

ARTICLE

Received 28 Dec 2011 | Accepted 13 Apr 2012 | Published 15 May 2012

DOI: 10.1038/ncomms1846

# Serine-7 but not serine-5 phosphorylation primes RNA polymerase II CTD for P-TEFb recognition

Nadine Czudnochowski<sup>1,\*</sup>, Christian A. Böskén<sup>1,\*</sup> & Matthias Geyer<sup>1</sup>

Phosphorylation of RNA polymerase II carboxy-terminal domain (CTD) in hepta-repeats YSPTSPS regulates eukaryotic transcription. Whereas Ser5 is phosphorylated in the initiation phase, Ser2 phosphorylation marks the elongation state. Here we show that the positive transcription elongation factor P-TEFb is a Ser5 CTD kinase that is unable to create Ser2/Ser5 double phosphorylations, while it exhibits fourfold higher activity on a CTD substrate pre-phosphorylated at Ser7 compared with the consensus hepta-repeat or the YSPTSPK variant. Mass spectrometry reveals an equal number of phosphorylations to the number of hepta-repeats provided, yet the mechanism of phosphorylation is distributive despite the repetitive nature of the substrate. Inhibition of P-TEFb activity is mediated by two regions in Hexim1 that act synergistically on Cdk9 and Cyclin T1. HIV-1 Tat/TAR abrogates Hexim1 inhibition to stimulate transcription of viral genes but does not change the substrate specificity. Together, these results provide insight into the multifaceted pattern of CTD phosphorylation.

<sup>1</sup> Max-Planck-Institut für molekulare Physiologie, Abt. Physikalische Biochemie, Otto-Hahn-Straße 11, 44227 Dortmund, Germany. \*These authors contributed equally to this work. †Present address: Department of Biochemistry and Biophysics, University of California, San Francisco, California 94158, USA. Correspondence and requests for materials should be addressed to M.G. (email: matthias.geyer@mpi-dortmund.mpg.de).

After transcription initiation, RNA polymerase II (RNAPII) pauses approximately 25–50 nucleotides downstream of the transcription start site because of the action of negative transcription elongation factors<sup>1</sup>. Release from this block requires the positive transcription elongation factor P-TEFb<sup>2</sup>, which is a heterodimer composed of the cyclin-dependent kinase Cdk9 and the regulatory subunit Cyclin T. P-TEFb mediates the transition from transcription initiation to productive elongation of pre-mRNA transcripts by phosphorylation of the carboxy-terminal domain (CTD) of the largest subunit of RNAPII. In humans, the CTD comprises 52 repeats of the consensus sequence Y<sub>1</sub>S<sub>2</sub>P<sub>3</sub>T<sub>4</sub>S<sub>5</sub>P<sub>6</sub>S<sub>7</sub> that exhibits only some variations from the strict consensus towards the C-terminus<sup>3–7</sup>. Phosphorylation of Ser5 by Cdk7 and Cdk8 kinases of TFIIF and the mediator complex, respectively, has been described to be concomitant with transcription initiation, whereas Cdk9 of P-TEFb is suggested to phosphorylate Ser2, which marks the elongation phase of transcription<sup>3–8</sup>. In addition, Ser7 of the CTD can also be phosphorylated, which has been linked to small nuclear RNAs (snRNA) transcription and recruitment of the integrator complex<sup>9–13</sup>. Whereas some genes, for example, those on CpG islands, can directly proceed to the elongation state by recruiting P-TEFb to RNAPII<sup>14</sup>, genome-wide studies suggest that the majority of genes in higher eukaryotes are under the control of promoter-proximal pausing<sup>15–17</sup>.

Studies in budding yeast, whose RNAPII contains 26 hepta-repeats of the consensus CTD sequence, have recently identified gene class-specific CTD phosphorylation patterns and a widespread co-occurrence of such CTD marks at various stages of transcription<sup>18</sup>. Particularly, Ser7 phosphorylation is found early in transcription initiation and retained until transcription termination in all RNAPII-dependent genes. In line with these observations, the dynamics of CTD phosphorylations were found to be not scaled to the gene length but differ among genes with different promoter structures and expression levels<sup>19</sup>. High levels of Ser7 and Ser5 phosphorylations occur at the transcription start site, whereas Ser2 phosphorylation levels increase until they reach their peak about 600–1,000 nucleotides downstream of the start site<sup>20</sup>. These modifications suggest a dual gradient model for increasing Ser2 and decreasing Ser5 phosphorylation across genes that regulate transcription<sup>3,4</sup>. In addition to these serine phosphorylations, arginine methylation at one specific residue of the CTD was shown to lead to misregulation of snRNAs as well as small nucleolar RNAs expression<sup>21</sup>.

