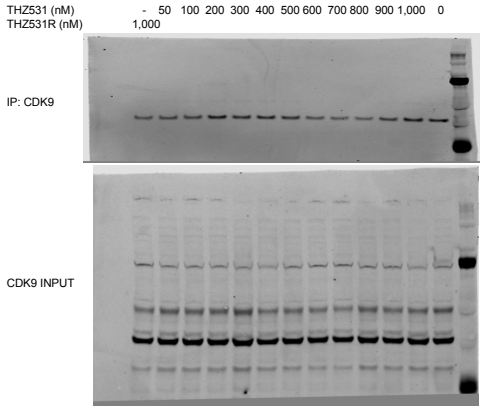


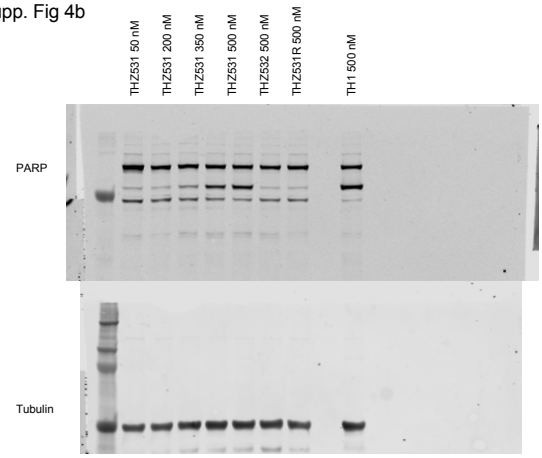
Fig. 5e



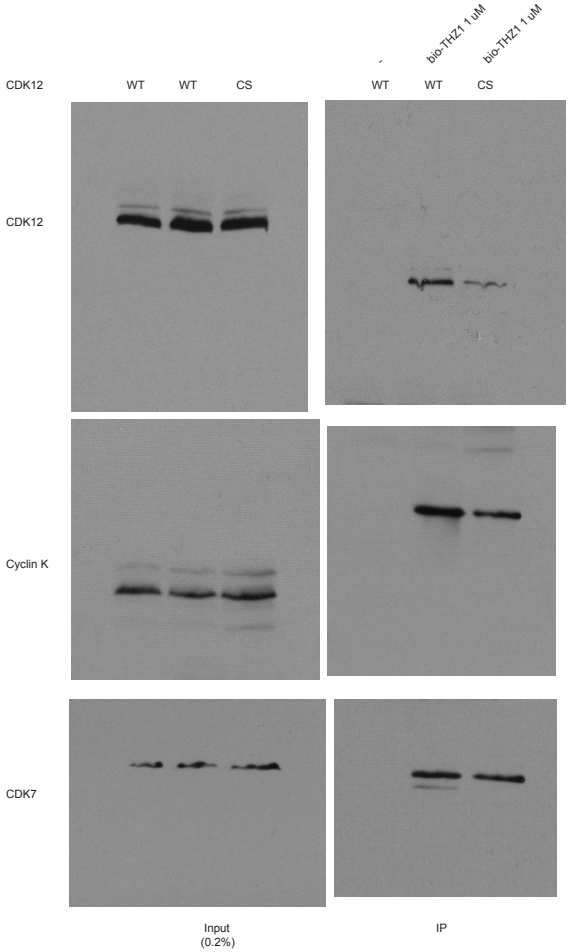
Supp. Fig 1b



Supp. Fig 4b



Supp. Fig 5c



Nature Chemical Biology: doi:10.1038/nchembio.2166

Supplementary Figure 11 | CDK12 genomic sequence for genome editing

Modified CDK12 genomic sequence used as repair template in genomic editing experiments.

Nature Chemical Biology: doi:10.1038/nchembio.2166

The modified CDK12 genomic sequence (RefSeq Accession NM_016507) was cloned into pUC57-AMP by Genewiz and used as the repair template for genome editing:

GCCATGAGGA**GTCGAC**ATTACACTAGAATGTTGATACATTGAATATGACTTGAAACATATAAGGGTTTTA
CTGAAATTTGGGAACTCCTATTATAGAGAGATTTATAGTAAGATTTGACCTACTTGGTCTGATTGTATGA
GAGATTCTGGACTAAGTTATACGTAGTGCTGAACTCCACAGATGAGCAAGTTGATATCCAGTCATAGAA
ATTCTGTGGGCAGAAAGGGGAAAGAACTAGTCTTTGGTCCTCACAAACCAAATTACACATAAGTTGATTT
TGCACAGAGATGTTTTGATCATGATACATGGTTCCATGGTTTAAGTCACCGCCTTTCCTCGACCTTCTGT
ATACATTAACAAGCCAGCCATTACTGTCCTCGTCTTTATTCCCTACTGAAA
CTTGACCAATACTGTTCTCCTTTTTTGTTGTTATGCCAAGGAAAGACAGTATTTATATGGGAATTTATATA
GCTGGTCTGTACCTAGTATTAGCAAGATCCTCCTTTCTCACTATGTAACTTTCGTACTTTTATTTCCTCAG
ACTACCATCTGCCATGATTTATTTTACACTGCTGTTACCTACTTTTACTAACTTTTTTGGTTATTTTTGTTT
TCCTTATTCTTTACATTTCCCACAGTCTTTGCCTTCCCATTTTAATCCTTTGCTCCTCCATTCATTCACTG
CTGCATTCCCTTACATATTTCCCCTTTGCTTTGTCTTTTTCCAGCCTCCCCCACTGGCA**A**GA**C**T**C**CCA**C**G
AG**C**T**C**TGGAGTAAGAAACG**C**CGACGTCAGCGACAAAGTGGTGTTGTAGTCGAAGAGCCACCTCCATCC
AAAACTTCTCGAAAAGAAACTACCTCAGGGACAAGTACTGAGCCTGTGAAGAACAGCAGCCCAGCACC
ACCTCAGCCTGCTCCTGGCAAGGTGGAGTCTGGGGCTGGGGATGCAATAGGTCAGTGCCAGAATGGG
GCCTTTGTGCTTTTGCTAAGCGTATTTGGCAGGTTTTGAAGGTCAAGTGTAAGGAGTTTGTGTGTGTGT
CTGTTTTTTAGTTTTTGGTCCAGTGAACCATTAGGAAAGAGGTTAATGTCCTCATTATGGTGAGAAGGAG
ATTAACCAAGTTTCGTACATATGGTCCAGAAAGTCTACAATAGATCATTTTCCTTTCCTGAAATATGTACG
GAACACTGTATCTCAATTAGATGTGTGGCATAAATCAGTGAATTAGGTAAGCCAAATTTAATTTAAGGTT
AGTTCACCCATGACTTTATGAAACACATAATTTGGAGCCTTGTTTATATTAGATTTGATAAGACCTTCATA
GAATCTAGTAGTGTTAAATTACAATTGCTAAATGTTGTTCTATAGACTGTTATTTTGATCTAAATGTATCTG
AAACTATGTCATATTGAACTGTTATAACAAAGTAGTTCTCCTGTGTTATGACCAATTCTTCGGGACAAAA
GATACTCTTTTTCTGGTTCACTTTCCCCTCTTC**GAATTC**TCGAGTACCG

