

## CELL BIOLOGY

# A novel landscape of nuclear human CDK2 substrates revealed by in situ phosphorylation

Yong Chi<sup>1,2</sup>, John H. Carter<sup>3</sup>, Jherik Swanger<sup>1</sup>, Alexander V. Mazin<sup>4</sup>, Robert L. Moritz<sup>2</sup>, Bruce E. Clurman<sup>1\*</sup>

Cyclin-dependent kinase 2 (CDK2) controls cell division and is central to oncogenic signaling. We used an “in situ” approach to identify CDK2 substrates within nuclei isolated from cells expressing CDK2 engineered to use adenosine 5′-triphosphate analogs. We identified 117 candidate substrates, ~40% of which are known CDK substrates. Previously unknown candidates were validated to be CDK2 substrates, including LSD1, DOT1L, and Rad54. The identification of many chromatin-associated proteins may have been facilitated by labeling conditions that preserved nuclear architecture and physiologic CDK2 regulation by endogenous cyclins. Candidate substrates include proteins that regulate histone modifications, chromatin, transcription, and RNA/DNA metabolism. Many of these proteins also coexist in multi-protein complexes, including epigenetic regulators, that may provide new links between cell division and other cellular processes mediated by CDK2. In situ phosphorylation thus revealed candidate substrates with a high validation rate and should be readily applicable to other nuclear kinases.

## INTRODUCTION

Advanced mass spectrometry (MS) and quantitative phosphoproteomics enable the identification of large sets of protein phosphorylation sites to comprehensively identify protein kinase substrates (1–4). However, despite many advances, determining which phosphorylation sites are direct and physiologic targets of a specific kinase remains challenging. Analog-sensitive kinases (AS-kinases) are an important tool for discovering kinase substrates. AS-kinases contain a mutation in a conserved “gatekeeper” residue that normally functions to restrict active site access to adenosine 5′-triphosphate (ATP; but not other nucleotides) (3, 5–7). Replacing the gatekeeper residue with a smaller amino acid renders AS-kinases capable of using bulky ATP analogs that cannot be used by normal kinases. By using bulky ATP analogs in combination with AS-kinases in various ways, the activity of an AS-kinase can be isolated from that of other cellular kinases.

In vitro applications of AS-kinases use ATP analogs to directly tag substrates with a label that distinguishes them from other cellular phosphoproteins [reviewed in (3, 6)]. However, because ATP analogs cannot enter intact cells, these approaches require in vitro conditions that can be prone to artifacts. For example, the use of highly active recombinant AS-kinases and cell lysates (in which cellular contexts are lost) can produce false-positive identifications that may greatly increase the experimental work needed to validate candidate substrates. In contrast, in vivo applications of AS-kinases typically identify proteins that are differentially phosphorylated in cells treated with highly specific AS-kinase inhibitors, sometimes in the context of AS mutations engineered into endogenous kinases (3, 6). While these methods are not prone to in vitro artifacts, they cannot readily distinguish direct substrates from phosphorylations that may be indirectly regulated by AS-kinase inhibition because substrates are not directly labeled by the AS-kinase.

<sup>1</sup>Divisions of Clinical Research and Human Biology, Fred Hutchinson Cancer Research Center, 1100 Fairview Avenue N. Seattle, WA 98109, USA. <sup>2</sup>Institute for Systems Biology, 401 Terry Avenue N. Seattle, WA 98109, USA. <sup>3</sup>Division of Hematology/Medical Oncology, Oregon Health & Science University School of Medicine, 3181 S.W. Sam Jackson Park Rd. Portland, OR 97239-3098, USA. <sup>4</sup>Department of Biochemistry and Molecular Biology, Drexel University College of Medicine, 245 N 15th Street, Philadelphia, PA 19102-1192, USA.

\*Corresponding author. Email: bclurman@fredhutch.org

Cyclin-dependent kinases (CDKs) are holoenzymes consisting of catalytic (CDK) and regulatory (cyclin) subunits that regulate cellular process by phosphorylating complex substrate networks (8, 9). Many mammalian CDKs—including CDK1, CDK2, CDK5, CDK7, and CDK9—have proven amenable to AS mutations and have been studied using both direct and indirect approaches (10–18). Cdk2 is activated by cyclin E and cyclin A. Cyclin E-Cdk2 regulates cell cycle re-entry, G<sub>1</sub> progression, and S phase entry, whereas cyclin A-Cdk2 acts later in the cell cycle, where it coordinates S phase progression and functions in G<sub>2</sub> and M phase cells. In previous work, we developed an in vitro approach using recombinant cyclin A/AS-CDK2 and thiophosphate labeling to identify >100 candidate cyclin A-CDK2 substrates in human cell lysates (19). The use of a thiophospho-ATP analog allowed a biochemical enrichment-based strategy and prevented ATP analog hydrolysis in cell lysates. Shokat and colleagues (10) also developed a thiophosphate-based method to study AS-CDK1 in vitro, which used alternate thiophosphate chemistry. These studies revealed candidate substrates common to both CDK2 and CDK1 and unique candidates for each kinase. However, the extent to which any of these proteins is predominantly a CDK1 versus CDK2 substrate in vivo is not known.

Because of its crucial roles in normal and neoplastic signaling, we sought to identify high confidence CDK2 substrates using conditions that maintain near-physiologic CDK2 activity and preserve nuclear context and architecture. In this approach, termed “in situ phosphorylation,” we first stably expressed ectopic AS-CDK2 in cells, which is activated by endogenous cyclins and thus subject to near-normal regulation. We then isolated nuclei from cells expressing either wild-type CDK2 (WT-CDK2) or analog-sensitive CDK2 (AS-CDK2) and performed in situ substrate thiophosphorylation by incubating the nuclei with an ATP- $\gamma$ -S analog. These conditions allow substrates to be phosphorylated in conditions that better preserve CDK2’s normal subcellular interactions with its substrates. We subsequently used biochemical enrichment and MS to identify candidate substrates.

We identified ~150 AS-CDK2-specific thiophosphopeptides and 117 candidate CDK2 substrates. Remarkably, ~43% of these proteins are known CDK substrates, indicating that a high proportion of the candidates are physiologic CDK2 substrates. Moreover, we found

Copyright © 2020  
The Authors, some  
rights reserved;  
exclusive licensee  
American Association  
for the Advancement  
of Science. No claim to  
original U.S. Government  
Works. Distributed  
under a Creative  
Commons Attribution  
NonCommercial  
License 4.0 (CC BY-NC).

Downloaded from <http://advances.sciencemag.org/> on November 18, 2020

that each of the candidates we tested was directly phosphorylated by CDK2 in vitro or in vivo, further supporting the idea that candidates identified through in situ phosphorylation contain a high proportion of bona fide CDK2 substrates. Many previously unidentified candidates are chromatin-associated proteins with roles in histone modification and chromatin remodeling; DNA metabolism, damage, and repair; and transcription and RNA metabolism, and whose identification was likely allowed by maintaining the nuclear contexts within which CDK2 normally functions. In summary, in situ conditions led to the efficient and more confident identification of CDK2 substrates. These methods should be readily applicable to other nuclear kinases with complex substrate networks.

## RESULTS

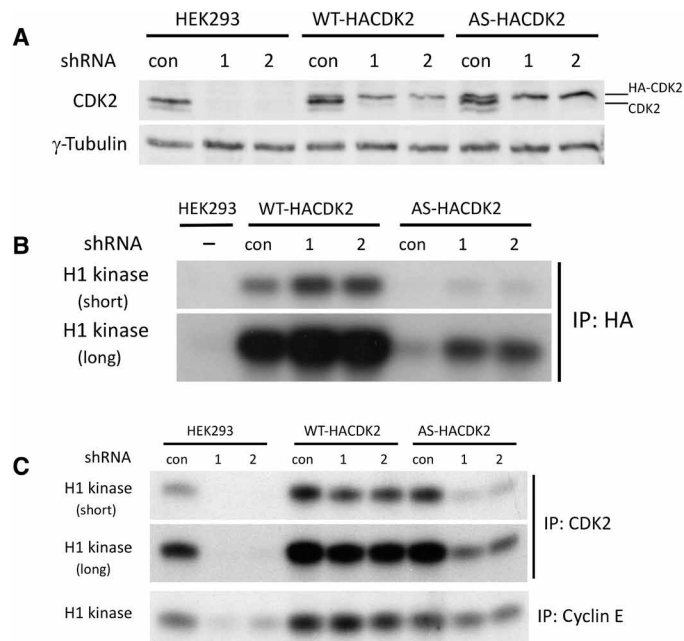
### Developing AS-CDK2 cells and substrate thiophosphorylation in isolated cell nuclei

We used retroviral transduction to generate pools of human embryonic kidney–293 (HEK293) cells that stably express either WT-HACDK2 or AS-HACDK2 [hemagglutinin (HA) tagged] at levels slightly less than endogenous CDK2 (Fig. 1A). A previous study revealed that AS-CDK2 has impaired cyclin A binding that prevents AS-CDK2 from effectively competing with endogenous CDK1 and CDK2 for cyclin A binding (12). Accordingly, AS-CDK2 had less kinase activity than WT-CDK2, as assayed by immunoprecipitation kinase assays using antibodies against the HA-tag (specific to ectopic CDK2), total CDK2, or cyclin E (Fig. 1, B and C). To overcome this defect, we inhibited endogenous CDK2 by using lentiviruses encoding a short hairpin RNA (shRNA) targeting the CDK2 3' untranslated region (3'UTR; which is not present in the retroviral CDK2 complementary DNAs), which silenced endogenous CDK2 to almost undetectable levels (Fig. 1A). As predicted, endogenous CDK2 knockdown increased AS-CDK2 activity, presumably by enhancing cyclin association (Fig. 1, B and C). However, AS-HACDK2 activity remained lower than WT-HACDK2 activity even after endogenous CDK2 knockdown (Fig. 1B). Silencing endogenous CDK2 thus yielded detectable AS-CDK2 activity that was somewhat less than endogenous CDK2 activity, rather than supraphysiologic activity.