Here we investigated the specificity and activity of P-TEFb for CTD substrates by analytical means as a complementary technique to the use of antibodies raised against designed CTD epitopes. Using various CTD templates we applied mass spectrometry and kinase activity assays for quantification of CTD modifications. Intriguingly, we found that every single hepta-repeat of a CTD substrate can be phosphorylated one time by P-TEFb, though, P-TEFb is unable to phosphorylate a CTD that has been pre-phosphorylated at Ser5 or Ser2 sites. Phosphorylation of Ser7 instead results in a four-fold higher catalytic activity for such modified templates. Inhibition of P-TEFb activity by its cellular regulation factor Hexim1 requires two distinct regions, a central PYNT motif and a C-terminal cyclin T-binding domain (TBD). Both regions act synergistically on the P-TEFb subunits Cdk9 and Cyclin T1, but are ineffective independently. Addition of 7SK snRNA to Hexim1 significantly increases this inhibitory effect. The HIV-1 Tat–TAR complex abrogates Hexim1 inhibition on P-TEFb to stimulate transcription of viral genes but does not change the substrate specificity. These observations provide insights into the kinetics and substrate specificity of the transcription elongation factor P-TEFb, its phosphorylation signature on the RNAPII CTD and regulation by cellular factors.

## Results

**The number of CTD phosphorylations equals the hepta-repeats.** We set up an *in vitro* phosphorylation assay using recombinant

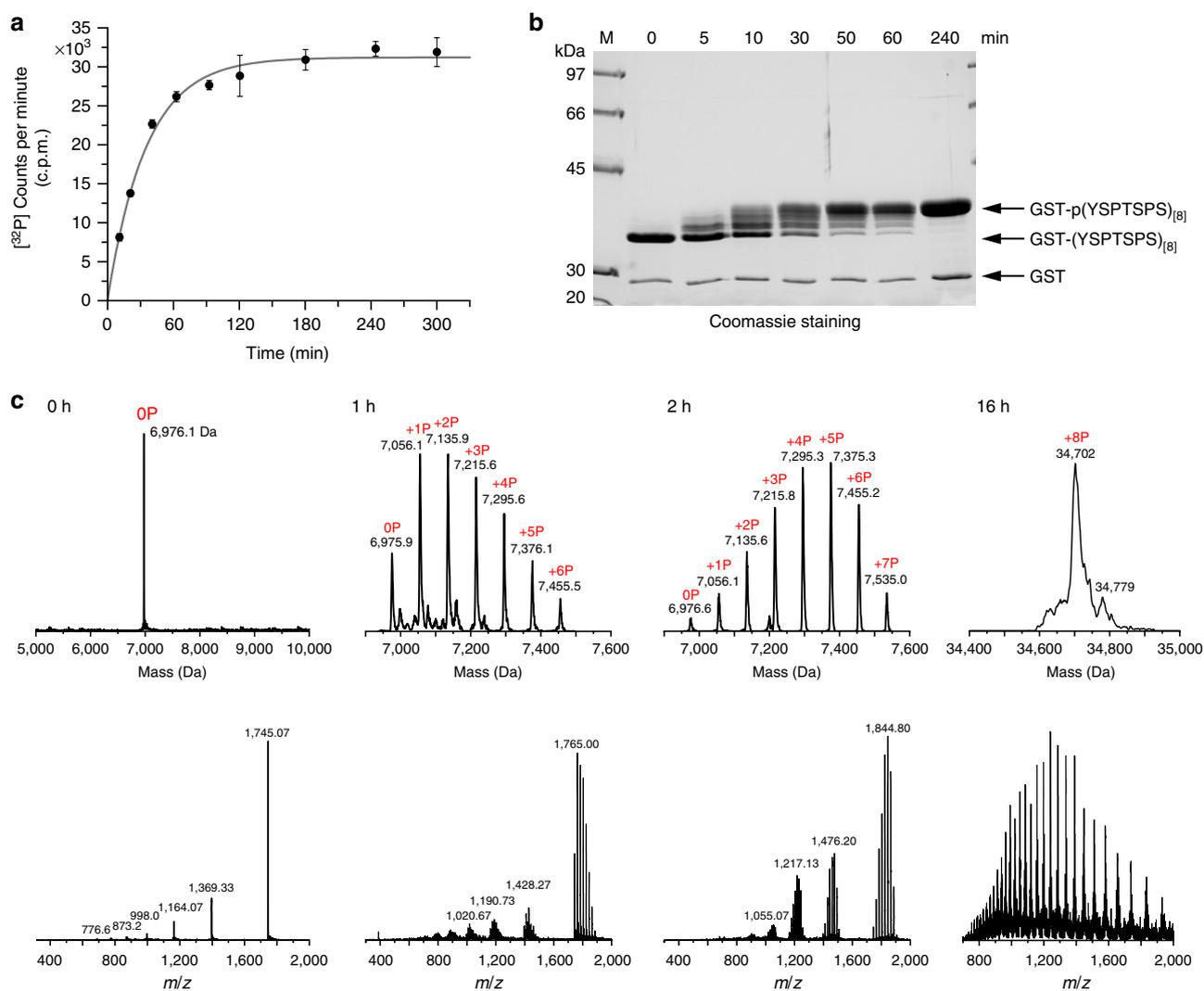
P-TEFb and CTD substrates containing 5 to 13 consensus hepta-repeats or all 52 hepta-repeats of human wild-type CTD (Methods). The homogeneity and functional integrity of the P-TEFb complex was assayed by mass spectrometry, size exclusion chromatography and radioactive filter-binding experiments to confirm its catalytic activity (Supplementary Fig. S1). At a temperature of 30 °C, saturation of the CTD phosphorylation reaction was achieved within several hours of incubation, using twofold molar excess of the co-substrate ATP over the concentration of CTD substrate multiplied by the number of hepta-repeats contained in the substrate (Fig. 1a). The increasing phosphorylation status of a glutathione S-transferase (GST)–CTD fusion protein and the gain in mass upon each phosphorylation is also visible as retained migration behaviour in SDS-polyacrylamide gel electrophoresis (PAGE) analysis (Fig. 1b).

