Structure and activation mechanism of the yeast RNA Pol II CTD kinase CTDK-1 complex

Yihu Xie[a], Christopher L. Lord[b], Bradley P. Clarke[a], Austin L. Ivey[a], Pate S. Hill[a], W. Hayes McDonald[a,c], Susan R. Wente[b], and Yi Ren[a,1]

[a]Department of Biochemistry, Vanderbilt University School of Medicine, Nashville, TN 37232; [b]Department of Cell and Developmental Biology, Vanderbilt University School of Medicine, Nashville, TN 37232; and [c]Mass Spectrometry Research Center, Vanderbilt University School of Medicine, Nashville, TN 37232

Edited by Hao Wu, Harvard Medical School, Boston, MA, and approved December 15, 2020 (received for review September 11, 2020)

The C-terminal domain (CTD) kinase I (CTDK-1) complex is the primary RNA Polymerase II (Pol II) CTD Ser2 kinase in budding yeast. CTDK-1 consists of a cyclin-dependent kinase (CDK) Ctk1, a cyclin Ctk2, and a unique subunit Ctk3 required for CTDK-1 activity. Here, we present a crystal structure of CTDK-1 at 1.85-Å resolution. The structure reveals that, compared to the canonical two-component CDK-cyclin system, the third component Ctk3 of CTDK-1 plays a critical role in Ctk1 activation by stabilizing a key element of CDK regulation, the T-loop, in an active conformation. In addition, Ctk3 contributes to the assembly of CTDK-1 through extensive interactions with both Ctk1 and Ctk2. We also demonstrate that CTDK-1 physically and genetically interacts with the serine/arginine-like protein Gbp2. Together, the data in our work reveal a regulatory mechanism of CDK complexes.

cyclin-dependent kinase | X-ray crystallography | transcription | mRNA processing

RNA Polymerase II (Pol II) transcription is regulated by a group of cyclin-dependent kinases (CDKs) (1). The C-terminal domain (CTD) of the largest Pol II subunit consists of tandem repeats of the heptapeptide $Y_1S_2P_3T_4S_5P_6S_7$. CDKs phosphorylate the CTD, yielding specific patterns of CTD phosphorylation during the transcription cycle (2, 3). For example, Ser5 phosphorylation on Pol II is most abundant near the 5′-end of genes, whereas Ser2 phosphorylation accumulates on elongating Pol II. Phosphorylated Pol II CTD serves as a binding platform for factors required for transcription, chromatin modification, and cotranscriptional mRNA processing (2, 3). CTD phosphorylation is mediated by four CDKs in budding yeast (Kin28, Bur1, Ctk1, and Srb10) and six CDKs in humans (CDKs 7, 8, 9, 12, 13, and 19) (1). These transcriptional CDKs play critical roles in transcription and cotranscriptional processes.

In budding yeast, Ser2 phosphorylation of Pol II CTD is mainly mediated by the CTD kinase I (CTDK-1) complex (4). CTDK-1 consists of a kinase subunit Ctk1, a cyclin subunit Ctk2, and a unique subunit Ctk3 with unknown function (5–7). Deletion of Ctk1 reduces total Ser2 phosphorylation by ∼90% (8). Individual deletion of the CTDK-1 subunits results in similar slow-growth cold-sensitive phenotypes in yeast (6, 7). Catalytic activation of canonical CDK kinases is dependent on interaction with the regulatory cyclin subunit as well as phosphorylation of a regulatory loop, known as the T-loop. Intriguingly, both Ctk2 and Ctk3 regulatory subunits are suggested to be required for activating the Ctk1 kinase (9). Thus far, the only structural information about CTDK-1 comes from the crystal structure of the *Schizosaccharomyces pombe* Ctk3 (Lsg1) N-terminal domain (10), which resembles other known CTD-interaction domains (CIDs). However, it is not known how the CTDK-1 complex is assembled or importantly how Ctk1 is activated by both Ctk2 and Ctk3.

The closest evolutionary relatives of the CTDK-1 complex in humans are the CDK12•Cyclin K and CDK13•Cyclin K complexes, which have been recognized as tumor suppressors (11, 12). Several lines of evidence suggest that Ctk1 and CDK12/13 functions are intimately involved in mRNA biogenesis. Both Ctk1 and CDK12/13 contain an N-terminal extension, featuring serine/arginine-rich (SR) regions that are typically found in the SR protein family of splicing factors (13, 14). In humans, RNA processing factors including the SR protein SRSF1 were enriched in CDK12/13 purifications (15). In yeast, chromatin association of Ctk1 involves interactions with the nascent pre-mRNA (16). Ctk1 coimmunoprecipitates with the SR-like proteins Gbp2 and Hrb1 (17, 18). Thus far, evidence for a direct molecular connection of Ctk1 and CDK12/13 to SR proteins remains lacking. Beyond transcription and mRNA biogenesis, proper function of Ctk1 in yeast and CDK12/CDK13 in humans is critical for the regulation of DNA damage-response genes (19, 20). Ctk1 and CDK12 also play additional roles in translation through phosphorylation of translation machinery components (21, 22). Despite their shared functional features, the activation mechanism of CTDK-1 is likely to be uniquely different from CDK12•Cyclin K and CDK13•Cyclin K, since Ctk3 is required for Ctk1 activity and its ortholog does not exist in humans.

Here, we reveal the activation mechanism of CTDK-1 by determining a crystal structure of the Ctk1•Ctk2•Ctk3 complex at 1.85-Å resolution. The structure reveals that Ctk1 contains the conserved kinase domains (N-lobe and C-lobe), and Ctk2 interaction with Ctk1 resembles that of canonical cyclin-CDK complexes. Importantly, a highly conserved helix from the Ctk3 subunit stabilizes the T-loop of Ctk1 in an active conformation. In addition, Ctk3 features extensive interactions with both the Ctk1 kinase subunit and the Ctk2 cyclin subunit. Thus, Ctk3, together with Ctk2, are both integral structural constituents of

Significance

Gene transcription from DNA to RNA by RNA Polymerase II (Pol II) is regulated by a group of cyclin-dependent kinases (CDKs). In order to coordinate the molecular events during transcription, CDKs enzymatically modify Pol II. Here, we determined the crystal structure of an important CDK involved in Pol II regulation, the yeast C-terminal domain kinase-1 (CTDK-1) complex, to describe its architecture in atomic detail. Importantly, the CTDK-1 structure reveals a mechanism for activation of the enzyme. In addition, we show that CTDK-1 interacts with an RNA splicing factor, suggesting a connection between transcriptional regulation by CTDK-1 and subsequent mRNA processing. Together, our results provide insights into the function and regulatory mechanism of CDKs.

Author contributions: Y.X., C.L.L., B.P.C., S.R.W., and Y.R. designed research; Y.X., C.L.L., B.P.C., A.L.I., P.S.H., W.H.M., and Y.R. performed research; Y.X., C.L.L., B.P.C., W.H.M., S.R.W., and Y.R. analyzed data; and Y.X., C.L.L., B.P.C., and Y.R. wrote the paper.