Since the cell membrane is impermeable to nucleotides, we isolated nuclei using hypotonic methods to help preserve nuclear integrity and determined that ATP- $\gamma$ -S enters isolated nuclei and that proteins could be readily labeled by incubating nuclei with ATP- $\gamma$ -S (table S1). We then used nuclei isolated from WT- and AS-HACDK2 cells in conjunction with  $N^6$ -(2-phenylethyl)-ATP- $\gamma$ -S (PE-ATP- $\gamma$ -S) to directly label nuclear CDK2 substrates (Fig. 2A). Cells have high concentrations of ATP (>1 mM) that competes with PE-ATP- $\gamma$ -S for kinase occupancy. We thus included  $Mn^{2+}$  in the in situ kinase reaction, which improves thiophosphate labeling by allowing more effective competition with cellular ATP (20). Last, CDK2 substrates that are already highly phosphorylated by CDK2 in vivo may not be additionally labeled by the thiophospho-ATP analog in situ. We thus briefly pretreated cells with a CDK inhibitor (roscovitine) before harvesting and isolating nuclei to inhibit endogenous CDK2 and allow substrate dephosphorylation before labeling in situ (Fig. 2A).

### Thiophosphopeptide recovery and identification of high confidence nuclear CDK2 substrates

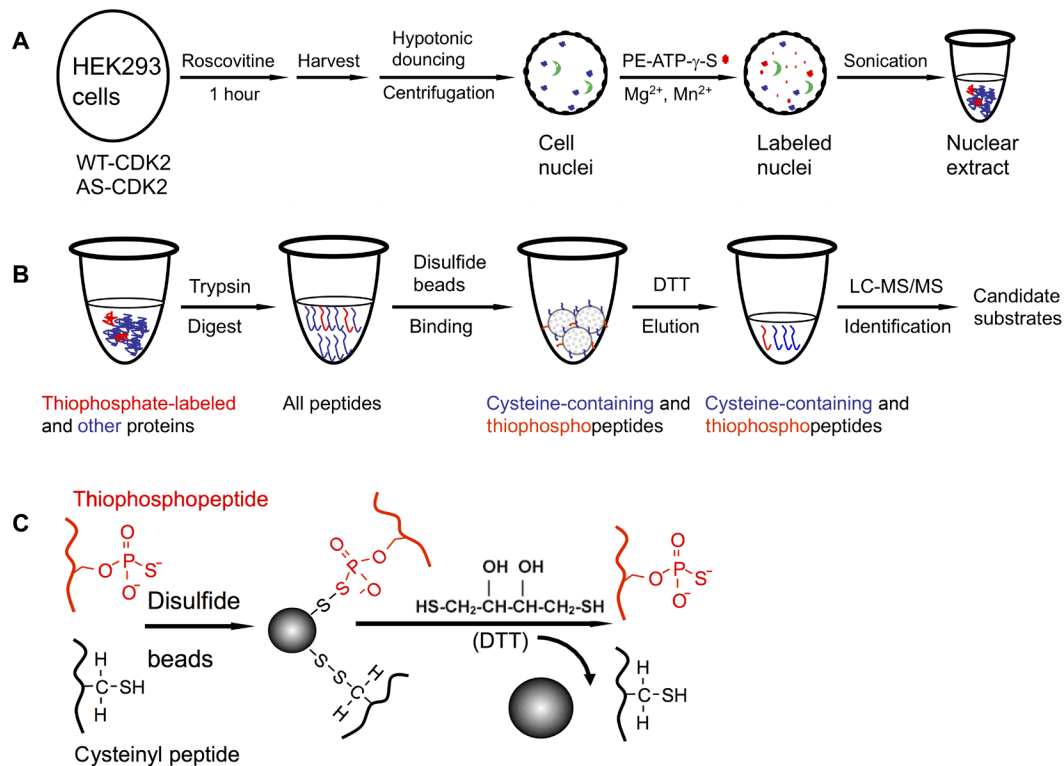
Nuclei were lysed after thiophosphorylation and digested with trypsin, and the resultant tryptic peptides containing either thiophosphate



**Fig. 1. Characterization of AS-CDK2 activity.** (A) HEK293 cells [control (con)] and HEK293 cells stably expressing WT-HACDK2 and AS-HACDK2 were transduced with lentiviral vectors expressing control or one of two different CDK2 shRNAs (#1 and #2) for 2 days. Cell lysates were immunoblotted with anti-CDK2 and anti- $\gamma$ -tubulin (loading control) antibodies. The positions of CDK2 and ectopic HA-CDK2 are indicated. (B and C) Lysates from the above cells were immunoprecipitated (IP) with anti-HA (B), anti-CDK2, and anti-cyclin E antibodies (C), and histone H1 kinase assays were carried out in the presence of  $\gamma$ - $^{32}P$ -ATP.

or cysteine were bound to a disulfide resin (Thiopropyl Sepharose 6B) as previously described (Fig. 2, B and C) (19)). We developed a new method to elute thiophosphopeptides from the disulfide beads with dithiothreitol (DTT), which retains the thiophosphate tag within the eluted peptides (thereby providing a unique mass signature for AS-CDK2 substrates) and allows the elution of cysteine-containing thiophosphopeptides. Although DTT elutes a large excess of peptides that simply contain cysteine, current MS instruments easily handle the additional sample complexity, and these peptides were readily excluded from subsequent analyses. Moreover, the mildly acidic conditions used favor the binding and elution of thiophosphopeptides over cysteine-containing peptides (21).

We performed a large-scale experiment, in which four replicates of cell nuclei from WT- or AS-HACDK2 cells were labeled with PE-ATP- $\gamma$ -S (see Materials and Methods). We identified 4592 unique peptides, of which 464 were unique thiophosphopeptides (about 10% of the total eluted peptides), and almost all (~99%) of the remaining peptides contained cysteine (table S2). Among the 326 unique “stripped” thiophosphopeptides (distinguished only on the basis of primary sequence), 166 thiophosphopeptides were found only in AS-CDK2 samples, 32 only in control WT-CDK2 samples, and 128 in both (table S3). Candidate CDK2 substrates were defined as containing thiophosphopeptides that (i) were present only in AS-CDK2 samples and (ii) have primary sequences that contain proline-directed CDK phosphorylation motifs (serine or threonine followed by a proline, termed “S/T-P”). Applying these criteria, 156 of 166 (93%) of the CDK2-AS-specific thiophosphopeptides had at least one S/T-P site, which corresponds to 117



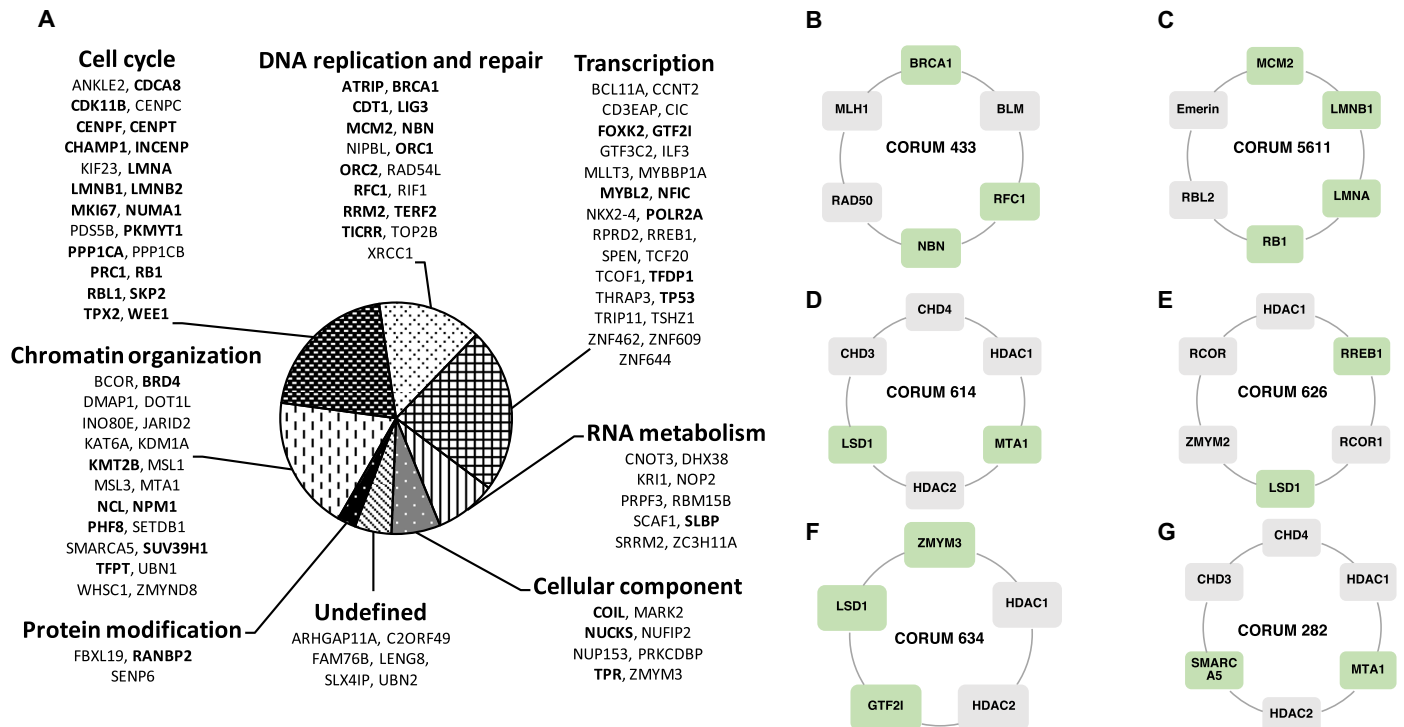
**Fig. 2. Schema for identification of CDK2 substrates using an in situ nuclear labeling assay.** (A) Scheme for in situ labeling. Asynchronously growing HEK293 cells stably expressing WT-CDK2 (AS-CDK2) were pretreated with roscovitine and harvested by trypsinization. Nuclei were isolated as depicted, incubated with PE-ATP- $\gamma$ -S to allow thiphosphate labeling of proteins, and subsequently lysed (see Materials and Methods). (B) Scheme for thiphosphopeptide isolation and identification. Labeled nuclear extracts were digested with trypsin. The resulting peptides were incubated with disulfide beads, which capture both thiphosphopeptides and cysteine-containing peptides. Beads were eluted with dithiothreitol (DTT), and the eluted peptides were subjected to liquid chromatography coupled to tandem MS to identify the thiphosphopeptides. Specific thiphosphopeptide identifications that carried proline-directed phosphorylation motifs and were present in the AS-CDK2 cells but not control WT-CDK2 cells were considered as candidate CDK2 substrates. (C) Schematic of thiphosphopeptide and cysteine-containing peptide elution from disulfide beads by DTT.