For a quantitative analysis, we used electrospray ionization mass spectrometry (ESI-MS) to determine the number of phosphorylation events at various time points. Towards the endpoint of the reaction under saturating conditions, the number of phosphorylations as determined by mass spectrometry equals the number of hepta-repeats provided in the CTD substrate (Fig. 1c). Although this result may seem trivial, it confirms a high specificity of P-TEFb for the CTD substrate given that three canonical serines are present in each hepta-repeat, two of them as the preferred Ser-Pro recognition motif for Cdk family kinase phosphorylation<sup>22,23</sup>. It also indicates that a continuous phosphorylation signature can be achieved by P-TEFb, leaving no gaps in the modification pattern of the CTD hepta-repeat structure. These results were confirmed by the potent phosphorylation of a full-length CTD containing all 52 hepta-repeats including the acidic C-terminus (Supplementary Fig. S2). Although the highly negatively charged CTD was not accessible to mass spectrometry analysis, the large band shift upon phosphorylation and the passage into saturation suggest full phosphorylation of the substrate.

## A distributive CTD phosphorylation mechanism by PTEfb.

Mass spectra recorded in the course of the reaction provided additional insights into the phosphorylation mechanism (Fig. 1c). Assuming that all phosphorylated CTD substrates are similarly susceptible to ionization, the phosphorylation spectra recorded represent a histogram of the number of molecules containing no, one, two etc., phosphorylations, albeit without any information about the spatial arrangement of these modifications on the periodic template. The asymmetric distribution of the mass spectra recorded at 1 and 2 h after reaction start indicate a binominal distribution function with two boundaries, ‘0’ and ‘8’ in this experiment. From the statistics of mass increase it can be concluded that the mechanism of CTD phosphorylation by P-TEFb is distributive, rather than processive, as might be suspected from the repetitive nature of the substrate. An illustration of a distributive versus a processive reaction mechanism on a periodic matrix as the CTD and the corresponding histograms of mass distributions are shown in Supplementary Fig. S3. However, it cannot be excluded that a consecutive phosphorylation mechanism applies to the mode of P-TEFb action, meaning that once a first phosphorylation mark has been set on a substrate template, the following (upstream or downstream) hepta-repeat will be phosphorylated preferentially due to an increased recognition affinity.

**CTD-pSer7 is the preferred P-TEFb substrate.** To analyse the preference of P-TEFb for particular sequence compositions and phosphorylation signatures, we designed a series of CTD peptides, each containing three hepta-repeats (Fig. 2a). The peptides contained the unmodified consensus CTD sequence (cons. CTD<sub>[3]</sub>), phosphoryl groups at Ser2, Ser5 or Ser7 throughout (pS2-CTD<sub>[3]</sub>, pS5-CTD<sub>[3]</sub> and pS7-CTD<sub>[3]</sub>, respectively) or lysines at position 7 (K7-CTD<sub>[3]</sub>). The hepta-repeat sequence YSPSPK is indeed the prevailing alteration from the consensus CTD in higher eukaryotes, occurring eight



**Figure 1 | P-TEFb phosphorylates the CTD in a distributive mechanism equally to the number of hepta-repeats.** (a) Time course of CTD phosphorylation in a P-TEFb kinase assay. A GST-CTD<sub>[13]</sub> fusion protein containing 13 hepta-repeats (10  $\mu$ M) was incubated with 0.1  $\mu$ M P-TEFb and 300  $\mu$ M [<sup>32</sup>P]- $\gamma$ -labelled ATP. Saturation of the phosphorylation reaction was achieved within 3 h. Data represent the mean  $\pm$  range from two independent experiments. (b) SDS-PAGE analysis of CTD phosphorylation by P-TEFb. At concentrations of 0.1  $\mu$ M P-TEFb, 100  $\mu$ M GST-CTD<sub>[8]</sub> and 3 mM ATP the substrate was fully phosphorylated after 4 h at 30  $^{\circ}$ C. (c) ESI-MS analysis of CTD phosphorylation in a time course experiment. A CTD peptide containing eight hepta-repeats was used as P-TEFb substrate. Shown are time points at the beginning and after 1, 2 and 16 h of the reaction. For the last time point, a GST-CTD<sub>[8]</sub> substrate was used for better ionization properties in the ESI-MS analysis.

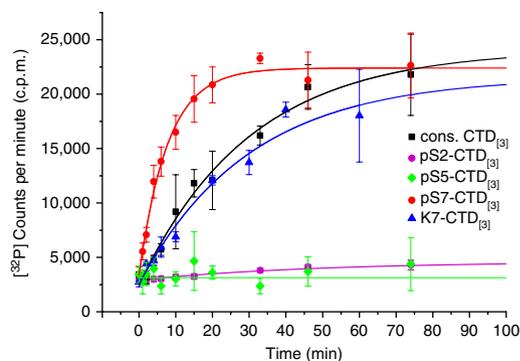
times in human RNAPII. The peptides were marked at the C-terminus with a double arginine motif separated by a polyethylene glycol spacer for better ionization properties in the ESI-MS analysis.