The authors declare no competing interest.

This article is a PNAS Direct Submission.

[1]To whom correspondence may be addressed. Email: yi.ren@vanderbilt.edu.

This article contains supporting information online at https://www.pnas.org/lookup/suppl/doi:10.1073/pnas.2019163118/-/DCSupplemental.

Published January 11, 2021.

the CTDK-1 complex and are required to activate the Ctk1 kinase. We also demonstrate through biochemical and genetic studies that CTDK-1 functionally interacts with the SR protein Gbp2. Together, our studies reveal a unique activation mechanism of CTDK-1 and present a direct molecular link between CTDK-1 and an SR splicing factor.

Results

Structure Determination of a Ctk1•Ctk2•Ctk3 Complex. To determine the structure of the CTDK-1 complex, we first sought to identify a minimal core complex suitable for crystallization. We found that each individual subunit of CTDK-1 is required for solubility of the complex; therefore Ctk1, Ctk2, and Ctk3 subunits were coexpressed in insect cells. We were able to purify a CTDK-1 complex consisting of Ctk1 (residues 1 to 508) with a C-terminal truncation of the low-complexity region (residues 509 to 528), full-length Ctk2 (residues 1 to 323), and full-length Ctk3 (residues 1 to 296), hereafter denoted by Ctk1•Ctk2•Ctk3 (Fig. 1*A*). Next, we truncated the intrinsically disordered N-terminal extension (residues 1 to 158) of Ctk1, as this region is predicted to contain no secondary structure. The resulting Ctk1ΔN•Ctk2•Ctk3 complex was soluble and less prone to degradation compared to Ctk1•Ctk2•Ctk3. To examine whether the truncated Ctk1ΔN•Ctk2•Ctk3 complex maintains the enzymatic activity, we performed in vitro kinase assays. For the purpose of the kinase assay, various CTDK-1 complexes were coexpressed with a CDK-activating kinase (*S. pombe* Csk1), as phosphorylation of the T-loop (T338) of Ctk1 is required to generate an active enzyme (*SI Appendix*, Fig. S1 *A* and *B* and Table S1) (23, 24). Ctk1•Ctk2•Ctk3 readily phosphorylates the CTD of the yeast RNA Pol II subunit Rpb1, as observed in the appearance of a retarded band on SDS/PAGE, which disappeared upon phosphatase treatment (Fig. 1*B* and *SI Appendix*, Fig. S1*A*). In addition, the kinase dead mutant Ctk1(D324N)•Ctk2•Ctk3 did not yield a change in CTD mobility, excluding the possibility that the observed activity came from contaminating proteins. We found that Ctk1ΔN•Ctk2•Ctk3 exhibited a similar level of kinase activity compared to Ctk1•Ctk2•Ctk3. Together, these data identified Ctk1ΔN•Ctk2•Ctk3 as a structural and functional unit suitable for structural characterization.

We obtained Ctk1ΔN•Ctk2•Ctk3 crystals grown in the P1 space group with one complex in the asymmetric unit. The

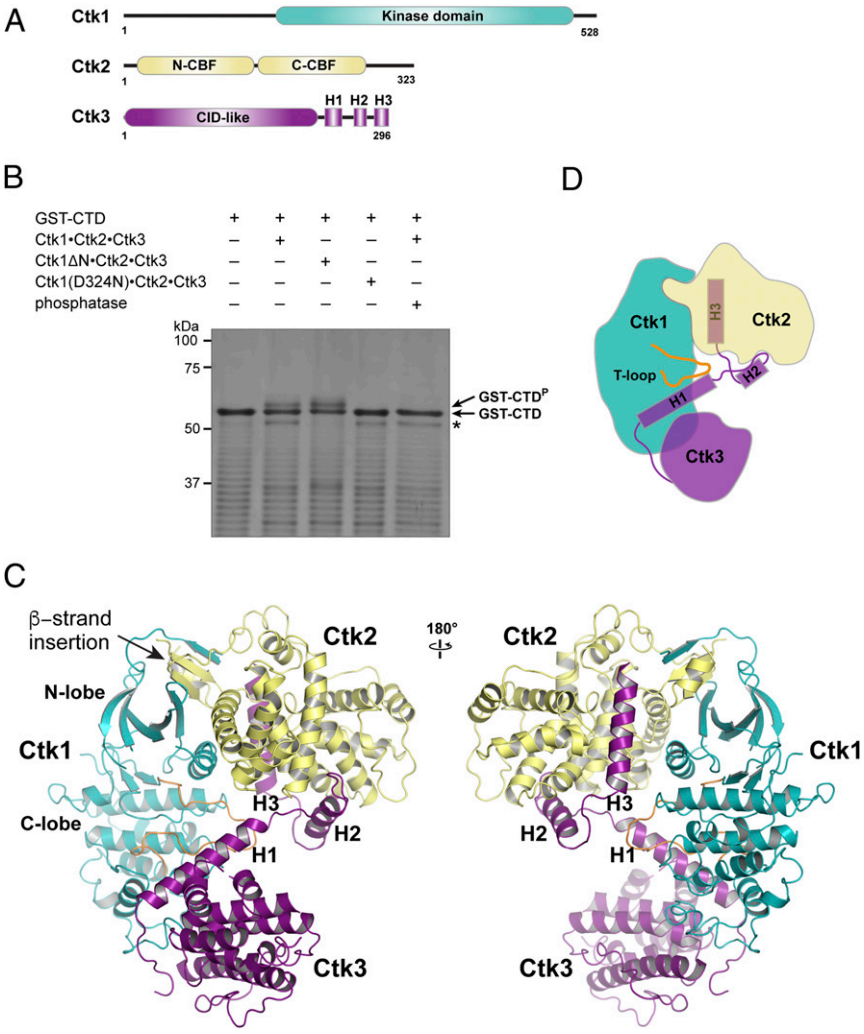

Fig. 1. Structure of a Ctk1•Ctk2•Ctk3 complex. (*A*) Schematic representation of the CTDK-1 subunits. (*B*) Identification of a catalytically active CTDK-1 core complex. GST-tagged yeast Pol II CTD (10 μM) was incubated with the indicated CTDK-1 complex (0.25 μM) and 5 mM ATP at 30 °C for 30 min. CTDK-1 activity was shown by a retarded band on SDS/PAGE, corresponding to phosphorylated GST-CTD. An asterisk denotes the stained band of Ctk1 or the Ctk1(D324N) mutant. (*C*) Crystal structure of the Ctk1ΔN•Ctk2•Ctk3 complex determined at 1.85-Å resolution. Ctk1ΔN, Ctk2, and Ctk3 are colored in cyan, yellow, and purple, respectively. The T-loop of Ctk1 is colored in orange. (*D*) Schematic diagram of the Ctk1ΔN•Ctk2•Ctk3 complex.