unique proteins. The Comet search engine localized the presumptive thiphosphorylation site to be the proline-directed S/T-P residues in 148 of 156 (95%) of the AS-specific S/T-P thiphosphopeptides (tables S3 and S4). Most (94%) of these 117 substrates are nuclear proteins, and only 33 of 117 had been found in our previous *in vitro* studies. About 95% of the phosphorylation sites carried the full CDK consensus motif S/T-P-X-R/K (where X is any amino acid), and ~95% of these phosphorylations have been reported in large-scale *in vivo* phosphoproteome analyses (22–25). We recovered six cysteine-containing thiphosphopeptides representing four unique proteins (tables S3 and S4). Fifty of 117 (43%) candidate substrates are known CDK substrates, indicating a remarkably high proportion of physiologic CDK substrates within the candidate list generated *in situ* (table S4).

In addition to known CDK substrates, we found a novel landscape of candidate CDK2 substrates, most of which function within key nuclear processes, including cell cycle control, histone modification and chromatin organization, DNA replication and repair, transcription, and RNA metabolism (Fig. 3A). The high proportion of candidates associated with chromatin modification, DNA metabolism, and transcription is likely the result of *in situ* conditions that preserve cyclin-CDK2 chromatin associations. For example, we identified 11 proteins that regulate histone modifications as candidate substrates, eight of which are previously unknown (DOT1L, JARID2,

KAT6A, LSD1, MSL1, MSL3, SETDB1, and NSD2) and three of which are known CDK substrates (KMT2B, PHF8, and SUV39H1). We also found candidates with broader roles in chromatin remodeling (e.g., BCOR, BCL11A, DMAP1, INO80E, and SMARCA5) and known CDK substrates with chromatin remodeling functions (e.g., BRD4). Similarly, we identified many proteins that regulate transcription and RNA metabolism, and these include many previously unknown candidate CDK2 substrates (see Discussion).

We further interrogated the relationships between the substrates in this network by searching protein-protein interaction databases. First, we manually curated the known physical interactions between all of the substrates with one another by using BioGRID 3.5 (<https://thebiogrid.org>) (table S5) (26). We found extensive physical interactions between many of the CDK2 substrates, reinforcing the master regulatory roles that CDK2 plays in processes such as cell division and DNA repair and replication. Next, we examined whether and how candidate substrates coexist within known multiprotein complexes by using the Corum database (<https://mips.helmholtz-muenchen.de/corum/>) (27). Many known and candidate substrates coexist within these defined Corum complexes in both expected and unexpected ways, suggesting that CDK2 may globally regulate these complexes through phosphorylating multiple subunits thereof (Fig. 3, B and C, and see Discussion). One particularly notable finding was the co-occurrence of candidate CDK2 substrates in complexes with crucial



**Fig. 3. Functional categorization of CDK2 candidate substrates.** (A) Candidate substrates are grouped based on select gene ontology processes and with manual curation. Boldfaced proteins denote known CDK substrates. (B to G) Examples of CDK2 substrates and candidate substrates (green subunits) co-occurring in known Corum complexes. The Corum complex identification number is shown in the middle of each depicted complex. The following complexes are shown: (B) BASC complex, 6 of 12 components depicted; (C) Emerin complex 24, 6 of 6 components depicted; (D) NuRD complex, 6 of 7 components depicted; (E) LSD1 complex, 6 of 14 components depicted; (F) XFIM complex, 5 of 5 components depicted; and (G) SNF2h-cohesin-NuRD complex, 6 of 16 components depicted.

epigenetic functions, particularly those that contain both LSD1 and the histone deacetylase 1/2 (HDAC1/2) (Fig. 3, D to G, and see Discussion). References for each of these specific complexes can be found on the Corum website.

### In vitro validation of select candidate CDK2 substrates

The high proportion of known CDK substrates found in our screen suggested that many candidates may also be bona fide CDK2 substrates. We thus determined whether a group of nine candidates (representing proteins that were readily immunoprecipitated from transfected cells) were phosphorylated by recombinant cyclin A-CDK2 in vitro. These include (i) NSD2, a SET-domain histone lysine methyltransferase; (ii) DMAP1, a component of the NuA4 histone acetyltransferase complex; (iii) LSD1/KDM1A, a histone H3 Lys-4 demethylase; (iv) DOT1L, a methyltransferase that acts on Lys<sup>79</sup> of histone H3; (v) BCL11A, a transcription factor (TF) associated with chromatin remodeling; (vi) PRPF3, a component of the precatlytic spliceosome; (vii) MSL3, a chromatin remodeling protein and regulator of histone H3 acetylation; and (viii) GTF3C2, a general TF required for RNA polymerase III-mediated transcription. Last, we recently validated Rad54 phosphorylation on both T31 and S59 by CDK2 in vitro and found that S49 phosphorylation inhibits Rad54's branch migration activity (28).

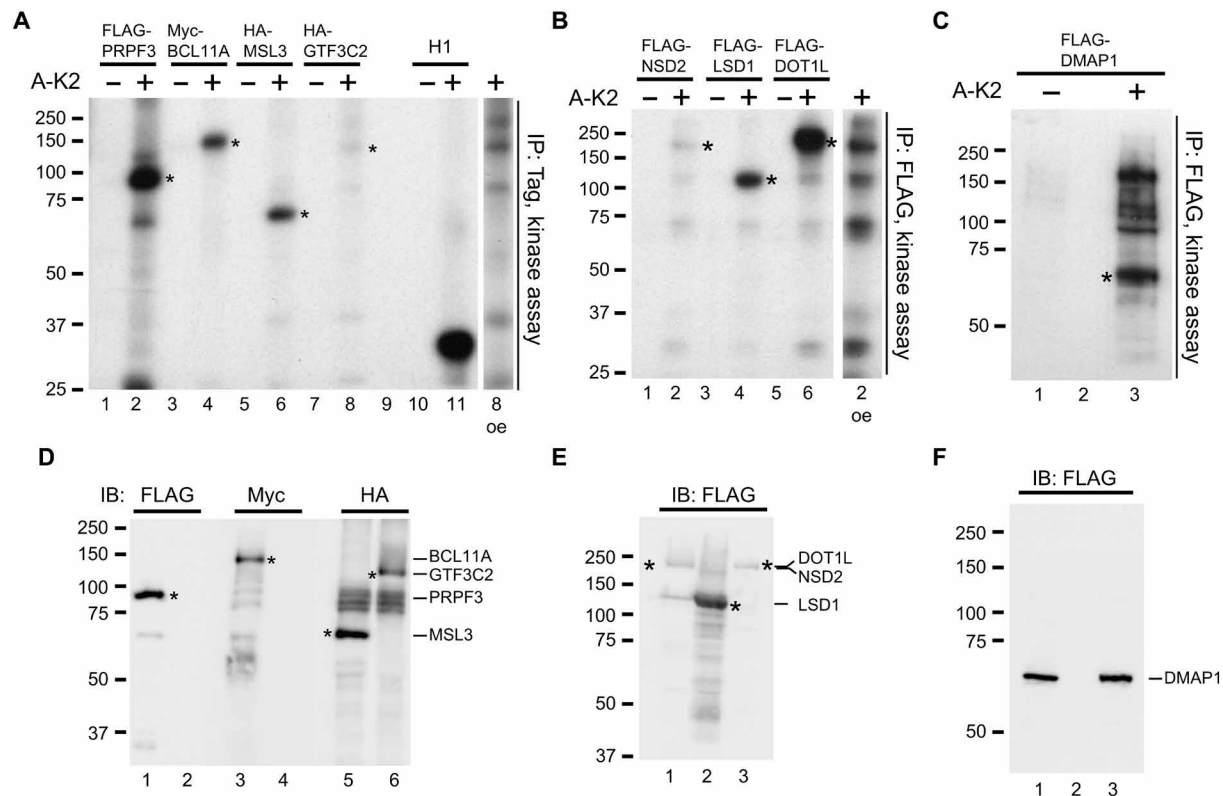
We immunoprecipitated each of these nine proteins from transfected HEK293 cells and tested their ability to be phosphorylated by recombinant cyclin A-CDK2 in vitro using <sup>32</sup>P-ATP (Fig. 4). All

nine of these in situ candidates were validated to be direct CDK2 substrates in vitro. In the case of DMAP1, it was necessary to elute the immunoprecipitated protein for it to be phosphorylated (Fig. 4, C and F), whereas the other substrates were readily phosphorylated while bound to beads (Fig. 4, A and B). Immunoblots of the immunoprecipitated substrates are shown in Fig. 4 (D to F).