After 4 h incubation with P-TEFb, the consensus CTD substrate showed up to three phosphorylations corresponding to the three hepta-repeats provided as indicated by mass increases (Fig. 2b). Intriguingly, we found that P-TEFb is unable to phosphorylate a CTD substrate that contained complete phosphorylation at Ser5. Likewise, a peptide-containing phosphoryl group at Ser2 was not phosphorylated after 4 h incubation. In sharp contrast, P-TEFb readily phosphorylated a CTD template pre-phosphorylated at position Ser7. The alternate K7 hepta-repeats again were phosphorylated, while phosphorylation of either Tyr1 or the recently identified Thr4 residues<sup>24,25</sup> abrogated P-TEFb activity for these pre-modified substrates (Supplementary Fig. S4). Notably, a similar recognition preference for the phosphorylation of cons. CTD, pS7-CTD and K7-CTD but not pS2- and pS5-CTD was seen for native P-TEFb that contained besides full-length Cdk9 also full-length CycT1 (Supplementary Fig. S5).

**Kinetics of P-TEFb-mediated CTD phosphorylations.** From a time course of the reaction, a  $K_{cat}/K_M$  value of  $2.15 \times 10^4 \text{ M}^{-1} \text{ s}^{-1}$  was determined for the pS7-CTD<sub>[3]</sub> substrate, whereas P-TEFb showed a  $K_{cat}/K_M$  value of  $5.5 \times 10^3 \text{ M}^{-1} \text{ s}^{-1}$  for the consensus CTD using an enzyme concentration of 0.1  $\mu$ M (Fig. 3). Surprisingly, the CTD sequence modified with a lysine residue at position 7 exhibited with a  $K_{cat}/K_M$  value of  $5.2 \times 10^3 \text{ M}^{-1} \text{ s}^{-1}$  similar substrate susceptibility for P-TEFb as the consensus sequence. Using longer CTD templates of 9 or 13 repeats uniformly modified with lysines at position 7, as found, for example, in *Plasmodium falciparum*<sup>5</sup>, we noticed, however, a retarded phosphorylation activity of P-TEFb. Such effect could result from a changed accessibility of the serines, possibly caused by ionic interactions between lysines and phosphate groups.

**Reverse kinase order does not induce double phosphorylations.** As an alternative progression of P-TEFb function, we thought to test Cdk kinase activities in reverse order such that after incubation with P-TEFb the CTD substrate was subjected to Cdk7 to test for higher phosphorylation modes. Whereas the Cdk7/CycH/MAT1 subunits





**Figure 3 | Kinetics of P-TEFb-mediated phosphorylation of different CTD substrates.** P-TEFb exhibits a fourfold higher enzymatic activity for a CTD substrate phosphorylated at Ser7 (pS7-CTD<sub>[3]</sub>) compared with the unmodified consensus sequence (cons. CTD<sub>[3]</sub>). A CTD-containing lysine at position 7 (K7-CTD<sub>[3]</sub>) showed a similar susceptibility for P-TEFb as the consensus sequence, whereas Ser2 and Ser5 pre-phosphorylated substrates (pS2-CTD<sub>[3]</sub> and pS5-CTD<sub>[3]</sub>) were abrogated for additional phosphorylation modifications by P-TEFb. All data were reported as the mean  $\pm$  s.d. from three independent experiments.

a peptide containing Ser5 at its centre (Fig. 4e). To assign the site of phosphorylation in more detail, again a 13-mer peptide was generated but with the central Ser2 residue acetylated in order to prevent its phosphorylation. Again up to two phosphorylations were seen by analytical HPLC upon incubation with P-TEFb (Fig. 4f). The same peptide stretch containing Ser5 acetylation but not Ser2 showed instead a majority of remaining unphosphorylated peptide, whereas only a minor fraction with one additional phosphorylation was detected (Fig. 4g). These data suggest that P-TEFb preferentially phosphorylates Ser5 on a RNAPII CTD substrate, whereas a small ambiguity for Ser2 phosphorylation may remain.

In a second set of experiments, the directionality of CTD phosphorylation by P-TEFb was addressed. Using CTD substrate peptides of three hepta-repeats that contained only one Ser5 phosphorylation at either end of the template (Fig. 5a), we found that P-TEFb exhibited an almost fourfold higher catalytic activity for the substrate when the phosphorylation mark was placed at the C-terminus (Fig. 5b,c). Such differences in the amino- or carboxy-terminal extension kinetics to pre-existing phosphorylation sites could indicate a preferred directionality of the enzyme during the phosphorylation reaction.

**Two regions in Hexim1 are required for P-TEFb inhibition.** The cellular protein Hexim1 is an inhibitor of P-TEFb activity by an as yet unknown mechanism<sup>2</sup>. Using the *in vitro* kinase assay, we set out to determine the molecular requirements of P-TEFb inhibition quantitatively. Under standard conditions, 0.1  $\mu$ M P-TEFb was mixed with 100  $\mu$ M GST-CTD<sub>[8]</sub> and 100  $\mu$ M ATP and incubated for 10 min before the solution was spotted onto filter paper and the reaction stopped by immersing the filter paper in phosphoric acid. Under these conditions full-length Hexim1 at a concentration of 2.5  $\mu$ M inhibited P-TEFb activity to 28% without addition of 7SK snRNA. In contrast, 2.5  $\mu$ M of the 332 nucleotides encompassing 7SK showed no effect on P-TEFb, a result also obtained with the 66-mer 5'-hairpin finger that was previously identified to interact with Hexim1 (ref. 26; Fig. 6a). The addition of a Hexim1/7SK snRNA complex instead inhibited P-TEFb activity almost completely (3%), as did flavopiridol (2%). These results confirm a role of 7SK snRNA as a scaffold for Hexim1-mediated P-TEFb inhibition.