1. Green highlighting indicates the introduced desired TCC mutation, which codes for serine (C1039S), replacing TGC which codes for cysteine (C1039, WT)
2. Yellow highlighting indicates wobble mutations introduced to remove Cas9 –targeting cysteines, 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

Sanger sequence of confirmed mutant C1039S allele:

TGATCCGCCTGCCTCTACCTCCCAAAGTGCTGGGATTACAGGCATGAGCTGCCGCGCCCAGCCAAATT
CATTTATTACACTAGAATGTTGATACATTGAATATGACTTGAAACATATAAGGGTTTTACTGAAATTTGGG
AACTCCTATTATAGAGAGATTTATAGTAAGATTTGACCTACTTGGTCTGATTGTATGAGAGATTCTGGAC
TAAGTTATACGTAGTGCTGAACTCCACAGATGAGCAAGTTGATATCCAGTCATAGAAATTCTGTGGGCA
GAAAGGGGAAAGAACTAGTCTTTGGTCCTCACAAACCAAATTACACATAAGTTGATTTTGCACAGAGAT
GTTTTGATCATGATACATGGTTCCATGGTTTAAGTCACCGCCTTTCCTCGACCTTCTGTATACATTAACA
AGCCAGCCATTACTGTCCTCGTCTTTATTCCCTACTGAAACTTGACCAATACTGTTCTCCTTTTTTGTTGT
TATGCCAAGGAAAGACAGTATTTATATGGGAATTTATATAGCTGGTCTGTACCTAGTATTAGCAAGATCC
TCCTTTCTCACTATGTAACTTTCGTACTTTTATTTCCTCAGACTACCATCTGCCATGATTTATTTTACACTG
CTGTTACCTACTTTTACTAACTTTTTTGGTTATTTTTGTTTTCCTTATTCTTTACATTTCCCACAGTCTTTGC
CTTCCCATTTTAATCCTTTGCTCCTCCATTCATTCACTGCTGCATTCCCTTACATATTTCCCCTTTGCTTT
GTCTTTTTCCAGCCTCCCCCACTGGCAAGACTCCCACGAGCTCTGGAGTAAGAAACGCCGACGTCAGC
GACAAAGTGGTGTTGTAGTCGAAGAGCCACCTCCATCCAAAACTTCTCGAAAAGAAAC

Supplementary Tables

Supplementary Table 1: Intracellular KiNATiv™ profiling assay identifies CDK12 and CDK13 as major intracellular targets of THZ531.