structure was determined by molecular replacement single-wavelength anomalous diffraction (MR-SAD) using a human CDK12•Cyclin K complex as a search model and anomalous signals obtained from an Hg-derivative crystal. The Ctk1ΔN•Ctk2•Ctk3 model was refined to 1.85-Å resolution with R_{work} and R_{free} values of 18.4% and 22.2%, respectively (*SI Appendix*, Table S2). The structure of Ctk1ΔN•Ctk2•Ctk3 readily reveals the role of Ctk3 in CTDK-1 architecture (Fig. 1 *C* and *D* and *SI Appendix*, Fig. S2). The Ctk1 kinase and the Ctk2 cyclin feature canonical domain arrangements like other CDKs and cyclins with Ctk1 in a nucleotide-free form. Ctk3 has extensive interactions with both Ctk1 and Ctk2. The N-terminal region of Ctk3 folds into a helical bundle that bears similarity to the CID domain. This CID-like domain is connected to three extended helices, H1, H2, and H3. The CID-like domain and H1 bind to Ctk1 with H1 in direct contact with the T-loop of Ctk1, whereas H2 and H3 bind at the interface of two Ctk2 cyclin-box fold (CBF) domains, N-CBF and C-CBF. Overall, ∼1,400 Å^2 and 1,900 Å^2 of interface area are employed for the Ctk3–Ctk1 and Ctk3–Ctk2 interactions, respectively.

Ctk3 Stabilizes an Active Conformation of the Ctk1 T-Loop. Comparison of the Ctk1ΔN•Ctk2•Ctk3 structure with CDK12•CyclinK reveals that the Ctk1 T-loop is in an active conformation (Fig. 2 *A*–*C*) (25, 26). The T-loop has a well-characterized role of regulating the CDK kinase activity in a manner dependent on the binding of cyclin and the phosphorylation state of the conserved threonine residue (T338 in Ctk1) (24). In our structure, the T-loop of Ctk1 is well ordered and the electron-density maps at 1.85-Å resolution allow unambiguous assignment of the residue side chains (*SI Appendix*, Fig. S3). The conformation of the T-loop resembles that of the active CDK12, particularly at the N-terminal conserved DFG motif at the ATP binding pocket and the C-terminal segment featuring the phosphorylation regulated T338 (Fig. 2 *A*–*C*). Ctk1 and CDK12 diverge at the tip of the T-loop distal to the ATP binding pocket, presumably because the former contains a shorter T-loop by two residues. Serendipitously, a citrate molecule was cocrystallized with Ctk1ΔN•Ctk2•Ctk3 near the position of phosphorylated T893 (TPO893) in CDK12 (*SI Appendix*, Fig. S3). This citrate molecule appears to serve the role of a phosphate group by forming salt bridges with two arginine residues R305 and R329. Importantly, our structure indicates that Ctk3 plays a key role in stabilizing the T-loop through its H1 helix. Specifically, H238, R235, and E232 of Ctk3 form hydrogen bonds or salt bridges with D336, N339, and R340 of the T-loop, respectively. These interactions are in immediate proximity to the phosphorylation regulated T338 residue and stabilize the T-loop in an active conformation. In addition, K242 on H1 is in proximity to T338 and may contribute to a positively charged environment to accommodate phosphorylated T338.

The Ctk3–Ctk1 interface at the T-loop is reinforced by a continuous hydrophobic interaction network mediated by the N-terminal segment of H1 and the N-terminal CID-like domain of Ctk3 (Fig. 2 *A* and *D*). At the N-terminal segment of H1, L227 and M231 interact with W413, M416, and I417 of Ctk1. The CID-like domain employs two C-terminal helices (α6 and α7) to bind a hydrophobic surface (Y409, F414, F415, M418, and P419) at the tip of the Ctk1 C-lobe (Fig. 2*D* and *SI Appendix*, Fig. S4). The extensive hydrophobic interaction network makes significant contributions to anchor the Ctk3 H1 helix in close proximity to the T-loop. Taken together, the CTDK-1 structure data demonstrate a regulatory mechanism by which Ctk3, through extensive interactions with Ctk1, primes the Ctk1 kinase by promoting the T-loop in an active conformation.

It is worth noting that although Ctk3 overall exhibits low sequence homology, the H1 helix is the most conserved segment within the protein (*SI Appendix*, Fig. S2). The charged residues on H1 mediating Ctk1 interactions (E232, D234, R235, and H238) are identical among Ctk3 proteins in *Saccharomyces cerevisiae*, *S. pombe*, *Candida glabrata*, and *Komogataella pastoris*. Regarding the CID-like domain, the overall folding is similar between *S. cerevisiae* Ctk3 in our structure and *S. pombe* Ctk3 (Lsg1) (10), but the α7 helix of *S. pombe* Ctk3 is shorter and lacks some of the key residues (F202, F203, and F206) involved in Ctk1 binding (*SI Appendix*, Fig. S4). The Ctk3 orthologs in *C. glabrata* and *K. pastoris* likely interact with Ctk1 in a similar manner to *S. cerevisiae* Ctk3 based on the sequence alignment (*SI Appendix*, Fig. S2). In all these species, the H1 helix contains a high percentage of identical residues. This highly conserved H1 helix highlights the critical role of Ctk3, which, as suggested by our structure, is to stabilize an active conformation of Ctk1.

Ctk3–Ctk2 Interface. Ctk3 binds to the Ctk2 cyclin through its H2 and H3 extended helices. H2 and H3 form a hairpin at the interface of the Ctk2 N-CBF and C-CBF. The H2 helix packs against a pair of helices in C-CBF (Fig. 3*A*). Specifically, hydrophobic residues F261, L264, W265, and F268 line the interface with Ctk2, and at the C terminus of H2, E269 forms salt bridges with R162 of Ctk2. In addition, the extended loop connecting H1 and H2 makes significant contacts with Ctk2.

H3 of Ctk3 is sandwiched between Ctk1 and Ctk2 mainly through interactions with the N-CBF of Ctk2 (Fig. 3*B*). This interface is dominated by hydrophobic interactions involving Y280, I283, L286, A290, and Y294 of H3. Several residues of H3 also interface with Ctk1 (*SI Appendix*, Fig. S2), with D278 forming a salt bridge with the K330 on the Ctk1 T-loop. Intriguingly, H3 is reminiscent of an N-terminal helix of Cyclin A in the CDK2•Cyclin A complex that contributes to the CDK2–Cyclin A interaction (Fig. 3*C*) (27). Ctk2, as well as its human ortholog Cyclin K, lack this N-terminal helix in Cyclin A. Interestingly, a recent structure of the human CDK-activating kinase CDK7•Cyclin H•MAT1 complex shows that a C-terminal helix of the regulatory subunit MAT1 occupies a similar position in proximity to the CDK7 T-loop (28). Our results demonstrate that H3 of Ctk3 likely contributes, in addition to the CID-like domain and H1, to the observed conformation of the Ctk1 T-loop.