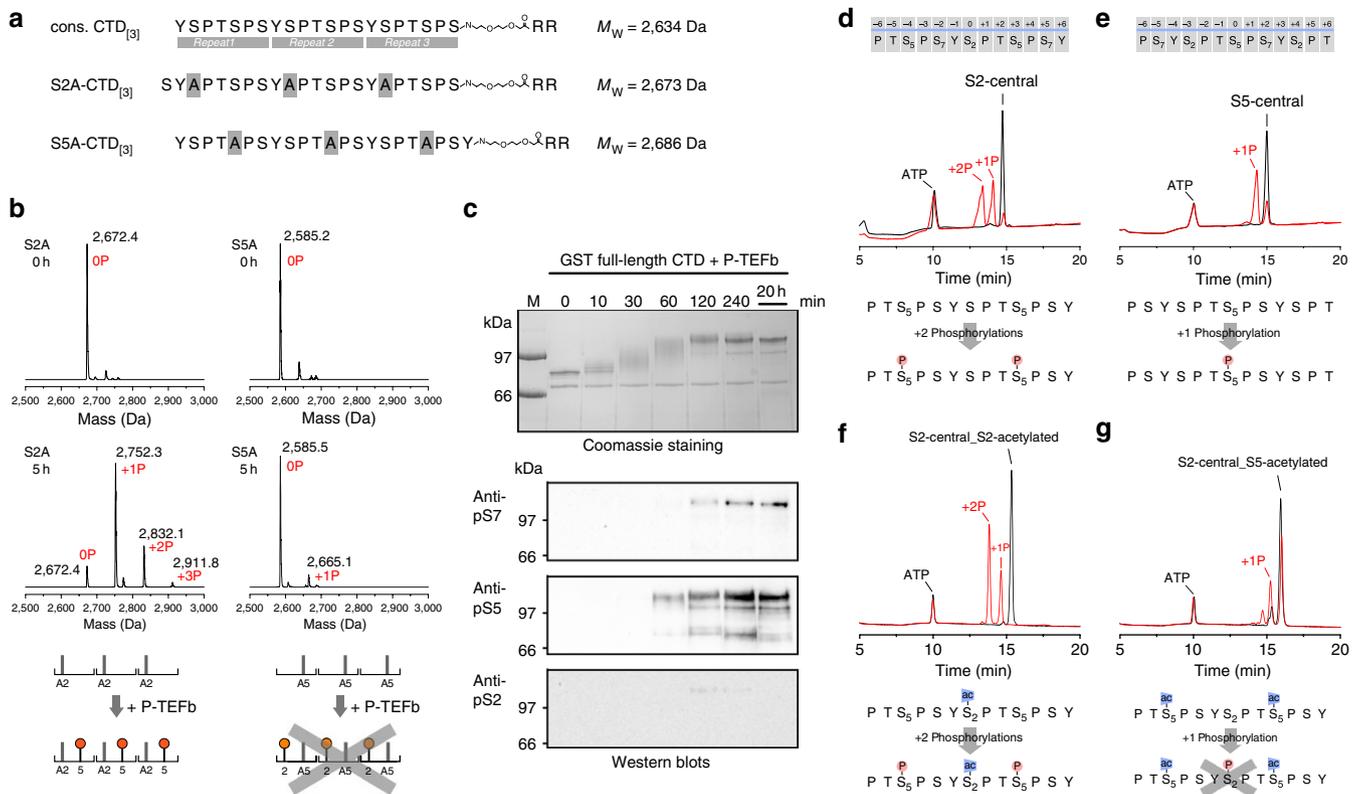
We next fragmented Hexim1 in the C-terminal Cyclin TBD (residues 255–359) and the remaining regulatory segment 1–254.

Interestingly, neither fragment inhibited P-TEFb activity (Fig. 6b), although the TBD was shown before to interact with CycT1 at a site required in the Cdk2–CycA complex for substrate recognition<sup>27</sup>. Further N-terminal elongations of the TBD revealed a marked increase in the inhibitory potential of a construct starting at position 194 compared with position 207, suggesting that residues 194–206 are necessary for the inhibitory function of Hexim1 (Fig. 6b). Point mutations identified a PYNT motif at position 202 as the inhibitory region of Hexim1, in accordance with previous studies *in vivo*<sup>28</sup>. The largest effect, however, was not contributed by Y203, as suggested in analogy to Cdk2–CycA inhibition by p27<sup>Kip</sup>, but by T205, whose mutation to glycine almost fully abrogated the inhibitory potential of Hexim1 (88% relative activity, Fig. 6c). The homogeneity of the various Hexim1 mutant proteins tested is shown in Supplementary Fig. S7.