Nature Chemical Biology: doi:10.1038/nchembio.2166

Jurkat cells treated with indicated compounds at 1μM

Kinase	Reference	Sequence	Labeling Site	THZ-5-31-1 1μM	THZ-5-31-1-R 1μM
CDK12	UniRef100_Q9NYV4	DIKCSNILLNNSGQIK	Lys2	88.1	8.1
CDK13	UniRef100_Q14004	DIKCSNILLNNR	Lys2	83.9	-5.6
AurA	UniRef100_O14965	DIKPENLLLGSAGELK	Lys2	54.4	-13.5
PLK1	UniRef100_P53350	DLKLGNLFLNEDLEVK	Lys2	53.3	16.1
AurA	UniRef100_O14965	FILALKVLFK	Lys1	51.7	-2.9
CLK1	UniRef100_P49759	LTHTDLKPENILFVQSDYTEAYNPK	Lys2	42.6	4.5
JNK1, JNK2, JNK3	UniRef100_P45983, UniRef100_P53779, UniRef100_P45984	DLKPSNIVVK	Lys2	42.3	2.6
PLK1	UniRef100_P53350	CFEISDADTKEVFAGKIVPK	Lys1	38.9	3
NEK4	UniRef100_P51957	DLKTQNVFLTR	Lys2	38.1	33.3
CDK7	UniRef100_P50613	DKNTNQIVAIKK	Lys1	34	7
PCTAIRE2	UniRef100_Q00537	DLKPQNLLINEK	Lys2	32.9	-7.3
CDK7	UniRef100_P50613	DLKPNNLLLDENGVLK	Lys2	32.8	6.1
SLK	UniRef100_Q9H2G2	DLKAGNILFTLDGDIK	Lys2	30.6	25.4
CHK1	UniRef100_B4DT73	LSKGDDGLEFK	Protein Kinase Domain	25.3	6.3
MAP3K5	UniRef100_Q99683	DIKGDNVLINTYSGVLK	Lys2	24.8	17.8
IRAK4	UniRef100_Q9NWZ3	DIKSANILLDEAFTAK	Lys2	24.2	18.6
TAO1, TAO3	UniRef100_Q7L7X3, UniRef100_Q9H2K8	DIKAGNILLTEPGQVK	Lys2	23.3	32.5
GSK3A	UniRef100_P49840	DIKPQNLLVDPDTAVLK	Lys2	22.9	8.9
MAP3K2, MAP3K3	UniRef100_Q9Y2U5, UniRef100_Q99759	DIKGANILR	Lys2	22.5	18.2
AurA, AurB, AurC	UniRef100_O14965, UniRef100_Q9UQB9, UniRef100_Q96GD4	GKFGNVYLAR	ATP Loop	22.2	-16.5
GSK3B	UniRef100_P49841	DIKPQNLLIDPDTAVLK	Lys2	22	3.3
OSR1	UniRef100_O9JG9, UniRef100_O95747	DVKAGNILLGEDGSVQIADFGVSAFLATGGDITR	Lys2	21.1	20.5
SNRK	UniRef100_Q9NRH2	DLKPENVVFFEK	Lys2	19.5	-3.9
MST4, YSK1	UniRef100_O00506, UniRef100_Q9P289	DIKAANVLLSEQGDVK	Lys2	19.4	32
CDK9	UniRef100_P50750	DMKAANVLITR	Lys2	19.3	17.9
MAP2K3	UniRef100_P46734	DVKPSNVLINK	Lys2	19.3	6.6
MST3	UniRef100_Q9Y6E0	DIKAANVLLSEHGEVK	Lys2	18.9	7.1
GCK	UniRef100_Q12851	DIKGANILLTLQGDVK	Lys2	18.7	35.7
MAP3K4	UniRef100_Q9Y6R4	DIKGANIFLTSEGLIK	Lys2	18.6	37.3
PRPK	UniRef100_Q96S44	FLSGLELVKQGAEAR	ATP Loop	17.4	14.4
KHS1	UniRef100_Q9Y4K4	DIKGANILLTDHGDVK	Lys2	17	14
MAST3	UniRef100_O60307	DLKPDNLLITSLGHIK	Lys2	17	-29
HPK1	UniRef100_Q92918	DIKGANILINDAGEVKR	Lys2	16.7	27
IKKb	UniRef100_O14920	DLKPENIVLQQSGEQR	Lys2	16.5	4.5
LOK	UniRef100_Q94804	DLKAGNVLMTLEGDIR	Lys2	16.5	13.4
CHK2	UniRef100_O96017	DLKPENVLLSSQEEDCLIK	Lys2	16.4	-8
ITK	UniRef100_Q08881	FVLDDQYTSSTGTKFPVK	Activation Loop	16.1	-0.6
Erk2	UniRef100_P28482	DLKPSNLLNTTCDLK	Lys2	16	8.9
PI4KB	UniRef100_Q9UBF8	VPHTQAVVLNSKDK	ATP	15.9	-10.8
IKKa	UniRef100_O15111	DLKPENIVLQDVGGK	Lys2	15.6	4.8
MAP3K1	UniRef100_Q13233	DVKGANLLIDSTGQR	Lys2	15.2	20.9
MST1, MST2	UniRef100_Q13188, UniRef100_Q13043	DIKAGNILLNTEGHAK	Lys2	14.9	7.2
CaMK2a, CaMK2b, CaMK2d,	UniRef100_Q53H78, UniRef100_Q13557, UniRef100_Q13555, UniRef100_Q9UQM7	DLKPENLLLASK	Lys2	14.3	8.4
Erk1	UniRef100_P27361	DLKPSNLLNTTCDLK	Lys2	14.3	6
TAO2	UniRef100_Q9UL54	DVKAGNILLSHSGLVK	Lys2	14.3	27.7
GPRK5	UniRef100_P34947	DLKPENILLDDYGHIR	Lys2	14.2	0.7
RSK2 domain2	UniRef100_P51812	DLKPSNILYVDESGNPESIR	Lys2	14	-1.8
p38a	UniRef100_Q16539	DLKPSNLAVNEDCELK	Lys2	13.9	3.3
PIP5K3	UniRef100_Q9Y2I7	GGKSGAAFYATEDDRFILK	ATP	13.5	3.7
CDK11, CDK8	UniRef100_P49336, UniRef100_Q9BWU1	DLKPANILVMGEGPER	Lys2	13.3	-4.2
NLK	UniRef100_Q9UBE8	DIKPGNLLVNSNCVLK	Lys2	13.2	-1.7
MARK1, MARK2	UniRef100_Q7KZI7, UniRef100_Q9P0L2	EVAVKIDK	Lys1	12.5	-2.9
ZC1/HGK, ZC2/TNIK, ZC3/MI	UniRef100_O95819, UniRef100_Q9UKE5, UniRef100_Q8N4C8	DIKGQNVLLTENAEVK	Lys2	12.8	15
MARK2, MARK3	UniRef100_P27448, UniRef100_Q7KZI7	DLKAENLLLDADMNIK	Lys2	12.3	-11.7