### In vivo validation of select candidate CDK2 substrates

We next determined whether any of three of these previously unknown substrates (RAD54, DOT1L, and LSD1) are phosphorylated by CDK2 in vivo by either generating phosphosite-specific antibodies or using commercial phospho-specific antibodies. Because we identified a RAD54 phosphopeptide containing T31 in our substrate screen, we generated an anti-Rad54 antibody that detects phosphorylated T31 in a phosphatase-sensitive manner (fig. S1A). The anti-pT31 antibody also exhibited a small amount of reactivity against unphosphorylated Rad54, which was eliminated when the epitope was deleted (fig. S1B). Endogenous RAD54 was hypophosphorylated on T31 in asynchronous cells, but T31 phosphorylation rapidly increased when cells were treated with the phosphatase inhibitor okadaic acid, suggesting that endogenous Rad54 T31 phosphorylation is highly labile and opposed by a phosphatase (fig. S1C). Phospho-T31 was also readily detectable on ectopic Rad54 in vivo (without okadaic acid treatment); this was reduced by brief CDK inhibition with roscovitine (Fig. 5A) and, conversely, increased by cyclin A-CDK2 overexpression (Fig. 5B). Residual Rad54 T31 phosphorylation in roscovitine-treated



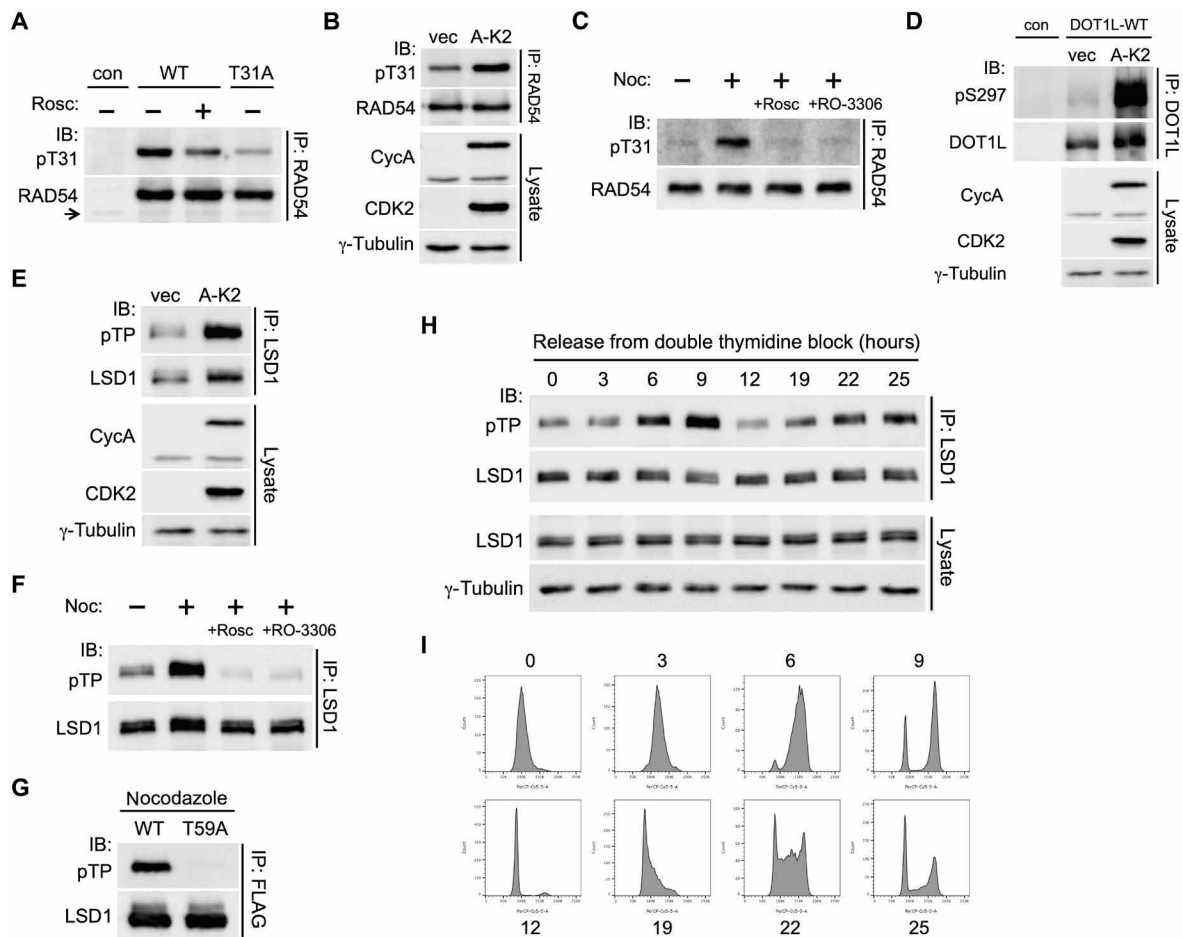


**Fig. 4. Validation of select CDK2 candidate substrates in vitro.** (A) FLAG-tagged PRPF3, Myc-tagged BCL11A, and HA-tagged MSL3 and GTF3C2 were expressed in HEK293 cells and immunoprecipitated from cell lysate using antibodies against the tag. Washed immunoprecipitates were phosphorylated using purified cyclin A–CDK2 (or no kinase control) in the presence of  $^{32}$ P-ATP and detected by autoradiography. Histone H1 kinase assay was used as control. oe, overexpression. Lane 9 was empty. Asterisk (\*) marks the expected band of the substrates. (B) FLAG-tagged NSD2, LSD1, and DOT1L were analyzed as in (A). (C) FLAG-tagged DMAP1 was analyzed as in (A), except that the washed immunoprecipitates were first eluted with a FLAG peptide and fractions of the eluate were used in the kinase assay. Lane 2 was empty. (D and E) Immunoblots (IB) of a fraction of the corresponding immunoprecipitates shown in (A) and (B) using the antibodies against the tag. In (D), lanes 2 and 4 were empty. (F) Immunoblot of the gel in (C) after autoradiography using FLAG antibody.

cells likely reflects either incomplete T31 dephosphorylation during the brief period of CDK2 inhibition or a roscovitine-insensitive kinase that also phosphorylates T31. Because CDK2 directly phosphorylates T31, these data are consistent with the idea that CDK2 directly phosphorylates RAD54 on T31 in vivo, although it is formally possible that another T31 kinase is up-regulated by cyclin A–CDK2 overexpression. Last, because CDK activity is highest in mitosis, we studied endogenous Rad54 in cells arrested in prometaphase with nocodazole (see flow profiles in fig. S1D) and found high levels of T31 phosphorylation that was almost completely inhibited by two CDK inhibitors, roscovitine and RO-3306 (a CDK1 inhibitor) (Fig. 5C). Thus, endogenous RAD54 is phosphorylated in mitosis on T31, likely by CDK2, CDK1, or both. Because we used a pT31-specific antibody, we were unable to assess Rad54 S49 phosphorylation in vivo.

To study DOT1L, we made a phospho-specific antibody against its pS297 epitope. We chose S297 instead of the S775 site identified in our screen because S297 is a CDK consensus site that falls within a conserved catalytic domain of DOT1L. We found that the antibody was specific for S297 (fig. S1F) and that coexpression of DOT1L with cyclin A–CDK2 strongly stimulated DOT1L S297 phosphorylation (Fig. 5D). We were unable to detect endogenous DOT1L protein and thus could not study its phosphorylation.