A Ctk1–Ctk2 Interface. The Ctk1ΔN•Ctk2•Ctk3 structure also reveals an additional interface between the Ctk1 kinase and the Ctk2 cyclin (Fig. 4) as compared to CDK12•Cyclin K and other CDK complexes. CDKs have a two-lobed structure. The N-lobe of CDKs typically is comprised of an antiparallel five-stranded β-sheet. Interestingly, the canonical β-sheet of the Ctk1 N-lobe is extended by two β-strands inserted by Ctk2, and by another β-strand contributed by the N terminus of Ctk1, together forming an expanded eight-stranded sheet. The β-strand insertion by Ctk2 is formed by the very N terminus of Ctk2 and the C terminus of Ctk2. The C-terminal Ctk2 β-strand is connected to an α-helix situated at the concave face of the intermolecular β-sheet. Within this Ctk2 helix, E313 forms hydrogen bonds with N202 and N204 of the β-sheet, and L307 and I310 mediates hydrophobic interactions with the β-sheet. Together, the above Ctk1–Ctk2 interactions generate a significantly expanded CDK-cyclin interface (∼1,900 Å^2) compared to that of CDK12•Cyclin K (∼980 Å^2).

CTDK-1 Interacts Physically and Genetically with the SR Protein Gbp2. Ctk1 has a poorly understood N-terminal extension, which contains seven SR or RS dipeptides. Similarly, CDK12 and CDK13 contain similar N-terminal extensions of SR domains (13). As SR domains are typically found in SR splicing factors, we speculate that the N-terminal extension of Ctk1 may be involved in mRNA biogenesis. Ctk1 isolated from yeast cells associates with the SR protein Gbp2 (17). Gbp2 is one of the three SR proteins in yeast and has been implicated in the quality control of splicing (29). We sought to examine whether CTDK-1 directly binds to Gbp2 using in vitro GST pull-down assays with purified recombinant

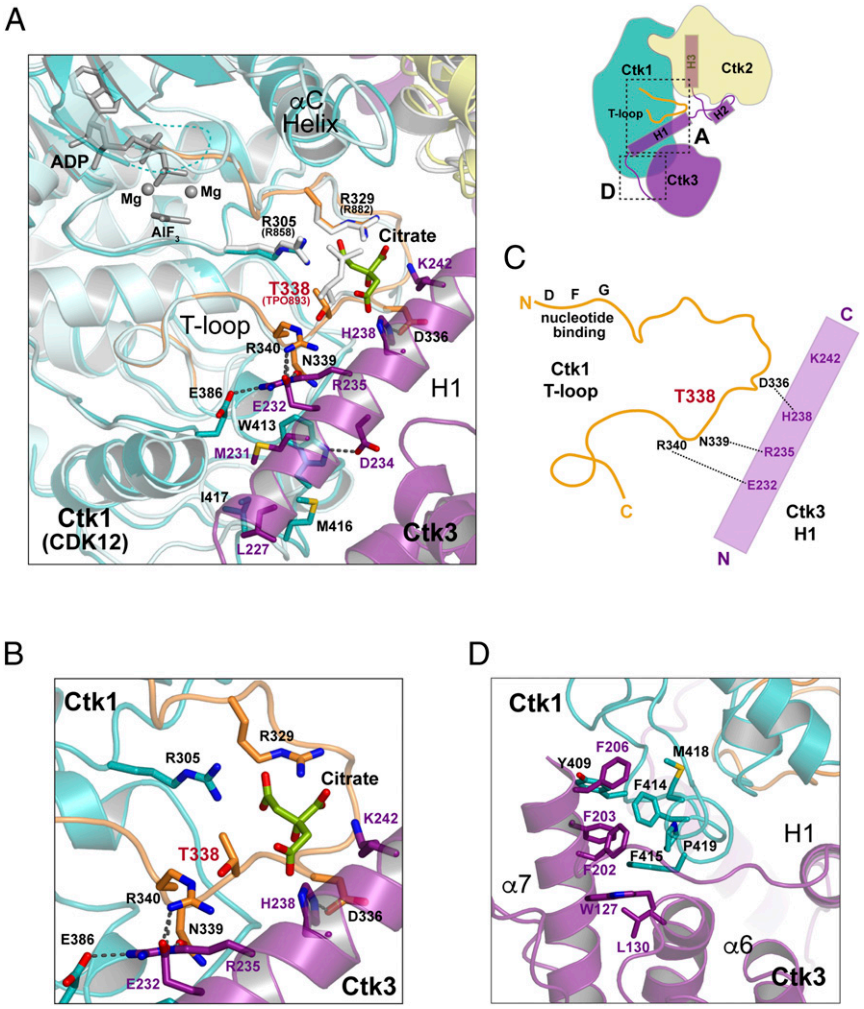

Fig. 2. Ctk3 stabilizes the T-loop of Ctk1 in an active conformation. (*A*) The conformation of the Ctk1 T-loop (orange) resembles that of an active CDK12•Cyclin K complex (PDB ID code 4NST). Key residues at the Ctk1–Ctk3 interface are shown as sticks. A citrate molecule (green) is located near T338. In CDK12•Cyclin K, selective residues, including the phosphorylated T893 (TPO893), are shown as white sticks. The nucleotide ADP•AlF_3 bound to CDK12 is colored gray. (*B*) Close-up view of the Ctk1 T-loop. (*C*) Schematics of the interactions between the Ctk1 T-loop and the Ctk3 H1 helix. (*D*) The interface between the CID-like domain of Ctk3 and Ctk1.

proteins. Gbp2 contains an N-terminal SR domain followed by three RRM domains. We used a truncated Gbp2ΔN (residues 107 to 427) consisting of the three RRM domains, as full-length Gbp2 expressed in *Escherichia coli* was not soluble. We found that purified Gbp2ΔN directly binds to the Ctk1•Ctk2•Ctk3 complex (Fig. 5*A*). Interestingly, their interaction requires the N-terminal extension of Ctk1, as Gbp2ΔN does not bind to the Ctk1ΔN•Ctk2•Ctk3 complex under our GST pull-down conditions. To examine the functional relevance of the interaction between CTDK-1 and Gbp2, we tested whether they have genetic interactions (Fig. 5*B*). We found that the *ctk1Δ gbp2Δ* double-knockout strain exhibited a more severe growth defect compared to either *gbp2Δ*, which exhibits normal growth, or to *ctk1Δ*, which exhibits a more mild growth defect. Together, our results demonstrate that the CTDK-1 complex physically interacts with the SR protein Gbp2 in a manner dependent on the N-terminal extension of Ctk1, and that deletion of both *CTK1* and *GBP2* in yeast results in a synthetic growth defect.