PKN1	UniRef100_Q16512	VLLSERPSGELFAIKALK	Lys1	11.7	0.9
PKN2	UniRef100_Q16513	DLKLDNLLLDTEGFVK	Lys2	11.7	-2.5
MAP2K4	UniRef100_P45985	DIKPSNILLDR	Lys2	11.3	-7.7
MSK1 domain1	UniRef100_O75582	DIKLENILLDSNGHVVLTDFGLSK	Lys2	11.1	0.3
MST2	UniRef100_Q13188	ESGQVVAIKQVPVESDLQEIIK	Lys1	11	-9.8
PYK2	UniRef100_Q14289	YIEDEDYYKASVTR	Activation Loop	10.5	7.8
Erk5	UniRef100_Q13164	DLKPSNLLVNENCELK	Lys2	10	4.6
RSK1 domain2	UniRef100_Q15418	DLKPSNILYVDESGNPECLR	Lys2	10	-8.4
SGK3	UniRef100_Q96BR1	IVYRDLKPENILLDSVGHVVLTDFGLCK	Lys2	9.7	12.4
STLK3	UniRef100_Q9UEW8	DLKAGNILLGEDGSVQIADFGVSAFLATGGDVTR	Lys2	9.7	6.6
CaMK1d	UniRef100_Q8IU85	DLKPENLLYYSQDEESK	Lys2	9.6	-10
CCRK	UniRef100_Q8IZL9	DLKPANLLISASGQLK	Lys2	9.3	7.3
NEK8	UniRef100_Q86SG6	DLKTQNILLDK	Lys2	9.3	6.4
MARK4	UniRef100_Q96L34	DLKAENLLLDAEANIK	Lys2	9	-5.6
MST1	UniRef100_Q13043	ETGQIVAIKQVPVESDLQEIIK	Lys1	8.8	-6.9
p38a	UniRef100_Q16539	QELNKTIWEVPER	Protein Kinase Domain	8.8	-1.1
ROCK1, ROCK2	UniRef100_O75116, UniRef100_Q13464	DVKPDNMLLDK	Lys2	8.8	-8.3
CDK10	UniRef100_Q15131	DLKVSNLLMTDK	Lys2	8.7	-0.7
ITK	UniRef100_Q08881	VAIKTIR	Lys1	8.5	-6.7
GCN2	UniRef100_Q9P2K8	LDGCCYAVKR	Lys1	8.2	-0.8
PKD2	UniRef100_Q9BZL6	DVAVKVIDK	Lys1	8.2	-9
RSK1 domain1, RSK2 domain	UniRef100_P51812, UniRef100_Q15418, UniRef100_Q15349	DLKPENILLDEEGHIK	Lys2	8.1	-13.2
JAK1 domain2	UniRef100_P23458	IGDFGLTKAIETDKEYYTVK	Activation Loop	8	-16.6
PCTAIRE2, PCTAIRE3	UniRef100_Q00537, UniRef100_Q07002	SKLTENLVALKEIR	Lys1	8	-1
NEK3	UniRef100_P51956	SKNIFLTQNVLK	Activation Loop	7.9	-2.7
FER	UniRef100_P16591	QEDGGVYSSSGLKQIPIK	Activation Loop	7.8	6.7
CHK1	UniRef100_B4DT73	DIKPENLLLDER	Lys2	7.5	-28
DNAPK	UniRef100_P78527	KGGSWIQEINVAEK	ATP	7.4	6.3
DRAK1	UniRef100_Q9UEE5	DVKVILDLKPQNILLTSESPLGDIK	Lys2	7.3	15.5
SNRK	UniRef100_Q9NRH2	VAVKVIDK	Lys1	7.1	-1
JNK2	UniRef100_P45984	YQQLKPIGSGAQGIVCAAFDTVLGINVAVKK	Lys1	6.8	11
PKN1	UniRef100_Q16512	DLKLDNLLLDTEGYVK	Lys2	6.7	-5
SRPK1, SRPK2	UniRef100_P78362, UniRef100_Q96SB4	FVAMKVVK	Lys1	6.6	-28.4
LKB1	UniRef100_Q15831	DIKPGNLLLTTGGTLK	Lys2	6.5	6.9
RSKL1	UniRef100_Q96S38	VLGVIDKVLLVMDTR	ATP	6.5	-3.6
MLK1	UniRef100_P80192	DLKSSNILILQK	Lys2	6.2	10.3
CDK6	UniRef100_Q00534	DLKPQNILVTSSGQIK	Lys2	5.7	0.7
MAP3K2	UniRef100_Q9Y2U5	ELAVKQVQFDPDSPETSKEVNALECEIQLLK	Lys1	5.5	-11.4
LATS1	UniRef100_O95835	ALYATKTLR	Lys1	5.4	-27.3
MAP2K3	UniRef100_P46734	HAQSGTIMAVKR	Lys1	5.2	7
NDR2	UniRef100_Q9Y2H1	DIKPDNLLLDAK	Lys2	5	-4.5
eEF2K	UniRef100_O00418	YIKYNSNSGFVR	ATP	4.9	19.5
MAP2K4	UniRef100_P45985	MVHKPSGQIMAVKR	Lys1	4.8	-5.2
SGK3	UniRef100_Q96BR1	FYAVKVLQK	Lys1	4.8	1
DGKH	UniRef100_Q86XP1	ATFSFCVSPLLVFVNSKSGDNQGVK	ATP	4.7	4.6
ULK1	UniRef100_O75385	DLKPQNILLSNPAGR	Lys2	4.5	-20.4
MAP2K6	UniRef100_P52564	DVKPSNVLINALGQVK	Lys2	4.4	1.8
MAP2K6	UniRef100_P52564	HVPSGQIMAVKR	Lys1	4.3	-7.3
NDR1	UniRef100_Q15208	DIKPDNLLLDAR	Lys2	4	-7.8
ILK	UniRef100_Q13418	ISMADVKFSFQCPGR	Protein Kinase Domain	3.4	10.1
TAK1	UniRef100_O43318	DLKPPNLLLVAGGTVLK	Lys2	3.3	-12.3
EphB2	UniRef100_P29323	FLEDDTSDPTYTSALGGKIPIR	Activation Loop	3.2	5.5
NEK6, NEK7	UniRef100_Q8TDX7, UniRef100_Q9HC98	DIKPANVFITATGVVK	Lys2	3.2	-6.2
MAP3K15, MAP3K5, MAP3K	UniRef100_Q99683, UniRef100_O95382, UniRef100_Q6ZN16	IAIKEIPER	Lys1	3.1	3.6
TLK1	UniRef100_Q9UKI8	YLNEIKPPIHYDLKPGNILLVDGTACGEIK	Lys2	3.1	-8.9
GCN2	UniRef100_Q9P2K8	DLKPVNIFLDSDDHVK	Lys2	3	-2.7
FGR	UniRef100_P09769	LIKDDEYNPDEGAKSFPIK	Activation Loop	2.9	4.6
PKCi	UniRef100_P41743	DLKLDNVLLDSEGHIK	Lys2	2.8	-11.3
MAP2K5	UniRef100_Q13163	DVKPSNMLVNTR	Lys2	2.7	5.2
MAP2K1, MAP2K2	UniRef100_P36507, UniRef100_Q02750	DVKPSNILVNSR	Lys2	2	-2