Last, we performed similar validation experiments for LSD1. Although we failed to make a phospho-specific antibody against T59 (the site we identified by MS), a commercial phospho-T/P antibody specifically detected T59 phosphorylation of LSD1 in vitro, as shown by the strong LSD1 phosphorylation after cyclin A–CDK2 treatment and the lack of detectable phosphorylation of an LSD1 T59A mutant (fig. S1E). Using this phospho-T/P antibody, we found that endogenous LSD1 is phosphorylated in vivo and strongly stimulated by cyclin A–CDK2, suggesting that cyclin A–CDK2 phosphorylates LSD1 on T59 in vivo (Fig. 5E). Endogenous LSD1 was also strongly phosphorylated in nocodazole-treated cells in a roscovitine-dependent and RO-3306–dependent manner (Fig. 5F). This phosphorylation was completely eliminated when an ectopic T59A LSD1 mutant was compared with ectopic WT LSD1, indicating that pT59 is also recognized by the pT/P antibody in vivo (Fig. 5G). We also studied synchronized cells to determine whether LSD1 T59 phosphorylation exhibited cell cycle regulation. By following cells arrested in early S phase by a double thymidine block and release, we found that T59 phosphorylation was lowest in G<sub>1</sub> phase, begins to rise in S phase, and peaks in G<sub>2</sub>–M phase (Fig. 5, H and I). This pattern is highly reflective of CDK2 activity during the cell cycle, although we were unable to directly test this because CDK2 inhibition prevents cell cycle progression after cell synchronization and release. In



**Fig. 5. CDK2-dependent phosphorylation of RAD54, DOT1L, and LSD1 in vivo and cell cycle regulation of LSD1 T59 phosphorylation.** (A) Myc-tagged RAD54 WT and T31A mutant were immunoprecipitated from transfected HEK293 cells and immunoblotted for T31 phosphorylation and total RAD54 protein. Roscovitine (Rosc) was added 2 hours before harvest. Untransfected cells, control. Arrow denotes endogenous RAD54. (B) HEK293 cells were cotransfected with myc-tagged RAD54 and myc-tagged cyclin A-HA-CDK2 or vector control. RAD54 was immunoprecipitated and immunoblotted for T31 phosphorylation and total RAD54 protein (top), and the cell lysates were immunoblotted for cyclin A and CDK2 expression (bottom).  $\gamma$ -Tubulin—loading control. (C) HEK293 cells were treated with nocodazole (Noc) for 17 hours, and endogenous RAD54 was immunoprecipitated and immunoblotted for T31 phosphorylation and for total RAD54 protein. Roscovitine or RO-3306 was added 1 hour before harvest. “Noc-” indicates asynchronous cells. (D) FLAG-tagged DOT1L was analyzed as in (B). DOT1L phosphorylation was detected using an anti-phospho-S297 DOT1L antibody. Untransfected cells, control. (E) HEK293 cells were transfected with myc-cyclin A and HA-CDK2 or vector control. Endogenous LSD1 was immunoprecipitated and immunoblotted for T59 phosphorylation using anti-phospho-TP antibody and also for total LSD1 (top). Cell lysates were immunoblotted for cyclin A and CDK2 expression (bottom).  $\gamma$ -Tubulin, loading control. (F) Endogenous LSD1 was analyzed as in (C). (G) FLAG-tagged LSD1 WT and T59A mutant were immunoprecipitated from transfected HEK293 cells that were treated with nocodazole for 17 hours and immunoblotted for T59 phosphorylation and total LSD1. (H) HEK293 cells were synchronized at the G<sub>1</sub>-S boundary by double thymidine block and released. Cells were harvested at indicated times. Endogenous LSD1 was immunoprecipitated and immunoblotted for T59 phosphorylation using anti-phospho-TP antibody and also for total LSD1 (top). Cell lysates were immunoblotted for LSD1 and  $\gamma$ -tubulin (bottom). (I) Cell cycle profiles at the indicated times using flow cytometry.

summary, these data support that Rad54, DOT1L, and LSD1 are in vivo substrates of CDK2.

## DISCUSSION

CDK2 has critical roles in cell division, DNA replication, DNA damage and repair, and cell cycle checkpoints. Moreover, CDK2 mediates many oncogenic pathways and responses to anti-neoplastic agents. While many CDK2 substrates in these pathways have been identified, crucial aspects of normal and neoplastic CDK2 function remain unknown. We thus developed in situ phosphorylation to enable future mechanistic studies by identifying previously unknown physio-

logic substrates. In addition to many known CDK2 substrates, we identified ~70 previously undiscovered candidate substrates. Validation studies of a set of nine candidates confirmed that each was a CDK2 substrate, suggesting that many (or most) candidates will be found to be bona fide substrates. In situ phosphorylation identified a very different set of CDK2 substrates from those found in vitro, including many chromatin-associated proteins that we speculate were discovered because in situ kinase reactions are performed under conditions in which the normal subcellular relationships between CDK2 and its nuclear substrates are preserved.

One particularly interesting set of candidates includes enzymes that modify histones and chromatin. CDK2 affects histone abundance

and activity in several ways: (i) by direct histone phosphorylation (29), (ii) by regulating histone transcription through NPAT phosphorylation (30), and (iii) through phosphorylation of histone-modifying enzymes, such as EZH2, which controls epigenetic gene silencing (31). Here, we identify several candidate CDK2 substrates with important roles in histone modification (DOT1L, JARID2, KAT6A, LSD1, MSL1, MSL3, SETDB1, and NSD2) or broader roles in chromatin remodeling (e.g., BCOR, DMAP1, INO80E, and SMARCA5).

Another interesting group of candidate substrates is composed of transcriptional regulators and TFs, which include proteins with roles in cancer (e.g., BCL11A and AF9) and more general transcriptional functions (e.g., GTF2I). We also found several proteins with roles in splicing and RNA metabolism (e.g., PRP3, PRP16/DHX38, and SRRM2). Again, we speculate that CDK2-mediated phosphorylation of these transcriptional and mRNA regulators may thus provide important mechanisms linking control of cell division with gene expression. Last, while CDK2's roles in DNA replication, damage, and repair are well known, we identified previously undiscovered putative substrates with important roles within these well-studied pathways, including RIF1, RAD54, and XRCC1. Future studies are needed not only to validate these candidates but also to understand how they are regulated by CDK2-mediated phosphorylation during the cell cycle.

As noted above, many of the CDK2 substrates we found coexist within known complexes, suggesting that CDK2 may regulate these complexes by targeting multiple components for phosphorylation, perhaps within the context of the complexes themselves. For example, while many cell cycle and DNA repair proteins are found in complexes with canonical functions (e.g., BASC complex; Fig. 3B), these proteins are also found in complexes with unknown functions, such as the Emerin 24 complex (Fig. 3C). Studies of these atypical complexes may reveal new functions for CDK2.

One particularly interesting finding involves the presence of multiple CDK2 substrates within complexes with epigenetic functions, many of which contain both the LSD1 histone demethylase and the HDAC1/2 histone deacetylases. These complexes regulate gene expression through coordinated chromatin modification, including histone demethylation and acetylation. In addition to LSD1 and HDACs, these complexes also contain unique subunits, which include previously undiscovered CDK2 substrates, including RREB1, MTA1, SMARCA5, ZMYM3, and GTF2I (Fig. 3, D to G). The presence of multiple CDK2 substrates within these epigenetic regulators suggests that CDK2 may directly affect their function via direct phosphorylation of these components, perhaps in cell cycle-specific ways. There are numerous additional examples of multiple CDK substrates contained within complexes that regulate transcription, cell division, ribosomal biogenesis, and other cellular processes. We speculate that these substrates link the control of cell division with the regulation of gene expression at many levels via their phosphorylation by CDK2.

Validation work is a bottleneck for all proteomic screens, and it is highly desirable to improve the ability of these screens to identify physiologic substrates. While we initially attempted to label CDK2 substrates in living cells using different delivery systems, we were unable to deliver ATP- $\gamma$ -S analogs into live cells in an efficient and noninvasive way. We thus sought to preserve nuclear contexts while also allowing direct substrate labeling by performing *in situ* kinase assays in isolated nuclei. This *in situ* approach has notable advantages,

including (i) maintenance of nuclear structures and the cellular contexts within which CDKs normally function, (ii) near-physiologic CDK2-AS abundance and activity, and (iii) maintaining thiophosphate signatures after elution, which allows more confident identification of peptides phosphorylated by the AS-kinase. As a result, the list of candidate CDK2 substrates was highly enriched for known CDK substrates, and all of the candidates that we studied were validated to be phosphorylated by CDK2. We thus believe that it is likely that many (or most) of the candidates that remain untested will prove to be bona fide CDK2 substrates.

Although we confirmed a high rate of validation, several factors may still contribute to false-positive substrate identifications. For example, nearly half of the thiophosphopeptides we identified were found in both CDK-AS and CDK2-WT kinase assays, indicating that some AS-CDK2-independent background labeling occurs, with a signal-to-noise ratio of about 1 (table S3). This background likely results from cellular kinases known to use bulky ATP analogs (32). While it is also possible that PE-ATP- $\gamma$ -S was hydrolyzed during the kinase reaction (thereby generating ATP- $\gamma$ -S that can be used by many cellular kinases), we previously found that ATP- $\gamma$ -S and PE-ATP- $\gamma$ -S are stable in whole-cell lysates, and this is likely also true in isolated nuclei (19). Despite this background, by requiring that thiophosphopeptides contain S/T-P motifs to be considered as substrates, the vast majority of background thiophosphopeptides were eliminated from subsequent analyses. The few thiophosphopeptides present in the "WT-CDK2-only" samples indicates that our MS analyses did not reach saturation with respect to peptide identifications. Thus, some "AS-specific" candidate substrates may drop out if they are found in WT-CDK samples with additional replicates. However, thus far, our validations have not revealed many false-positive identifications. Last, because our methods require isolating intact nuclei, we could only identify CDK2 substrates in interphase cells. This limitation may explain why we did not detect elongation factor 2, which is a cytoplasmic protein phosphorylated by CDK2 only in mitosis (33).