Discussion

We determined the CTDK-1 structure to reveal the molecular architecture of this trimeric cyclin-dependent kinase complex, and importantly, how the unique Ctk3 subunit contributes to the assembly of CTDK-1 and promotes an active conformation of the T-loop of the Ctk1 kinase. We show that Ctk3 reinforces the interaction of the Ctk1 kinase and the Ctk2 cyclin through extensive interactions with both subunits, featuring ~3,300 $Å^2$ of interface area in total. Most prominently, the T-loop of Ctk1, particularly the residues in immediate proximity to the phosphorylation regulated T338 residue, are stabilized in an active conformation through interactions with the most conserved segment of Ctk3. Together, the data of our CTDK-1 structure reveals a regulatory mechanism of CDK complexes.

The most well-documented function of the CTDK-1 complex is to phosphorylate the RNA Pol II CTD during transcription elongation. We show the evidence for a direct interaction between CTDK-1 and the SR protein Gbp2 (Fig. 5*A*). Gbp2 binds to pre-mRNAs and the spliceosome during splicing and is implicated in splicing quality control (29). Gbp2 isolated from yeast associates with a core mRNA export machinery, the TREX complex (30, 31). TREX prepares export of nuclear mRNAs by recruiting the principal mRNA export receptor NXF1•NXT1, which facilitates mRNA translocation through the nuclear pore complex to the cytoplasm for translation (32–34). Our studies

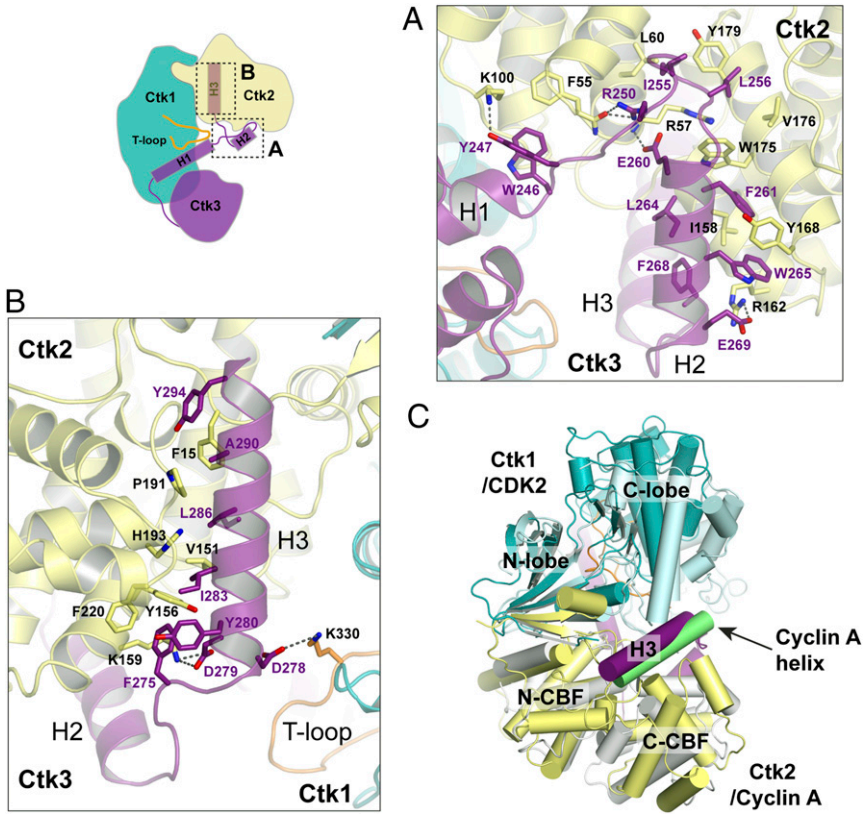

Fig. 3. The Ctk2–Ctk3 interface. (*A*) The binding interface between the H2 helix of Ctk3 and Ctk2. (*B*) The binding interface involving the H3 helix of Ctk3. D278 of H3 forms a salt bridge with K330 on the T-loop of Ctk1. (*C*) Alignment of Ctk1ΔN•Ctk2•Ctk3 and CDK2•Cyclin A (PDB ID code 1JST). Ctk1ΔN•Ctk2•Ctk3 is colored as in Fig. 1*C*. CDK2 and Cyclin A are colored light cyan and white, respectively. The location of the H3 helix of Ctk3 resembles that of the N-terminal helix of Cyclin A (green).

(Fig. 5*B*), as well as those of others (17), demonstrate that CTDK-1 genetically interacts with Gbp2 and the TREX complex. While the synthetic growth defect of the *ctk1Δ gbp2Δ* double mutant could indicate a parallel function of *CTK1* and *GBP2*, together with the physical interaction between CTDK-1 and Gbp2 demonstrated by both in vitro GST-pull down assays (Fig. 5*A*) and immunoprecipitation from yeast lysates (17), our results more likely suggest a link of CTDK-1 function to Gbp2-associated splicing/mRNA export events. Furthermore, the interaction between CTDK-1 and Gbp2 requires the N-terminal extension of Ctk1 (Fig. 5*A*). The N-terminal extension of Ctk1, as well as its human orthologs CDK12 and CDK13, contains SR domains that are typically found in SR splicing factors. Intriguingly, CDK11, another CDK that contains a long N-terminal extension with SR domains, has been shown to associate with the TREX complex (13, 35). Thus, it appears that the N-terminal extension in CDKs may correlate with their connections to nuclear mRNA biogenesis processes. It is worth noting that the evolution of CDKs and SR proteins are both divergent. Yeast have only three SR-like proteins, whereas humans have a dozen, SRSF1 to SRSF12 (14, 36). We speculate that the mechanism of CTDK-1 in yeast is a prototype for yet-to-be-defined molecular connections between CDKs and SR proteins in humans.

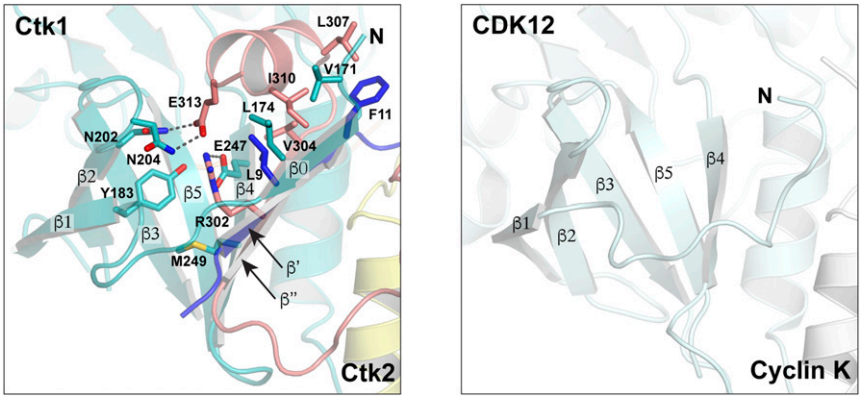

Fig. 4. A Ctk1–Ctk2 interface. Comparison of Ctk1ΔN•Ctk2•Ctk3 (*Left*) and its human ortholog CDK12•Cyclin K (right, PDB ID code 4NST). Ctk2 inserts two strands (β′ and β′′) into the β-sheet of the Ctk1 N-lobe, forming an intermolecular β-sheet. Residues at the Ctk1–Ctk2 interface are shown as sticks.