MAP2K7	UniRef100_O14733	DVKPSNILLDER	Lys2	1.8	-9.6
RSK2 domain1	UniRef100_P51812	LTDFGLSKESIDHEKK	Activation Loop	1.7	2.8
CLK3	UniRef100_P49761	YEIVGNLGEGTFGKVVECLDHAR	ATP Loop	1.3	1.7
IKKb	UniRef100_O14920	WHNQETGEQAAIKQCR	Lys1	1.2	-1.3
PAN3	UniRef100_Q58A45	VMDPTKILTGK	ATP	1.1	-6.9
PITSLRE	UniRef100_P21127	DLKTSNLLLSHAGILK	Lys2	1	-1.9
ZAP70	UniRef100_P43403	ISDFGLSKALGADDSYYTAR	Activation Loop	0.8	-6.4
CK1a	UniRef100_P48729	DIKPDNFLMGIGR	Lys2	0.7	-4.2
NEK9	UniRef100_Q8TD19	DIKTLNIFLTK	Lys2	0.5	-2.2
PIK3CD	UniRef100_O00329	TKVNWLAHNVSKDNRQ	ATP	0.5	-5.3
PI4KB	UniRef100_Q9UBF8	LLSVIVKCGDDLRQELLAFQVLK	ATP	0.3	7.3
RSK2 domain1	UniRef100_P51812	VLGQGSFGKVFLVK	ATP Loop	0.3	13.9
CK2a2	UniRef100_P19784	DVKPHNVMIDHDQK	Lys2	0.2	-7.9
PKR	UniRef100_P19525	DLKPSNIFLVDTK	Lys2	0.2	5.1
AMPKa1, AMPKa2	UniRef100_P54646, UniRef100_Q96E92	DLKPENVLLDAHMNAK	Lys2	0.1	-2.6
MAP3K3	UniRef100_Q99759	ELASKQVQFDPDSPETSKEVSALECEIQLLK	Lys1	0.1	5.7
MPSK1	UniRef100_O75716	DLKPTNILLGDEGDPVLMDLGSMNQACIHVEGSR	Lys2	0.1	-1.7
CDK5	UniRef100_Q00535	DLKPQNLLINR	Lys2	0	-1.4
CSK	UniRef100_P41240	VAVKCIK	Lys1	-0.1	-9
CDK5	UniRef100_Q00535	NRETHEIVALKR	Lys1	-0.3	1.2
MASTL	UniRef100_Q96GX5	GAFGKVYLGQK	ATP Loop	-0.3	-2.7
MARK3, MARK4	UniRef100_Q96L34, UniRef100_P27448	EVAIKIIDK	Lys1	-0.4	-7.1
TLK1	UniRef100_Q9UKI8	YAAVKHIQLNK	Lys1	-0.6	6.4
JAK1 domain2	UniRef100_P23458	YDPEGDNTGEQVAVKSLKPESGGNHIADLKK	Lys1	-1	0.6
p70S6K, p70S6Kb	UniRef100_Q9UBS0, UniRef100_P23443	GGYGKVFQVR	ATP Loop	-1	8.4
TLK2	UniRef100_Q86UE8	YVAVKIHQLNK	Lys1	-1.1	4.8
ULK3	UniRef100_D3DW67	EVVAIKCVAK	Lys1	-1.3	-2.1
MLK3	UniRef100_Q16584	DLKSNNILLQPIESDDMEHK	Lys2	-1.5	-23.4
PEK	UniRef100_Q9NZJ5	DLKPSNIFFTMDDVVK	Lys2	-1.6	-1.3
TLK2	UniRef100_Q86UE8	YLNEIKPPIHYDLKPGNILLVNGTACGEIK	Lys2	-1.7	-7.1
MST4	UniRef100_Q9P289	TQQVVAIKQVLEAEDEIEDIQQEITVLSQCDSSYVTK	Lys1	-1.9	-22.2
PKD3	UniRef100_O94806	NIVHCDLKPENVLLASAEPFPQVK	Lys2	-2.5	4.2
CaMK2g	UniRef100_Q13555	TSTQEYAAKIINTK	Lys1	-2.6	14.3
CSK	UniRef100_P41240	VSDFGLTKEASSTQDTGKLPVK	Activation Loop	-2.7	-2.5
CK1d, CK1e	UniRef100_P49674, UniRef100_P48730	DVKPDNFLMGLGKK	Lys2	-2.9	-7.1
IKKe, TBK1	UniRef100_Q14164, UniRef100_Q9UHD2	DIKPGNIMR	Lys2	-3.2	-17.5
IRAK4	UniRef100_Q9NWZ3	GYVNNTTVAVKK	Lys1	-3.2	-1
LATS1	UniRef100_O95835	DIKPDNILIDR	Lys2	-3.6	-31.5
PRP4	UniRef100_Q13523	AAGIKQDFSWIFPNLR	Protein Kinase Domain	-3.8	7.2
p38d, p38g	UniRef100_O15264, UniRef100_P53778	DLKPGNLAVNEDCELK	Lys2	-4.1	-5.9
GCK	UniRef100_Q12851	DTVTSELAAVKIVK	Lys1	-4.4	3.5
SLK	UniRef100_Q9H2G2	AQNKETSVLAAAKVIDTK	Lys1	-4.4	-5.7
CaMK1d	UniRef100_Q8IU85	LFAVKCIPK	Lys1	-4.5	-5.2
PCTAIRE1, PCTAIRE3	UniRef100_Q00536, UniRef100_Q07002	DLKPQNLLINER	Lys2	-4.6	2.5
MASTL	UniRef100_Q96GX5	LYAVKVVK	Lys1	-4.7	-7.3
JAK1 domain1	UniRef100_P23458	QLASALSYLEDKDLVHGNVCTKNLLLAR	Protein Kinase Domain	-5.1	3
AKT2, AKT3	UniRef100_Q9Y243, UniRef100_P31751	GTFGKVILVR	ATP Loop	-5.2	3.9
CaMKK2	UniRef100_Q96RR4	LAYNENDNTYYAMKVLSK	Lys1	-5.4	-2.2
p70S6K	UniRef100_P23443	DLKPENIMLNHQGHVK	Lys2	-5.4	5.2
p38b	UniRef100_Q15759	QELNKTVWEVPQR	Protein Kinase Domain	-5.5	-4.1
MARK3	UniRef100_P27448	EVAIKIIDKTQLNPTSLQK	Lys1	-5.7	1.5
KHS2	UniRef100_Q8IVH8	NVNTGELAAIKVIK	Lys1	-5.8	17.4
MARK2	UniRef100_Q7KZI7	EVAVKIIDKTQLNSSSLQK	Lys1	-5.8	6.2
CDC2	UniRef100_Q5H9N4	DLKPQNLLIDDK	Lys2	-5.9	-8.7
STLK5	UniRef100_Q7RTN6	YSVKVLPWLSPEVLQQNLQGYDAK	Activation Loop	-6	0.8
MLK2	UniRef100_O02779	DLKSINILILEAIENHNLADTVLK	Lys2	-6.1	2.3
LATS2	UniRef100_Q9NRM7	DIKPDNILIDLGHIK	Lys2	-6.3	-33.2
AMPKa1	UniRef100_Q96E92	IGHYILGDTLGVGTFGKVK	ATP Loop	-6.4	2.2