In conclusion, we describe an *in situ* implementation of AS-CDK2 for substrate discovery that allows CDK2 to phosphorylate its substrates under near-physiologic conditions. We also developed methodological advances to identify candidate AS-CDK2 substrates more definitively and comprehensively. As a result, we found a high percentage of known CDK substrates and a high rate of validation for previously undiscovered candidates. Moreover, the enrichment for chromatin-associated proteins with important roles in processes such as gene expression and DNA and RNA metabolism suggests that *in situ* conditions may allow new types of substrates to be identified for CDKs and other nuclear kinases that function within specific nuclear contexts. These methods should be readily adaptable to study other nuclear kinases that are amenable to AS mutations.

## MATERIALS AND METHODS

### Experimental design

The objectives of this study were to develop a near physiologic environment for substrate phosphorylation by AS-CDK2. The primary study design provided two mechanisms to accomplish this goal. The first was to stably express AS-CDK2 at normal abundance and to allow the exogenous AS-CDK2 to be activated by endogenous cyclins, which are rate limiting for CDK2 activation. The second was to isolate nuclei and allow substrates to be phosphorylated

by AS-CDK2 in a nearly normal nuclear environment. We used substrate thiophosphorylation for several reasons: (i) We previously have characterized the ability of AS-CDK2 to use and hydrolyse ATP- $\gamma$ -S, (ii) we previously developed efficient biochemical methods to enrich and identify thiophosphopeptides by MS, and (iii) thiophosphate allows AS-CDK2 substrates to be easily distinguished from other cellular phosphopeptides. As shown in Fig. 2, the overall design consisted of the following steps: (i) stable expression of AS-CDK2 in HEK293A cells, (ii) inhibition of endogenous CDK2 by shRNA, (iii) treatment of cells with roscovitine to inhibit endogenous CDKs, (iv) isolation of nuclei and incubation with PE-ATP- $\gamma$ -S, (v) thiophosphopeptide enrichment and identification by MS, and (vi) substrate validation *in vitro* and *in vivo*.

### Cell culture, transfection, retroviral transduction, and drug treatments

The HEK293 was obtained from the American Type Culture Collection, tested for mycoplasma and genetically authenticated during the course of this study. HEK293 cells were cultured in Dulbecco's modified Eagle's medium (11965-092, Gibco), which was supplemented with 10% (v/v) fetal bovine serum (10437-028, Gibco), penicillin (100 U/ml), and streptomycin (100  $\mu$ g/ml; 15140-122, Gibco). Transient transfections were carried out using calcium phosphate. HEK293 cells stably expressing C-terminally HA-tagged WT-CDK2 or CDK2-F80A (AS-CDK2) were generated via retroviral transduction using a pBabe-based vector (puromycin) as previously described (34). For CDK2 knockdown via small hairpin RNA, cells were transduced with lentiviruses encoding nontargeting control shRNA or two different shRNAs directed to noncoding regions of CDK2 mRNA (see below). Cells were harvested 2 days after infection without selection, and shRNA expression was confirmed by green fluorescent protein coexpression. Nocodazole (M1404, Sigma-Aldrich) was used at 40 ng/ml, for 17 hours, roscovitine (R7772, Sigma-Aldrich) was used at 25  $\mu$ M for 1 to 2 hours where indicated, and RO-3306 (sc-358700; Santa Cruz Biotechnology) was used at 1  $\mu$ M for 1 hour. Okadaic acid (O-2220, LC Laboratories) was used at 200 nM for 90 min.

### Antibodies, plasmids, enzymes, and other reagents

The following antibodies were used: Myc tag (1:5; 9E10 hybridoma supernatant, prepared in house), HA-tag (12CA5 hybridoma supernatant; prepared in-house), FLAG M2 (1:4000; F1804, Sigma-Aldrich), CDK2 (1:1000; D-12, sc-6248, Santa Cruz Biotechnology), cyclin A (1:1000; gift from J. Roberts, Fred Hutchinson Cancer Research Center), cyclin E (1:1000; HE111, sc-248, Santa Cruz Biotechnology),  $\gamma$ -tubulin (1:1000; sc-17787; Santa Cruz Biotechnology), RAD54 (1:1000; F-11, sc-374598, Santa Cruz Biotechnology), LSD1 (1:5000; A300-215A, Bethyl Laboratories), DOT1L (1:2000; A300-953A, Bethyl Laboratories), and phospho-T/P (1:1000; 9391S; Cell Signaling Technology). Phospho-T31 RAD54 and phospho-S297 DOT1L antibodies were custom made by Yenzym Antibodies (San Francisco), raised against the synthetic phosphopeptides CEDWQPGLV-pT-PRKRK (RAD54) and CMRVVEL-pS-PLKGSVS (DOT1L).

Substrate expression plasmids were obtained as follows: FLAG-DMAP1 (C. Kim, Hanyang University), FLAG-NSD2 (J. Licht, University of Florida), HA-MSL3 (J. Cotes, Laval University), HA-GTF3C2 (R. White, University of York), FLAG-PRPF3 (S. M. Iguchi-Arigo, Hokkaido University), and N-terminal FLAG-DOT1L and C-terminal HA-DOT1L (Y. Zhang, Harvard). RAD54 was cloned into the pCS2<sup>+</sup> vector with a C-terminal 5 $\times$  myc tag. The deleted

sequence for myc-RAD54-T31A $\Delta$ 8 was EDWQPGLV. N-terminal 3 $\times$  FLAG-tagged LSD1 was cloned into a pCS2<sup>+</sup> vector. BCL11A was cloned into the pCS2<sup>+</sup> 5myc vector. Cyclin A and CDK2 expression plasmids were pCS2<sup>+</sup> 6 myc-cyclinA (gift from J. Roberts, Fred Hutchinson Cancer Research Center) and pCMV-CDK2HA (gift from S. van den Heuvel, Utrecht University, The Netherlands), respectively. All mutants were generated by site-directed mutagenesis using the QuikChange method (Stratagene) and sequenced. Lentiviral plasmids using the pGIPZ vector were purchased onsite, and the CDK2 3' UTR targeting sequences of the two shRNAs used were AGGATGAACAATTATATTT (#1) and AGGTTATATC-CAATAGTAG (#2). Lambda protein phosphatase (P0753S) and recombinant cyclin A-CDK2 were purchased from New England Biolabs. Recombinant glutathione S-transferase (GST)-cyclin E-CDK2 (33) and GST-Rb (19) were prepared in house. ATP- $\gamma$ -S was purchased from Millipore Sigma (catalog no. 119120). PE-ATP- $\gamma$ -S was purchased from BIOLOG Life Sciences Institute (catalog no. P026, Bremen, Germany).

### Immunoblotting, immunoprecipitation, kinase assays, cell synchronization, and flow cytometry

Immunoblotting, Immunoprecipitation, kinase assays, and flow cytometry were performed as described (34). Images of full gels for each figure are shown in fig. S2. For DMAP1 kinase assay, DMAP1 was first eluted from immunoprecipitates with FLAG peptide (150  $\mu$ g/ml; F4799, Millipore Sigma) and concentrated using a MicroCon-30 (MRCF0R030, Millipore Sigma-Aldrich). For cell cycle analysis of LSD1 phosphorylation, HEK293 cells were first synchronized by double thymidine block. Briefly, cells were treated with 2 mM thymidine (T9250, Millipore Sigma-Aldrich) for 16 hours, washed once with fresh medium, and released into fresh medium supplemented with 24  $\mu$ M 2'-deoxycytidine (951-77-9, Acros Organics) for 8.5 hours. Cells were then washed once with fresh medium and treated again with 2 mM thymidine for 18 hours before releasing into fresh medium containing 2'-deoxycytidine. Cells were harvested by trypsinization at various times up to 25 hours, and a fraction of the cells were fixed for flow cytometry analysis.

### Cell nuclei isolation and in situ kinase assay

HEK293 cells stably expressing WT-CDK2 or AS-CDK2 were grown on eight 15-cm plates (2 days after infection by CDK2 shRNA lentivirus) to ~80 to 90% confluence (~2  $\times$  10<sup>7</sup> cells per plate), treated with roscovitine for 1 hour, and harvested in four replicates (two plates per cell pellet) by trypsinization, and washed once with cold phosphate-buffered saline buffer. Each cell pellet (total of eight) was washed twice by resuspending the cells in 1 ml of cold hypotonic lysis buffer 1 [10 mM Hepes (pH 7.4), 10 mM KCl, and 2 mM MgCl<sub>2</sub>] followed by centrifugation in a microcentrifuge (2000 rpm, 1 min). Cells were resuspended again in 1 ml of hypotonic lysis buffer 1 and incubated on ice for 15 min. Swollen cells were lysed by pipetting up and down five to seven times using a 1-ml syringe with 26-gauge needle (305111, Becton, Dickinson and Company). Cell lysis and intact cell nuclei were verified by staining a small aliquot with trypan blue (T10282, Life Technologies). The slurry was slowly added to the surface of 6 ml of cold hypotonic lysis buffer 1 containing 30% sucrose (w/v) in a 15-ml Falcon tube. After centrifugation in a tabletop centrifuge at 1000g for 10 min, the buffer above the nuclei pellet was removed, and the pellet was washed three times by resuspending it in 1 ml of cold hypotonic lysis buffer 1 followed by



centrifugation in a microcentrifuge (3000 rpm, 1 min). The nuclei preparation was checked again by staining with trypan blue and microscope examination. The final nuclei pellet was resuspended in 1.5× volume of hypotonic lysis buffer 1 containing PE-ATP- $\gamma$ -S (final concentration of 0.5 mM) and MnCl<sub>2</sub> (final concentration of 0.5 mM final) and incubated at 30°C for 30 min. The nuclei slurry was occasionally mixed by tapping during the course of the reaction. After the reaction, the nuclei mix was briefly centrifuged (3000 rpm, 15 s) to remove most of the supernatant, and the pellet was flash-frozen in liquid nitrogen and stored or processed as described below. ATP- $\gamma$ -S labeling was done similarly starting with two 15-cm plates of WT-CDK2 cells and labeled at a final concentration of 0.5 mM.