The inhibitory effect of N-terminally truncated Hexim1 200–359 was indeed as potent as full-length Hexim1 bound to 7SK snRNA, supporting previous observations that the N-terminal part of Hexim1 could be self-inhibitory<sup>29</sup>. At concentrations of 100 nM P-TEFb, half-maximal inhibitory concentration (IC<sub>50</sub>) values of 67 nM for Hexim1 200–359 and 67 nM for Hexim1/7SK, respectively, were determined compared with an IC<sub>50</sub> of 339 nM for full-length Hexim1 (Fig. 6d). In binding experiments using isothermal titration calorimetry (ITC), the T205G mutation reduced the affinity of Hexim1 for P-TEFb about sevenfold to the portion of the Hexim1 TBD–CycT1 interaction (Table 1). These results support the synergistic contribution of the central PYNT motif in Hexim1 to the C-terminal TBD, which are required for Cdk9 kinase inhibition and CycT1 binding (Fig. 6e and Table 1).

**HIV-1 Tat/TAR does not change P-TEFb phosphorylations.** Addition of either HIV-1 Tat or the HIV-1 Tat/TAR ribonucleoprotein complex in increasing amounts to P-TEFb/Hexim1 led to recovery of the kinase activity in the substrate activity assay (Fig. 7a). This effect is supposed to occur by displacement of Hexim1 from its binding site on CycT1, as no ternary complex between these two P-TEFb regulating proteins with CycT1 is formed<sup>30,31</sup>. An indication for a change or weakening of the P-TEFb substrate specificity towards both Ser2 and Ser5 phosphorylation<sup>32,33</sup>, for example, by means of increasing numbers of phosphorylation, was not observed. Addition of Tat or Tat/TAR led to an approximately twofold higher activity of P-TEFb to a consensus CTD substrate, though the number of phosphorylations remained unchanged (Fig. 7b and Supplementary Figs S5 and S8). This result was confirmed on the full-length CTD where the viral protein–RNA complex did not induce a super-shift in the migration pattern of the phosphorylated substrate (Fig. 7c).

The substrate specificity of P-TEFb when bound to Tat remained unaltered compared with P-TEFb alone. First, western blot analysis revealed the same prevalence of P-TEFb–Tat for Ser5 phosphorylation as seen with P-TEFb alone (Fig. 7d and Supplementary Fig. S9). Likewise, using the synthesized CTD peptide library in a radioactive kinase assay, we found that the P-TEFb–Tat complex showed again a higher activity for a CTD substrate pre-phosphorylated at Ser7 compared with the consensus CTD sequence, whereas Ser2 or Ser5 pre-phosphorylations abrogated their recognition as potential substrate (Fig. 7e). After 15 min incubation, phosphorylation numbers were always increased in the presence of Tat, whereas control measurements confirmed that Tat did not act itself as a substrate. The moderate increase in P-TEFb activity could result, for example, from advanced ionic interactions of basic residues in Tat with the negatively charged substrate. In addition, the effect of Tat mutants that were described to specifically influence Tat functions<sup>34</sup> were analysed for their ability to recover P-TEFb activity from inhibition by Hexim1 (Fig. 7f). As before, addition of TAR RNA increased the effect of Tat on P-TEFb activation, but this effect was lost when the arginine-rich motif in Tat that mediates TAR binding was mutated



**Figure 4 | P-TEFb is a Ser5 CTD kinase.** (a) Design of CTD substrate peptides containing alanine mutations at either position 2 or 5 in each repeat. (b) ESI-MS analyses after 5 h incubation with P-TEFb showed that the S2A CTD mutant is still susceptible as P-TEFb substrate while no phosphorylation occurred for the S5A mutant. (c) Coomassie staining and western blot analysis of P-TEFb-mediated CTD phosphorylation. Full-length human CTD containing all 52 hepta-repeats was applied in time course experiments. Phospho-serine-directed antibodies reveal a preference of P-TEFb for Ser5 phosphorylation. (d) HPLC analysis on a 13-mer CTD peptide with Ser2 at its centre showed up to two phosphorylations by P-TEFb. (e) The same peptide length with Ser5 at its centre showed only one phosphorylation. (f) As control, the same peptide sequence as in d was synthesized with the central Ser2 residue being acetylated to prevent its phosphorylation. Again, two additional phosphorylations were achieved after P-TEFb incubation. (g) The same peptide sequence as in d was synthesized with the two distant Ser5 residues being acetylated. Only a minor fraction was detected that contained one additional phosphorylation.

to alanines. Mutation of cysteine residues in Tat that form two distinct zinc-fingers, one including an intermolecular zinc-finger with CycT1, diminished the stimulatory function of Tat. Likewise, an N-terminal truncation that largely omits the Tat transactivation domain almost fully abrogated Tat function. In the latter two experiments, addition of TAR even weakened the Tat effect, potentially due to the capturing of Tat from P-TEFb.