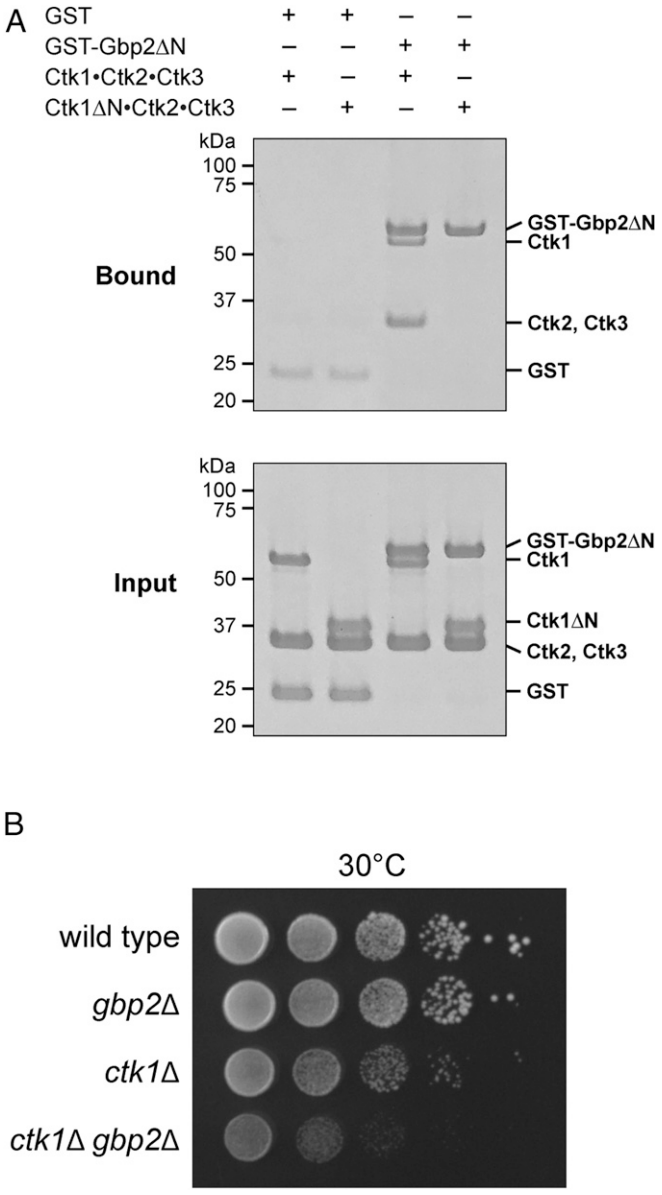

Fig. 5. Ctk1 interacts with the SR protein Gbp2 physically and genetically. (*A*) Ctk1•Ctk2•Ctk3 directly interacts with Gbp2. In vitro GST pull-down assays were performed with GST or GST-tagged Gbp2ΔN and the Ctk1•Ctk2•Ctk3 complex as indicated. (*B*) Serial dilutions of yeast demonstrating a synthetic growth defect of *ctk1Δ* and *gbp2Δ* in yeast.

Materials and Methods

Plasmids. Ctk1 (residues 1 to 508) or Ctk1ΔN (residues 159 to 508) was cloned into the pFastBac HTb vector (Thermo Fisher Scientific) for expression in insect cells with an N-terminal His tag. Full-length Ctk2 was cloned into a pFastBac Dual vector (Thermo Fisher Scientific) modified to contain a tobacco etch virus (TEV) protease cleavable N-terminal GST tag. Full-length Ctk3 was cloned into the pFastBac Dual vector for expression as an untagged protein. Gbp2ΔN (residues 107 to 427) and the CTD of the largest RNA Pol II subunit Rpb1 (residues 1549 to 1716) were cloned in a pGEX-4T-1 vector (GE Healthcare) modified to contain a TEV protease cleavable N-terminal GST tag.

Protein Expression and Purification. The yeast Ctk1•Ctk2•Ctk3 complexes were expressed in High Five insect cells (Thermo Fisher Scientific) by coinfection of recombinant baculoviruses expressing Ctk1, Ctk2, and Ctk3. For the kinase assay, CTDK-1 subunits were also coexpressed with the *S. pombe* CDK-activating kinase Csk1. The baculovirus encoding a full-length Csk1 with a Flag tag and a hexahistidine tag at the C terminus was a kind gift from Robert Fisher, Icahn School of Medicine at Mount Sinai, New York, NY (37). High-Five cells were harvested 48 h after infection. The cells were lysed with a cell disruptor (Avestin) in lysis buffer (50 mM Tris, pH 8.0, 500 mM NaCl, 0.5 mM TCEP, and a protease inhibitor mix containing 1 mM PMSF, 5 mg/L aprotinin, 1 mg/L pepstatin, and 1 mg/L leupeptin). CTDK-1 was first purified using Ni affinity chromatography (GE Healthcare). The eluted protein was then loaded onto a Glutathione Sepharose column (GE Healthcare), followed by on-column overnight digestion with TEV protease at 4 °C. CTDK-1 protein was obtained in the flow-through of the Glutathione Sepharose column, and was loaded onto a Superdex 200 column (GE Healthcare) equilibrated with 10 mM Tris, pH 8.0, 500 mM NaCl, and 0.5 mM TCEP.

The GST-Gbp2ΔN and GST-CTD proteins were expressed in *E. coli* Rosetta cells (Sigma-Aldrich). Protein expression was induced by 0.5 mM isopropyl-β-D-thiogalactopyranoside (IPTG) at 20 °C overnight. Cells were lysed in lysis buffer and purified using a Glutathione Sepharose column. GST-CTD was applied on a mono Q column (GE Healthcare) and eluted using a 50 mM to 300 mM NaCl gradient in a buffer containing 20 mM Tris, pH 8.0 and 0.5 mM TCEP. GST-Gbp2ΔN was purified on a Superdex 200 column equilibrated with 10 mM Tris, pH 8.0, 300 mM NaCl, and 0.5 mM TCEP.

All purified proteins were aliquoted and stored at −80 °C.