### Purification of thiophosphorylated peptides

The frozen nuclei pellet was resuspended in 0.4 ml of hypotonic lysis buffer 2 [30 mM Hepes (pH 7.4), 10 mM EDTA, and benzonase (25 U/ml; 70746, Millipore Sigma)]. After incubation on ice for 30 min, Tween-20 was added to a final concentration of 0.1%, and the sample was sonicated using 20× 1-s pulses. Nuclei debris was pelleted by centrifugation at 20,000g for 10 min. The supernatant was digested with sequencing grade modified trypsin (Promega) at 1:20 ratio (w/w), and thiophosphopeptides from the peptide mixture were purified by binding to 40  $\mu$ l of disulfide beads Thiopropyl Sepharose 6B (17042001, GE Healthcare) at pH 4.0 as previously described (19). Washed beads were eluted with 30  $\mu$ l of 25 mM DTT (pH ~4 without buffering) in 5% acetonitrile/95% H<sub>2</sub>O at room temperature for 30 min. The eluate was acidified with tris(2-carboxyethyl)phosphine and formic acid to a final concentration of 5 mM and 0.1%, respectively, and analyzed directly by MS.

### MS analysis and database search

Phosphopeptides samples were analyzed by Nanoflow liquid chromatography (NanoLC) and electrospray ionization tandem MS (MS/MS) using an LTQ-Orbitrap mass spectrometer (Thermo Fisher Scientific) interfaced with an Agilent 1100 Nano Pump with electronically controlled split flow. For ATP- $\gamma$ -S labeling, one sample was analyzed in duplicate MS runs, and for PE-ATP- $\gamma$ -S labeling, eight samples (four WT-CDK2 and four AS-CDK2) were analyzed in duplicate MS runs (16 MS runs in total). Peptides were loaded in sequence onto a 75  $\mu$ m (inner diameter) by 15 cm C18 microcapillary column, packed in-house with Magic C18 AQ 5- $\mu$ m resin (Michrom Bioresources), and resolved by a nonlinear gradient of 5 to 28% acetonitrile containing 0.1% formic acid at a flow rate of 300 nl/min over the course of 80 min. Each survey scan in the Orbitrap was followed by MS/MS scans of the top nine most intense precursor ions in the linear ion trap. Tandem spectra acquired were searched against a human Uniprot database (downloaded January 2015) with target decoy using the Comet algorithm (version 2014.02) (35). Peptide search parameters included precursor mass tolerance of 20 parts per million, one tryptic end for peptide, and differential mass modification to methionine (+15.999) due to oxidation and serine and threonine (+96.0329) due to thiophosphorylation. Search results were filtered using Trans Proteomic Pipeline (36) with a minimal iProphet (37) score of 0.75 and corresponding peptide false discovery rate (FDR) between 0.5 to 1%.

### Functional enrichment analysis of CDK2 substrates

A network made up of the candidate substrates were created by manually inputting the list into the STRING protein query within

Cytoscape (38, 39) and analyzed using the STRING functional enrichment tool with an enrichment FDR value cutoff of 0.05. Select enriched functional categories were generated on the basis of the Gene Ontology process category.

### SUPPLEMENTARY MATERIALS

Supplementary material for this article is available at <http://advances.sciencemag.org/cgi/content/full/6/16/eaaz9899/DC1>

[View/request a protocol for this paper from Bio-protocol.](#)

### REFERENCES AND NOTES

1. S. A. Johnson, T. Hunter, Kinomics: Methods for deciphering the kinome. *Nat. Methods* **2**, 17–25 (2005).
2. J. Ptacek, M. Snyder, Charging it up: Global analysis of protein phosphorylation. *Trends Genet.* **22**, 545–554 (2006).
3. K. Shah, H. Kim, The significant others: Global search for direct kinase substrates using chemical approaches. *IUBMB Life* **71**, 721–737 (2019).
4. S. N. Boyle, A. J. Koleske, Dissecting kinase signaling pathways. *Drug Discov. Today* **12**, 717–724 (2007).
5. K. Shah, Y. Liu, C. Deirmengian, K. M. Shokat, Engineering unnatural nucleotide specificity for Rous sarcoma virus tyrosine kinase to uniquely label its direct substrates. *Proc. Natl. Acad. Sci. U.S.A.* **94**, 3565–3570 (1997).
6. A. Koch, S. Hauf, Strategies for the identification of kinase substrates using analog-sensitive kinases. *Eur. J. Cell Biol.* **89**, 184–193 (2010).
7. P. J. Alaimo, M. A. Shogren-Knaak, K. M. Shokat, Chemical genetic approaches for the elucidation of signaling pathways. *Curr. Opin. Chem. Biol.* **5**, 360–367 (2001).
8. J. W. Harper, P. D. Adams, Cyclin-dependent kinases. *Chem. Rev.* **101**, 2511–2526 (2001).
9. D. O. Morgan, Cyclin-dependent kinases: Engines, clocks, and microprocessors. *Annu. Rev. Cell Dev. Biol.* **13**, 261–291 (1997).
10. J. D. Blethrow, J. S. Glavy, D. O. Morgan, K. M. Shokat, Covalent capture of kinase-specific phosphopeptides reveals Cdk1-cyclin B substrates. *Proc. Natl. Acad. Sci. U.S.A.* **105**, 1442–1447 (2008).
11. Y. Chi, B. E. Clurman, Mass spectrometry-based identification of protein kinase substrates utilizing engineered kinases and thiophosphate labeling. *Curr. Protoc. Chem. Biol.* **2**, ch100151 (2010).
12. K. A. Merrick, L. Wohlbold, C. Zhang, J. J. Allen, D. Horiuchi, N. E. Huskey, A. Goga, K. M. Shokat, R. P. Fisher, Switching Cdk2 on or off with small molecules to reveal requirements in human cell proliferation. *Mol. Cell* **42**, 624–636 (2011).
13. K. A. Merrick, S. Larochelle, C. Zhang, J. J. Allen, K. M. Shokat, R. P. Fisher, Distinct activation pathways confer cyclin-binding specificity on Cdk1 and Cdk2 in human cells. *Mol. Cell* **32**, 662–672 (2008).
14. L. Wohlbold, K. A. Merrick, S. De, R. Amat, J. H. Kim, S. Larochelle, J. J. Allen, C. Zhang, K. M. Shokat, J. H. J. Petrini, R. P. Fisher, Chemical genetics reveals a specific requirement for Cdk2 activity in the DNA damage response and identifies Nbs1 as a Cdk2 substrate in human cells. *PLoS Genet.* **8**, e1002935 (2012).
15. K.-H. Chang, F. Vincent, K. Shah, Deregulated Cdk5 triggers aberrant activation of cell cycle kinases and phosphatases inducing neuronal death. *J. Cell Sci.* **125**, 5124–5137 (2012).
16. C. Shi, K. Viccaro, H.-G. Lee, K. Shah, Cdk5-Foxo3 axis: Initially neuroprotective, eventually neurodegenerative in Alzheimer's disease models. *J. Cell Sci.* **129**, 1815–1830 (2016).
17. M. Sansó, R. S. Levin, J. J. Lipp, V. Y.-F. Wang, A. K. Greifengberg, E. M. Quezada, A. Ali, A. Ghosh, S. Larochelle, T. M. Rana, M. Geyer, L. Tong, K. M. Shokat, R. P. Fisher, P-TEFb regulation of transcription termination factor Xrn2 revealed by a chemical genetic screen for Cdk9 substrates. *Genes Dev.* **30**, 117–131 (2016).
18. S. Larochelle, K. A. Merrick, M. E. Terret, L. Wohlbold, N. M. Barboza, C. Zhang, K. M. Shokat, P. V. Jallepalli, R. P. Fisher, Requirements for Cdk7 in the assembly of Cdk1/cyclin B and activation of Cdk2 revealed by chemical genetics in human cells. *Mol. Cell* **25**, 839–850 (2007).
19. Y. Chi, M. Welcker, A. A. Hizli, J. J. Posakony, R. Aebersold, B. E. Clurman, Identification of CDK2 substrates in human cell lysates. *Genome Biol.* **9**, R149 (2008).
20. L. L. Parker, A. B. Schilling, S. J. Kron, S. B. H. Kent, Optimizing thiophosphorylation in the presence of competing phosphorylation with MALDI-TOF-MS detection. *J. Proteome Res.* **4**, 1863–1866 (2005).
21. S. W. Kwon, S. C. Kim, J. Jaunbergs, J. R. Falck, Y. Zhao, Selective enrichment of thiophosphorylated polypeptides as a tool for the analysis of protein phosphorylation. *Mol. Cell. Proteomics* **2**, 242–247 (2003).
22. P. Mertins, D. R. Mani, K. V. Ruggles, M. A. Gillette, K. R. Clauser, P. Wang, X. Wang, J. W. Qiao, S. Cao, F. Petralia, E. Kawaler, F. Mundt, K. Krug, Z. Tu, J. T. Lei, M. L. Gatzza, M. Wilkerson, C. M. Perou, V. Yellapantula, K.-L. Huang, C. Lin, M. D. McLellan, P. Yan,