Labeling Site Key

Lys1	Conserved Lysine 1
Lys2	Conserved Lysine 2
ATP Loop	ATP binding loop,
Activation Loop	Activation loop
ATP	ATP site in non-canonical kinase (e.g. lipid kinase)
Protein Kinase Domain	Other lysine within kinase domain, possibly not in ATP binding site
Other	Labeling of residue outside of the protein kinase domain, possibly not in ATP binding site

	>90% Inhibition
	75 - 90% Inhibition
	50 - 75% Inhibition
	35 - 50% Inhibition
	No change
	> 100% increase in MS signal
Not colored	Variable or weak data not reaching significance despite >35% inhibition

Note: This KiNativ dataset is the result of an analysis of duplicate treated samples and either duplicate or quadruplicate control samples. The % changes in MS signals being reported are statistically significant (Student T-test score <0.04). Additionally, the KiNativ method is performed using biological matrices (cells/tissues), which can be inherently variable, followed by a complex procedure of sample preparation and mass spec analysis. Thus, we recommend the use of independent, biological replicates to determine the inter-sample variability of the dataset (i.e., repeating the experimental conditions more than one time and with an independent sample of the biological matrix in order to validate the dataset). In addition, we strongly recommend that customers obtain supporting data through additional independent KiNativ studies or through orthogonal approaches prior to making critical project decisions.

The data reported was performed in a non-GLP manner and was not intended for regulatory submission. It was generated to provide scientific data for informational purposes only.

KiNativ project managers are available to discuss the results and assist customers in understanding the strengths and limitations of this KiNativ dataset. ActivX Biosciences, Inc. bears no responsibility for decisions made by customers based on this KiNativ dataset or for any experimental suggestions offered by KiNativ project managers.