Crystallization and Structure Determination. Crystals of Ctk1ΔN•Ctk2•Ctk3 were obtained at 20 °C by vapor diffusion in sitting drops using 0.5 μL protein complex and 0.5 μL reservoir solution containing 0.1 M sodium citrate pH 5.3, 0.3 M lithium chloride, and 17.5% PEG3350. Hg-derivative crystals were obtained by soaking the crystals with the reservoir solution supplemented with 1 mM Thimerosal for 16 h. The crystals were transferred in three steps of increasing glycerol concentration to a cryoprotectant solution containing 0.1 M sodium citrate pH 5.3, 0.3 M lithium chloride, 17.5% PEG3350, and 20% glycerol. X-ray diffraction data were collected at the 21-ID-D Beamline at the Advanced Photon Source, Argonne National Laboratory.

X-ray intensities were processed using the HKL2000 package (38). An initial electron density map was determined by MR-SAD using Phenix Autosol (39). The human CDK12•CyclinK (PDB ID code 4UN0) structure was used as a molecular replacement model. Anomalous signals were obtained from a Hg-derivative crystal. The anomalous signals extended to 3.8 Å, and 24 Hg sites were found. The overall Figure of Merit from MR-SAD phasing is 0.45, generating clearly interpretable density maps. Several rounds of modeling building in Coot (40) and refinement with phenix.refine (41) led to a Ctk1ΔN•Ctk2•Ctk3 model at 1.85-Å resolution. The Ctk1ΔN•Ctk2•Ctk3 model consists of Ctk1 (residues 170 to 191, 196 to 214, and 218 to 480), Ctk2 (residues 5 to 314), and Ctk3 (residues 5 to 21, 47 to 93, 103 to 207, and 215 to 295). Details of the data collection and refinement statistics are in *SI Appendix*, Table S2. Figures are prepared using PyMOL (Molecular Graphics System, Schrodinger).

GST Pull-Down Assays. For GST pull-down assays, 4 μM of GST or GST-tagged Gbp2ΔN was incubated with 2 μM of CTDK-1 complex as indicated in the binding buffer (20 mM Hepes pH 7.0, 125 mM NaCl, and 0.5 mM TCEP) at room temperature for 10 min. The reaction mixtures were then added into ∼15 μL glutathione resin in an Eppendorf tube and binding was allowed to proceed on ice for 30 min with gentle tapping to mix every 5 min. Beads were washed twice with 500 μL binding buffer and bound proteins were eluted and analyzed using Coomassie-stained SDS/PAGE gels. The experiments were repeated three times independently.

Kinase Assays. Kinase assays were conducted with 0.25 μM of CTDK-1 and 10 μM of purified GST tagged CTD in a reaction buffer containing 175 mM NaCl, 3 mM $MgCl_2$, 1 μM Na_3VO_4, and 5 mM ATP. The reactions were conducted at 30 °C for 30 min. Phosphatase treatment was conducted with Lambda phosphatase kit (New England Biolabs). Reactions were stopped by boiling in SDS/PAGE sample buffer. Phosphorylation was observed as a shift in mobility by Coomassie-stained SDS/PAGE. The experiments in Fig. 1*B* and *SI Appendix*, Fig. S1*B* were independently repeated three times and two times, respectively.

Yeast Strains and Spot Assay. The *ctk1Δ* strain was generated by initially transforming BY4743 with a PCR product derived using pFA6-KanMX to delete *CTK1*, while *gbp2Δ* strain was generated by initially transforming BY4743 with a PCR product derived using DHFR-3 to delete *GBP2*. Each resulting diploid strain was then sporulated and dissected to produce haploid versions of *ctk1Δ* (SWY6773, *Mat a his3Δ1 leu2Δ0 ura3Δ0 lys2Δ1*

met15Δ0 ctk1Δ::KAN) and *gbp2Δ* (SWY6777, *Mat alpha his3Δ1 leu2Δ0 ura3Δ0 lys2Δ1 gbp2Δ::HygB*) strains, which were verified using PCR. The *ctk1Δ gbp2Δ* strain was generated by initially crossing *ctk1Δ* and *gbp2Δ* and then transforming the resulting diploid strain with a *CEN CTK1-URA3* vector. The diploid was sporulated and dissected to isolate Kan- and Hyg-resistant spores. The *CTK1-URA3* vector was eliminated from the *ctk1Δ gbp2Δ* strain (SWY6798, *Mat alpha leu2Δ0 ura3Δ0 his3Δ1 lys2Δ1 ctk1Δ::KAN gbp2Δ::HygB*) by growing the cells on 5-FOA for two rounds.

For the spot assay, yeast cells were grown in YPD until midlog phase, then washed with Milli-Q water and resuspended at an OD600 1.0. After washing, 10-fold serial dilutions were prepared, and 5 μL of each was spotted on a YPD plate and grown at 30 °C. The experiments were repeated three times independently.

Data Availability. The atomic coordinates of the CTDK-1 crystal structure have been deposited in the RCSB Protein Data Bank, https://www.rcsb.org/ (RCSB ID code 7JV7).

ACKNOWLEDGMENTS. We thank Robert Fisher for a gift of the recombinant baculovirus encoding *Schizosaccharomyces pombe* Csk1; Renee Dawson for helpful discussions and comments on the manuscript; the staff at Advanced Photon Source beamline 21-ID-D for help with X-ray data collection; and the proteomics core of the Vanderbilt Mass Spectrometry Research Center for assistance. This work was supported by National Institute of General Medical Sciences Grants 5R35GM133743 (to Y.R.) and 5R37GM051219 (to S.R.W.), and funds from Vanderbilt University School of Medicine. B.P.C. is supported by NIH/National Cancer Institute Training Grant 5T32CA119925.