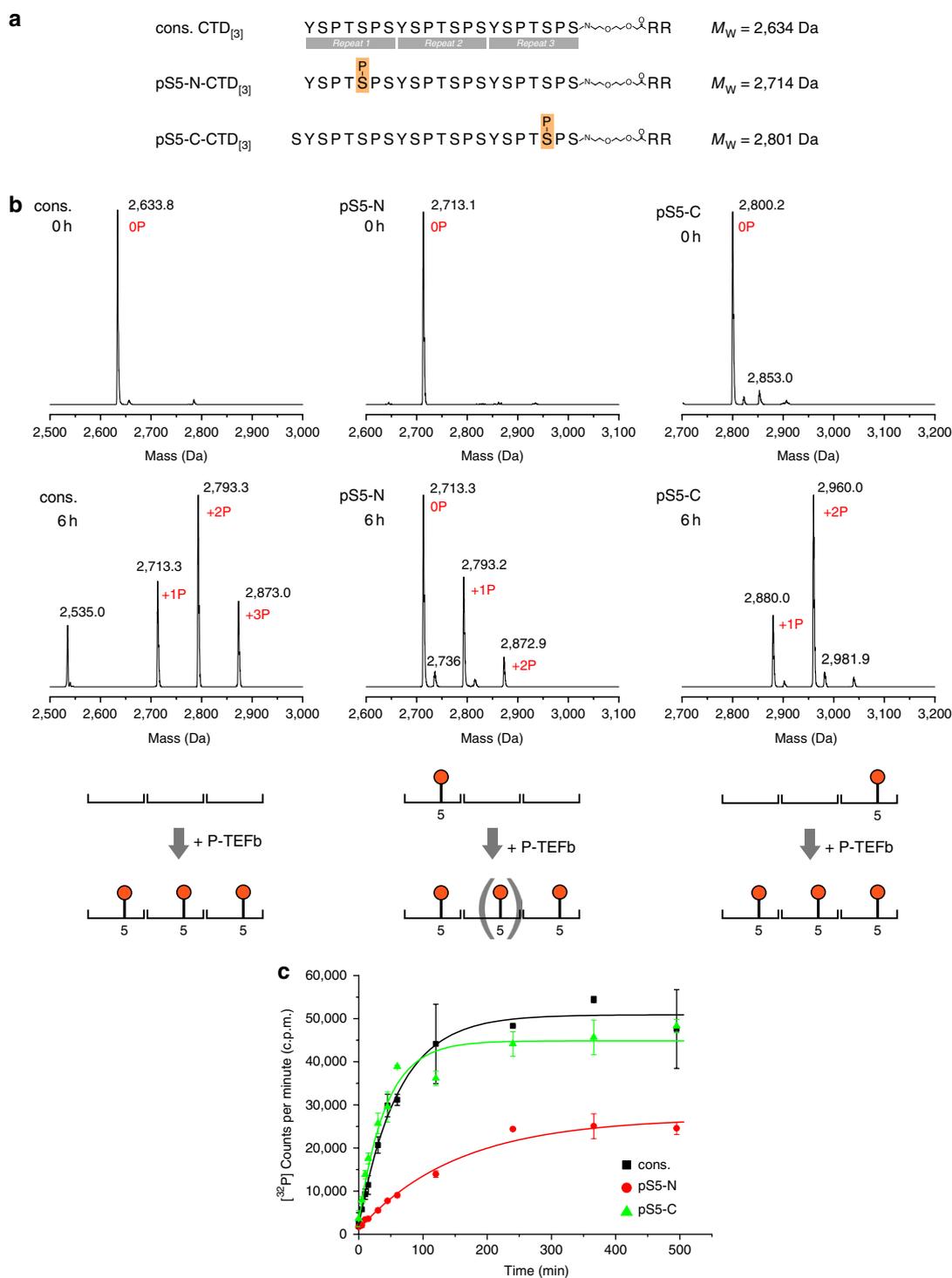
Finally, we investigated the effect of human AFF4 on the activity of P-TEFb. AFF4 is a subunit of the super elongation complex that is recruited by mixed lineage leukaemia (MLL) proteins to activate the expression of MLL target genes<sup>35–38</sup>. The N-terminal 300 amino acids of AFF4 were shown to specifically interact with P-TEFb<sup>39,40</sup>. Again, the substrate specificity was tested for the consensus CTD<sub>[3]</sub> as well as serine pre-phosphorylated peptides and K7-CTD<sub>[3]</sub>. Overall, the phosphorylation profile appeared similar as for HIV-1 Tat with a considerable stimulatory effect of P-TEFb–AFF4 for the consensus CTD (Supplementary Fig. S10). These results suggest similar molecular mechanisms for the stimulation of gene expression for the leukaemia-associated MLL proteins as for the viral Tat–TAR transactivation complex.

## Discussion

Since the identification of Cdk12/Cdk13 as the human orthologues of yeast Ctk1, there has been considerable debate about the substrate specificity of RNAPII transcription regulating kinases<sup>41,42</sup>.

Here we show by analytical means and western blot analysis that P-TEFb (Cdk9) specifically phosphorylates Ser5 of the hepta-repeat containing CTD. Although pre-phosphorylations of Ser7 even stimulate the catalytic activity of P-TEFb for CTD phosphorylation in the same hepta-repeat, Ser5 and Ser2 phosphorylations abrogate such double modifications. Instead, our data suggest that Ser5 and Ser2 phosphorylation of the CTD are two distinct events that are either spatially or temporally separated.