Nature Chemical Biology: doi:10.1038/nchembio.2166

MAP2K2	UniRef100_P36507	HQIMHRDVKPSNILVNSR	Lys2	-6.4	26.4
MSK2 domain1	UniRef100_O75676	DLKLENVLLDSEGHIVLTDFGLSK	Lys2	-6.5	-3.4
NEK1	UniRef100_Q96PY6	DIKSQNIFLTK	Lys2	-6.5	0.4
ZAK	UniRef100_Q9NYL2	WISQDKEVAVKK	Lys1	-6.5	-7.3
BRAF	UniRef100_P15056	DLKSNNIFLHEDLTVK	Lys2	-6.7	-6.4
AMPKa1, AMPKa2	UniRef100_P54646, UniRef100_Q96E92	VAVKILNR	Lys1	-7.1	-0.1
FER	UniRef100_P16591	TSVAVKTCKEDLPQELK	Lys1	-7.2	3.4
PIK3CD	UniRef100_O00329	VNWLAHNVSKDNRQ	ATP	-7.9	-5.6
Wnk1, Wnk2	UniRef100_Q9Y3S1, UniRef100_D3DUP1	GSFKTVYK	ATP Loop	-8.2	-18.2
SMG1	UniRef100_Q96Q15	SYPYLFKGLEDLHLDER	ATP	-9	-0.1
HPK1	UniRef100_Q92918	DKVSGDLVALKMVK	Lys1	-9.1	3.5
ATR	UniRef100_Q13535	FYIMMCKPK	ATP	-9.2	-8.7
PDK1	UniRef100_O15530	EYAIKILEK	Lys1	-9.4	-5
STLK6	UniRef100_Q9C0K7	SIKASHILISGDGLVTLSGLSHLHSLVK	Lys2	-9.7	-1.9
SRPK1	UniRef100_Q96SB4	IIHTDIKPENILLSVNEQYIR	Lys2	-9.8	5.5
IRAK1	UniRef100_P51617	AIQFLHQDSPSLIHGDIKSSNVLLDER	Lys2	-10.3	-16.7
MLK3	UniRef100_Q16584	GELVAVKAAR	Lys1	-11.1	-1.1
NEK7	UniRef100_Q8TDX7	AACLLDGVPVALKK	Lys1	-11.1	13.7
PKCa, PKCb	UniRef100_P05771, UniRef100_P17252	DLKLDNVMLDSEGHIK	Lys2	-11.1	-20.1
CaMK4	UniRef100_Q16566	DLKPENILVATPAPDAPLK	Lys2	-11.4	-14.9
RSK1 domain1	UniRef100_Q15418	LTDFGLSKEAIDHEKK	Activation Loop	-11.5	-5.2
PKD3	UniRef100_O94806	DVAIKVIDK	Lys1	-11.6	-2.8
AMPKa1	UniRef100_Q96E92	VGKHELTGHKVAVKILNR	Lys1	-11.8	-0.5
CDK2	UniRef100_P24941	DLKPQNLLINTEGAIK	Lys2	-11.8	2.1
ABL, ARG	UniRef100_P00519, UniRef100_P42684	LMTGDTYTAHAGAKFPIK	Activation Loop	-12	-14.2
GPRK6	UniRef100_P43250	DLKPENILLDDHGHIR	Lys2	-12	-28.3
NDR1	UniRef100_Q15208	DTGHVYAMKILR	Lys1	-12	-13.6
NEK9	UniRef100_Q8TD19	RTEDDSLVVVKEVDLTR	Lys1	-12.1	-6.1
PIP4K2C	UniRef100_Q8TBX8	VKELPTLKDMDFLNK	ATP	-12.1	-9.8
MST3	UniRef100_Q9Y6E0	VVAIKIIDLEEAEDEIEDIQQEITVLSQCDSPYVTK	Lys1	-12.3	-21.7
PKD1, PKD2	UniRef100_Q9BZL6, UniRef100_Q15139	NIVHCDLKPENVLLASADPFPQVK	Lys2	-12.8	1.9
CK2a1	UniRef100_P68400	GGPNIITLADIVKDPVSR	Protein Kinase Domain	-13.1	-5.3
ILK	UniRef100_Q13418	WQGNDIVVKVLK	Lys1	-13.3	6.6
PHKg2	UniRef100_P15735	ATGHEFAVKIMEVTAER	Lys1	-13.5	-0.9
MARK4	UniRef100_Q96L34	EVAIKIIDKTQLNPSSLQK	Lys1	-13.6	5.9
PKCi	UniRef100_P41743	IYAMKVVK	Lys1	-13.8	-21.3
CHK2	UniRef100_O96017	VAIKIISK	Lys1	-13.9	1.6
PCTAIRE1	UniRef100_Q00536	SKLTDNLVALKEIR	Lys1	-13.9	0.6
PI4KA, PI4KAP2	UniRef100_A4QPH2, UniRef100_P42356	SGTPMQSAAKAPYLAK	ATP	-14.4	8.8
MAP2K1	UniRef100_Q02750	IMHRDVKPSNILVNSR	Lys2	-14.5	19.3
p70S6Kb	UniRef100_Q9UBS0	DLKPENIMLSSQGHIK	Lys2	-14.6	-7.1
FRAP	UniRef100_P42345	IQSIAPSLQVITSKQRPR	ATP	-14.9	-11.6
ZAP70	UniRef100_P43403	QIDVAIKVLK	Lys1	-14.9	-1
MPSK1	UniRef100_O75716	LGEGGFSYVDLVEGLHDGHFYALKR	Lys1	-15	-12.5
FAK	UniRef100_Q05397	CIGEGQFGDVHQGIYMSPENPALAVAIKTCK	Lys1	-15.5	8.4
PRP4	UniRef100_Q13523	CNILHADIKPDNILVNESK	Lys2	-15.8	-2.4
CaMKK2	UniRef100_Q96RR4	DIKPSNLLVGEDGHIK	Lys2	-16.1	-8
AKT1	UniRef100_P31749	GTFGKVILVK	ATP Loop	-16.2	-1.2
MAPKAPK3	UniRef100_Q16644	QVLGLGVNGKVLECFHR	ATP Loop	-16.2	-6.8
DGKA	UniRef100_P23743	IDPVPNTHPLLVFVNPKSGGK	ATP	-16.4	-10.1
NDR2	UniRef100_Q9Y2H1	DTGHYAMKILR	Lys1	-16.9	-5.4
KHS1	UniRef100_Q9Y4K4	NVHTGELAAVKIIK	Lys1	-17.2	-4.3
LOK	UniRef100_O94804	NKETGALAAAKVIETK	Lys1	-17.2	-1.2
RAF1	UniRef100_P04049	DMKSNNIFLHEGLTVK	Lys2	-18.1	-11.8
CDC2	UniRef100_Q5HYN4	DLKPQNLLIDDKGTIK	Lys2	-18.4	-5.9
DNAPK	UniRef100_P78527	EHPFLVKGGEDLR	ATP	-18.6	-14.2
SMG1	UniRef100_Q96Q15	DTVTIHSVGGTITILPTKTKPK	ATP	-18.6	4.8
ULK3	UniRef100_D3DW67	NISHLDLKPQNILLSSLEKPHLK	Lys2	-18.7	10.3
YSK1	UniRef100_O00506	EVSAIKIDLEEAEDEIEDIQQEITVLSQCDSPYITR	Lys1	-19.6	-12.3
AurB	UniRef100_Q96GD4	SHFIVALKVLFK	Lys1	-19.7	-16.7
PIP4K2C	UniRef100_Q8TBX8	TLVIKEVSSEDIADMHSNLSNYHQYIVK	ATP	-20.2	-14.7
CK1g2	UniRef100_P78368	DVKPENFLVGRPGTK	Lys2	-20.3	-0.3
Wnk1, Wnk2, Wnk4	UniRef100_Q96J92, UniRef100_Q9Y3S1, UniRef100_D3DUP1	IGDLGLATLKR	Activation Loop	-21.3	-16.1
MARK2	UniRef100_Q7KZI7	FIVHRDLKAENLLLDADMNIK	Lys2	-21.6	-3.8
Wnk1, Wnk2, Wnk3	UniRef100_Q9Y3S1, UniRef100_D3DUP1, UniRef100_Q9BYP7	DLKCDNIFITGPTGSVK	Lys2	-22.8	-6.2
ABL, ARG	UniRef100_P00519, UniRef100_P42684	YSLTVAVKTLKEDTMEVEEFLK	Lys1	-22.9	-5.1
QSK	UniRef100_Q9Y2K2	VAAKIIDKTQLDEENLKK	Lys1	-22.9	-1.4
PIP4K2B	UniRef100_P78356	AKDLPTFKDNDFLNEGQK	ATP	-23	-2.4
MLKL	UniRef100_Q8NB16	APVAIKVFK	Lys1	-23.3	-6.2
BARK1	UniRef100_P25098	DLKSNNIFLHEGLTVK	Lys2	-23.6	-6.1
TYK2 domain2	UniRef100_P29597	IGDFGLAKAVPEGHEYYR	Activation Loop	-23.6	-37.4
IKKe	UniRef100_Q14164	SGELVAVKVFNTTSYLRPR	Lys1	-24.1	-14.4
PIK3CG	UniRef100_P48736	KKPLWLEFK	ATP	-24.1	7.9
TBK1	UniRef100_Q9UHD2	TGDLFAIKVFNNISFLRPVDVQMR	Lys1	-24.5	-15.5
CaMK2d	UniRef100_Q13557	IPTGQEYAAKIINTKK	Lys1	-24.9	6.5
PIK3C3	UniRef100_Q8NEB9	TEDGGKYPVIFKHGDDLRQDQLILQIISLMDK	ATP	-25.9	-3.5
STLK6	UniRef100_Q9C0K7	HTPTGTLVTIKITNLENCNEER	Lys1	-26.3	-9.5
ARAF	UniRef100_P10398	DLKSNNIFLHEGLTVK	Lys2	-26.4	-0.9
ZC1/HGK	UniRef100_Q95819	TGQLAAIKVMDVTEDEEEEIKLEINMLKK	Lys1	-26.7	9.3
ACK	UniRef100_Q07912	TVSVAVKCLKPDVLSQPEAMDDFIR	Lys1	-26.8	-23.5
RSK1 domain1	UniRef100_Q15418	KVTRPDSGHLYAMKVLK	Lys1	-27.1	-1.6
PIK3C3	UniRef100_Q8NEB9	TEDGGKYPVIFKHGDDLR	ATP	-27.9	-21
MAPKAPK3	UniRef100_Q16644	CALKLLYDSPK	Lys1	-28.8	-0.9
PIP4K2A	UniRef100_P48426	AKELPTLKDNDFINEGQK	ATP	-29.3	-7.5
PKR	UniRef100_P19525	IGDFGLVTSLKNDGKR	Activation Loop	-29.8	-9.3
STLK5	UniRef100_Q7RTN6	SVKASHILISVDGK	Lys2	-30.3	-3.3
CaMK4	UniRef100_Q16566	GTQKPYALKVLK	Lys1	-30.9	-11.6
LCK	UniRef100_P06239	EGAKFPIKWTAPEAINYGTFTIK	Activation Loop	-31.4	-14.4
CDK2	UniRef100_P24941	LTGEVVALKK	Lys1	-32	1.1
IRE1	UniRef100_O75460	DLKPHNILISMPNAHGK	Lys2	-33	-19.7
ITPK1	UniRef100_Q13572	ESIFFNSHNVSKPESSSVLTELDKIEGVFERPSDEVIR	ATP	-34.6	5.7
PKCa	UniRef100_P17252	KGTEELYAIKILK	Lys1	-34.6	-16.2
FYN	UniRef100_P06241	VAIKTLKPGTMSPESFLEEAQIMK	Lys1	-35.7	5
PIK3C2B	UniRef100_O00750	VIFKCGDDLRQDMLTLQMIR	ATP	-36	-26.3
FYN, SRC, YES	UniRef100_P12931, UniRef100_P07947, UniRef100_P06241	QGAKFPIKWTAPEAALYGR	Activation Loop	-42	10.8
ZC2/TNIK	UniRef100_Q9UKE5	TGQLAAIKVMDVTGDEEEEIKQEINMLKK	Lys1	-42.3	-13.9
PKCh	UniRef100_P24723	VKETGDLVAVKVLK	Lys1	-43.7	-13.2
CaMK4	UniRef100_Q16566	IVEHQVLMKTVCGTPGYCAPEILR	Activation Loop	-44.1	-14.8
PIK3CA	UniRef100_P42336	RPLWLNWENPDIMSELLFQNNEIIFKNGDDLRQDMLTLQIIR	ATP	-52.3	-48.4
PIK3CB	UniRef100_P42338	VFGEDSVGVIFKNGDDLRQDMLTLQMLR	ATP	-57.9	-5.5
PAK2	UniRef100_Q13177	IGQSASGTVFTATDVALGQEVAIKQINLQK	Lys1	-61	4