- S. R. Davies, R. R. Townsend, S. J. Skates, J. Wang, B. Zhang, C. R. Kinsinger, M. Mesri, H. Rodriguez, L. Ding, A. G. Paulovich, D. Fenyö, M. J. Ellis, S. A. Carr; NCI CPTAC, Proteogenomics connects somatic mutations to signalling in breast cancer. *Nature* **534**, 55–62 (2016).
23. S. A. Stuart, S. Houel, T. Lee, N. Wang, W. M. Old, N. G. Ahn, A phosphoproteomic comparison of B-RAFV600E and MKK1/2 inhibitors in melanoma cells. *Mol. Cell. Proteomics* **14**, 1599–1615 (2015).
24. K. Sharma, R. C. J. D'Souza, S. Tyanova, C. Schaab, J. R. Wiśniewski, J. Cox, M. Mann, Ultra-deep human phosphoproteome reveals a distinct regulatory nature of Tyr and Ser/Thr-based signaling. *Cell Rep.* **8**, 1583–1594 (2014).
25. T. Shirimizu, J. Adachi, S. Watanabe, T. Murakami, T. Kuga, S. Muraoka, T. Tomonaga, Identification of missing proteins in the neXtProt database and unregistered phosphopeptides in the phosphositeplus database as part of the chromosome-centric human proteome project. *J. Proteome Res.* **12**, 2414–2421 (2013).
26. R. Oughtred, C. Stark, B.-J. Breitkreutz, J. Rust, L. Boucher, C. Chang, N. Kolas, L. O'Donnell, G. Leung, R. McAdam, F. Zhang, S. Dolma, A. Willems, J. Coulombe-Huntington, A. Chatr-aryamontri, K. Dolinski, M. Tyers, The BioGRID interaction database: 2019 update. *Nucleic Acids Res.* **47**, D529–D541 (2019).
27. M. Giurgiu, J. Reinhard, B. Brauner, I. Dunger-Kaltenbach, G. Fobo, G. Frishman, C. Montrone, A. Ruepp, CORUM: The comprehensive resource of mammalian protein complexes-2019. *Nucleic Acids Res.* **47**, D559–D563 (2019).
28. N. Goyal, M. J. Rossi, O. M. Mazina, Y. Chi, R. L. Moritz, B. E. Clurman, A. V. Mazin, RAD54 N-terminal domain is a DNA sensor that couples ATP hydrolysis with branch migration of Holliday junctions. *Nat. Commun.* **9**, 34 (2018).
29. A. Contreras, T. K. Hale, D. L. Stenoien, J. M. Rosen, M. A. Mancini, R. E. Herrera, The dynamic mobility of histone H1 is regulated by cyclin/CDK phosphorylation. *Mol. Cell. Biol.* **23**, 8626–8636 (2003).
30. T. Ma, B. A. Van Tine, Y. Wei, M. D. Garrett, D. Nelson, P. D. Adams, J. Wang, J. Qin, L. T. Chow, J. W. Harper, Cell cycle-regulated phosphorylation of p220(NPAT) by cyclin E/Cdk2 in Cajal bodies promotes histone gene transcription. *Genes Dev.* **14**, 2298–2313 (2000).
31. S. Chen, L. R. Bohrer, A. N. Rai, Y. Pan, L. Gan, X. Zhou, A. Bagchi, J. A. Simon, H. Huang, Cyclin-dependent kinases regulate epigenetic gene silencing through phosphorylation of EZH2. *Nat. Cell Biol.* **12**, 1108–1114 (2010).
32. B. C. Kraybill, L. L. Elkin, J. D. Blethrow, D. O. Morgan, K. M. Shokat, Inhibitor scaffolds as new allele specific kinase substrates. *J. Am. Chem. Soc.* **124**, 12118–12128 (2002).
33. A. A. Hizli, Y. Chi, J. Swanger, J. H. Carter, Y. Liao, M. Welcker, A. G. Ryazanov, B. E. Clurman, Phosphorylation of eukaryotic elongation factor 2 (eEF2) by cyclin A-cyclin-dependent kinase 2 regulates its inhibition by eEF2 kinase. *Mol. Cell. Biol.* **33**, 596–604 (2013).
34. M. Welcker, J. Singer, K. R. Loeb, J. Grim, A. Bloecher, M. Gurien-West, B. E. Clurman, J. M. Roberts, Multisite phosphorylation by Cdk2 and GSK3 controls cyclin E degradation. *Mol. Cell* **12**, 381–392 (2003).
35. J. K. Eng, T. A. Jahan, M. R. Hoopmann, Comet: An open-source MS/MS sequence database search tool. *Proteomics* **13**, 22–24 (2013).
36. E. W. Deutsch, L. Mendoza, D. Shteynberg, T. Farrah, H. Lam, N. Tasman, Z. Sun, E. Nilsson, B. Pratt, B. Prazen, J. K. Eng, D. B. Martin, A. I. Nesvizhskii, R. Aebersold, A guided tour of the trans-proteomic pipeline. *Proteomics* **10**, 1150–1159 (2010).
37. D. Shteynberg, E. W. Deutsch, H. Lam, J. K. Eng, Z. Sun, N. Tasman, L. Mendoza, R. L. Moritz, R. Aebersold, A. I. Nesvizhskii, iProphet: Multi-level integrative analysis of shotgun proteomic data improves peptide and protein identification rates and error estimates. *Mol. Cell. Proteomics* **10**, M111.007690 (2011).
38. R. Saito, M. E. Smoot, K. Ono, J. Ruscheinski, P. L. Wang, S. Lotia, A. R. Pico, G. D. Bader, T. Ideker, A travel guide to cytoscape plugins. *Nat. Methods* **9**, 1069–1076 (2012).
39. P. Shannon, A. Markiel, O. Ozier, N. S. Baliga, J. T. Wang, D. Ramage, N. Amin, B. Schwikowski, T. Ideker, Cytoscape: A software environment for integrated models of biomolecular interaction networks. *Genome Res.* **13**, 2498–2504 (2003).

**Acknowledgments:** We thank members of the Clurman laboratory for helpful discussions during the course of this project. We thank S. Li, D. Shteynberg, Z. Sun, and members of the Moritz laboratory for technical and computational help throughout the method development. **Funding:** This publication was supported by grants 1R01CA193808 (B.E.C.), CA188347 (A.V.M.), and P30CA056036 and Cancer Center Support grant 2P30CA015704. This work was funded in part by the National Institutes of Health, National Institute of General Medical Sciences grants: R01GM087221 (R.L.M.). **Author contributions:** Y.C. designed the method, performed nuclei kinase assays and phosphopeptide purifications, carried out MS and data analyses, and substrate validation experiments. J.H.C. carried out substrate validation experiments. J.S. performed AS-CDK2 cell line characterization and kinase assays. A.V.M. provided scientific input. R.L.M. provided MS and software resources. Y.C., R.L.M., and B.E.C. designed the research project. Y.C. and B.E.C. wrote the manuscript. **Competing interests:** The authors declare that they have no competing interests. B.E.C. is the founder and equity holder of Coho Therapeutics, a startup company that is focused on protein degradation and is completely unrelated to the work presented in this paper. Coho Therapeutics did not provide any funding or support for this work. **Data and materials availability:** All data needed to evaluate the conclusions in the paper are present in the paper and/or the Supplementary Materials. Additional data (RAW files, search parameters, database file, and search results) has been deposited in the Institute for Systems Biology's Peptide Atlas repository and can be accessed via the following Weblink: [www.peptideatlas.org/PASS/PASS01510](http://www.peptideatlas.org/PASS/PASS01510). Additional data related to this paper may be requested from the authors.

Submitted 29 October 2019

Accepted 23 January 2020

Published 17 April 2020

10.1126/sciadv.aaz9899

**Citation:** Y. Chi, J. H. Carter, J. Swanger, A. V. Mazin, R. L. Moritz, B. E. Clurman, A novel landscape of nuclear human CDK2 substrates revealed by in situ phosphorylation. *Sci. Adv.* **6**, eaaz9899 (2020).

## A novel landscape of nuclear human CDK2 substrates revealed by in situ phosphorylation

Yong Chi, John H. Carter, Jherek Swanger, Alexander V. Mazin, Robert L. Moritz and Bruce E. Clurman

*Sci Adv* **6** (16), eaaz9899.

DOI: 10.1126/sciadv.aaz9899

### ARTICLE TOOLS

<http://advances.sciencemag.org/content/6/16/eaaz9899>

### SUPPLEMENTARY MATERIALS

<http://advances.sciencemag.org/content/suppl/2020/04/13/6.16.eaaz9899.DC1>

### REFERENCES

This article cites 39 articles, 12 of which you can access for free  
<http://advances.sciencemag.org/content/6/16/eaaz9899#BIBL>

### PERMISSIONS

<http://www.sciencemag.org/help/reprints-and-permissions>

Use of this article is subject to the [Terms of Service](#)

---

*Science Advances* (ISSN 2375-2548) is published by the American Association for the Advancement of Science, 1200 New York Avenue NW, Washington, DC 20005. The title *Science Advances* is a registered trademark of AAAS.

Copyright © 2020 The Authors, some rights reserved; exclusive licensee American Association for the Advancement of Science. No claim to original U.S. Government Works. Distributed under a Creative Commons Attribution NonCommercial License 4.0 (CC BY-NC).