1. P. K. Parua, R. P. Fisher, Dissecting the Pol II transcription cycle and derailing cancer with CDK inhibitors. *Nat. Chem. Biol.* **16**, 716–724 (2020).
2. J. P. Hsin, J. L. Manley, The RNA polymerase II CTD coordinates transcription and RNA processing. *Genes Dev.* **26**, 2119–2137 (2012).
3. S. Buratowski, Progression through the RNA polymerase II CTD cycle. *Mol. Cell* **36**, 541–546 (2009).
4. E. J. Cho, M. S. Kobor, M. Kim, J. Greenblatt, S. Buratowski, Opposing effects of Ctk1 kinase and Fcp1 phosphatase at Ser 2 of the RNA polymerase II C-terminal domain. *Genes Dev.* **15**, 3319–3329 (2001).
5. S. H. Ahn, M. Kim, S. Buratowski, Phosphorylation of serine 2 within the RNA polymerase II C-terminal domain couples transcription and 3′ end processing. *Mol. Cell* **13**, 67–76 (2004).
6. D. E. Sterner, J. M. Lee, S. E. Hardin, A. L. Greenleaf, The yeast carboxyl-terminal repeat domain kinase CTDK-I is a divergent cyclin-cyclin-dependent kinase complex. *Mol. Cell. Biol.* **15**, 5716–5724 (1995).
7. J. M. Lee, A. L. Greenleaf, CTD kinase large subunit is encoded by CTK1, a gene required for normal growth of Saccharomyces cerevisiae. *Gene Expr.* **1**, 149–167 (1991).
8. H. Qiu, C. Hu, A. G. Hinnebusch, Phosphorylation of the Pol II CTD by KIN28 enhances BUR1/BUR2 recruitment and Ser2 CTD phosphorylation near promoters. *Mol. Cell* **33**, 752–762 (2009).
9. G. Hautbergue, V. Goguel, Activation of the cyclin-dependent kinase CTDK-I requires the heterodimerization of two unstable subunits. *J. Biol. Chem.* **276**, 8005–8013 (2001).
10. W. Mühlbacher *et al*., Structure of Ctk3, a subunit of the RNA polymerase II CTD kinase complex, reveals a noncanonical CTD-interacting domain fold. *Proteins* **83**, 1849–1858 (2015).
11. B. Bartkowiak *et al*., CDK12 is a transcription elongation-associated CTD kinase, the metazoan ortholog of yeast Ctk1. *Genes Dev.* **24**, 2303–2316 (2010).
12. A. L. Greenleaf, Human CDK12 and CDK13, multi-tasking CTD kinases for the new millenium. *Transcription* **10**, 91–110 (2019).
13. M. Malumbres, Cyclin-dependent kinases. *Genome Biol.* **15**, 122 (2014).
14. J. L. Manley, A. R. Krainer, A rational nomenclature for serine/arginine-rich protein splicing factors (SR proteins). *Genes Dev.* **24**, 1073–1074 (2010).
15. K. Liang *et al*., Characterization of human cyclin-dependent kinase 12 (CDK12) and CDK13 complexes in C-terminal domain phosphorylation, gene transcription, and RNA processing. *Mol. Cell. Biol.* **35**, 928–938 (2015).
16. S. Battaglia *et al*., RNA-dependent chromatin association of transcription elongation factors and Pol II CTD kinases. *eLife* **6**, e25637 (2017).
17. E. Hurt, M. J. Luo, S. Röther, R. Reed, K. Strässer, Cotranscriptional recruitment of the serine-arginine-rich (SR)-like proteins Gbp2 and Hrb1 to nascent mRNA via the TREX complex. *Proc. Natl. Acad. Sci. U.S.A.* **101**, 1858–1862 (2004).
18. M. Windgassen, H. Krebber, Identification of Gbp2 as a novel poly(A)+ RNA-binding protein involved in the cytoplasmic delivery of messenger RNAs in yeast. *EMBO Rep.* **4**, 278–283 (2003).
19. D. Blazek *et al*., The Cyclin K/Cdk12 complex maintains genomic stability via regulation of expression of DNA damage response genes. *Genes Dev.* **25**, 2158–2172 (2011).
20. T. S. Winsor, B. Bartkowiak, C. B. Bennett, A. L. Greenleaf, A DNA damage response system associated with the phosphoCTD of elongating RNA polymerase II. *PLoS One* **8**, e60909 (2013).
21. B. Coordes *et al*., Ctk1 function is necessary for full translation initiation activity in Saccharomyces cerevisiae. *Eukaryot. Cell* **14**, 86–95 (2015).
22. S. H. Choi *et al*., CDK12 phosphorylates 4E-BP1 to enable mTORC1-dependent translation and mitotic genome stability. *Genes Dev.* **33**, 418–435 (2019).
23. K. M. Lee, J. E. Saiz, W. A. Barton, R. P. Fisher, Cdc2 activation in fission yeast depends on Mcs6 and Csk1, two partially redundant Cdk-activating kinases (CAKs). *Curr. Biol.* **9**, 441–444 (1999).
24. D. Ostapenko, M. J. Solomon, Phosphorylation by Cak1 regulates the C-terminal domain kinase Ctk1 in Saccharomyces cerevisiae. *Mol. Cell. Biol.* **25**, 3906–3913 (2005).
25. S. E. Dixon-Clarke, J. M. Elkins, S. W. Cheng, G. B. Morin, A. N. Bullock, Structures of the CDK12/CycK complex with AMP-PNP reveal a flexible C-terminal kinase extension important for ATP binding. *Sci. Rep.* **5**, 17122 (2015).
26. C. A. Bösken *et al*., The structure and substrate specificity of human Cdk12/Cyclin K. *Nat. Commun.* **5**, 3505 (2014).
27. A. A. Russo, P. D. Jeffrey, N. P. Pavletich, Structural basis of cyclin-dependent kinase activation by phosphorylation. *Nat. Struct. Biol.* **3**, 696–700 (1996).
28. B. J. Greber *et al*., The cryoelectron microscopy structure of the human CDK-activating kinase. *Proc. Natl. Acad. Sci. U.S.A.* **117**, 22849–22857 (2020).
29. A. Hackmann *et al*., Quality control of spliced mRNAs requires the shuttling SR proteins Gbp2 and Hrb1. *Nat. Commun.* **5**, 3123 (2014).
30. K. Strässer *et al*., TREX is a conserved complex coupling transcription with messenger RNA export. *Nature* **417**, 304–308 (2002).
31. S. Masuda *et al*., Recruitment of the human TREX complex to mRNA during splicing. *Genes Dev.* **19**, 1512–1517 (2005).
32. S. R. Carmody, S. R. Wente, mRNA nuclear export at a glance. *J. Cell Sci.* **122**, 1933–1937 (2009).
33. M. Stewart, Nuclear export of mRNA. *Trends Biochem. Sci.* **35**, 609–617 (2010).
34. Y. Xie, Y. Ren, Mechanisms of nuclear mRNA export: A structural perspective. *Traffic* **20**, 829–840 (2019).
35. V. Pak *et al*., CDK11 in TREX/THOC regulates HIV mRNA 3′ end processing. *Cell Host Microbe* **18**, 560–570 (2015).
36. R. Reed, H. Cheng, TREX, SR proteins and export of mRNA. *Curr. Opin. Cell Biol.* **17**, 269–273 (2005).
37. S. Larochelle *et al*., Dichotomous but stringent substrate selection by the dual-function Cdk7 complex revealed by chemical genetics. *Nat. Struct. Mol. Biol.* **13**, 55–62 (2006).
38. Z. Otwinowski, W. Minor, Processing of X-ray diffraction data collected in oscillation mode. *Methods Enzymol.* **276**, 307–326 (1997).
39. T. C. Terwilliger *et al*., Decision-making in structure solution using Bayesian estimates of map quality: The PHENIX AutoSol wizard. *Acta Crystallogr. D Biol. Crystallogr.* **65**, 582–601 (2009).
40. P. Emsley, B. Lohkamp, W. G. Scott, K. Cowtan, Features and development of Coot. *Acta Crystallogr. D Biol. Crystallogr.* **66**, 486–501 (2010).
41. P. V. Afonine *et al*., Towards automated crystallographic structure refinement with phenix.refine. *Acta Crystallogr. D Biol. Crystallogr.* **68**, 352–367 (2012).