Supplementary Table 2: Table of diffraction data collection and refinement statistics for CDK12-cyclin K and THZ531 co-crystal structure.

Nature Chemical Biology: doi:10.1038/nchembio.2166

Table 2 Data collection and refinement statistics (molecular replacement)

	$CDK12^{715-1052}$/cyclin K^{11-267}
Data collection	
Space group	$P2_1$
Cell dimensions	
a, *b*, *c* (Å)	*49.8, 148.7, 91.6*
α, β, γ (°)	90.0, 93.8, 90.0
Resolution (Å)	41.3-2.7(12.1-2.7)*
R_{merge}	0.071 (0.569)
$I / \sigma I$	12.5 (1.9)
Completeness (%)	98.7 (99.0)
Redundancy	3.0 (3.1)
Refinement	
Resolution (Å)	41.3-2.7
No. reflections	34,240 (2581)
R_{work} / R_{free}	22.1/26.2
No. atoms	
Protein	8993
Ligand/ion	80
B-factors	
Protein	61.1
Ligand/ion	87.0
R.m.s. deviations	
Bond lengths (Å)	0.0072
Bond angles (°)	1.1369

*Single crystal.
Values in parentheses are for highest-resolution shell.

Supplementary Table 3: GEO upload files

Nature Chemical Biology: doi:10.1038/nchembio.2166

ChIP-seq samples	Figure	GEO
Jurkat DMSO Pol II	4a,b,c,d; 5f,h; 6f; SF5; SF7a,b,c; SF9a,b	GSM1850204 (NEW)
Jurkat H3K27ac	4a,b,c,d; 6e,f; SF8e,f; SF9a,b	GSM1296384
Jurkat CDK7	SF8e; SF9a	GSM1296385
Jurkat CDK12	4a,b,c,d; 6f,g; SF9a,b,c	GSM1850203 (NEW)
Jurkat 50nM THZ531 Pol II	5f,h; 6f; SF7a,b	GSM1850205 (NEW)
Jurkat 500nM THZ531 Pol II	5f,h; 6f; SF7a,b; SF9a	GSM1850206 (NEW)
Jurkat Input DNA		GSM1296386
Jurkat Flavo Pol II	SF7b,c	GSM1224787
Jurkat THZ1 Pol II	SF9a	GSM1224785
Jurkat DMSO pSer2 Pol II	5b,f,g; SF7a	GSE72023
Jurkat 50 nM THZ531 pSer2 Pol II	5b,f,g; SF7a	GSE72023
Jurkat 500 nM THZ531 pSer2 Pol II	5b,f,g; SF7a	GSE72023
Expression microarrays	5c,d,g,h; 6a,c,g; SF6b,c; SF7c; S8a,c,f; SF9c	GSE72022
Previous THZ1 data from THZ1 paper (both 50 and 250 nM data)	SF6c	GSM1224822, GSM1224826, GSM1224827, GSM1224818, GSM1224819, GSM1224820, GSM1224821
Previous Flavopiridol data from THZ1 paper (250 nM data)	SF7c	GSM1224823, GSM1224822

Supplementary Data Sets

Supplementary Data Set 1: Mass spectrometry identifies CDK12-cyclin K and CDK13-cyclin K complexes as major targets of bioTHZ531 in Jurkat cell lysates.

See accompanying excel file

Supplementary Data Set 2: *In vitro* Ambit^{TM} binding assay shows THZ531 potently inhibits CDK13.

See accompanying excel file

Supplementary Data Set 3: Gene expression microarray data of THZ531, Flavopiridol, THZ1 – treated cells.

See accompanying excel file

Supplementary Data Set 4: Jurkat enhancers and super –enhancers identified by H3K27Ac ChIP-seq.

See accompanying excel file

Nature Chemical Biology: doi:10.1038/nchembio.2166