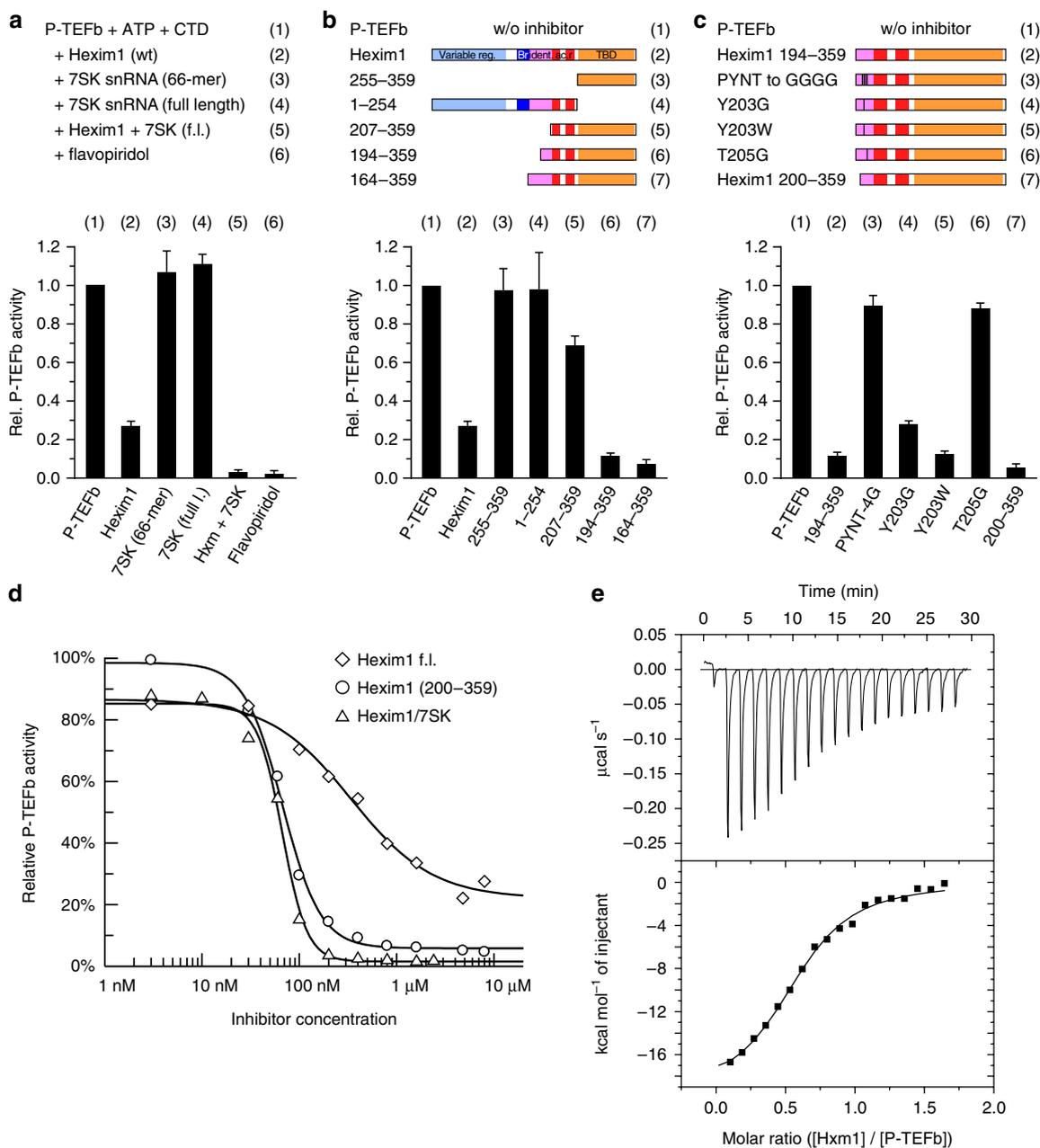
Cyclin-dependent kinases Cdk1 and Cdk2 typically phosphorylate substrates of the consensus type (S/T)Px(K/R)<sup>22,23</sup>. In this motif, the amino group of the lysine residue at position +3 is involved in ionic interactions with the phospho-threonine moiety within the T-loop activation segment of the kinase<sup>22</sup>. Following this recognition principle, phosphorylation of Ser5 by P-TEFb assigns Tyr1 a central role in the S<sub>5</sub>PSY<sub>1</sub> sequence motif. The aromatic tyrosine is indeed the most peculiar moiety in the hepta-repeats of otherwise low sequence complexity containing another three serines, two prolines and one threonine each. The tyrosine may thus account for the kinase–substrate specificity seen here for P-TEFb that secures exactly one phosphorylation per CTD repeat. The position +2 in this connotation, which can be serine, phospho-serine or lysine, instead seems less decisive but a negative charge appears to be beneficial. The stimulatory effect of Ser7 pre-phosphorylation for priming of the CTD as P-TEFb substrate gets even more accentuated as all other phospho-modifications ascribed to the CTD (pTyr1, pSer2,



**Figure 5 | Kinetics of P-TEFb activity on partially pre-phosphorylated CTD substrates.** (a) Design of CTD substrate peptides containing a Ser5 phosphorylation mark in either the N- or C-terminal hepta-repeat. The peptides were used for mass spectrometry and kinetic analyses. (b) ESI-MS analyses after 6 h incubation time with P-TEFb revealed a majority of two additional phosphorylations for the CTD substrate when the pSer5 mark was set in the C-terminal repeat (pS5-C-CTD<sub>[3]</sub>), whereas only one additional phosphorylation prevailed after the same incubation time when the Ser5 phosphorylation mark was set in the N-terminal repeat (pS5-N-CTD<sub>[3]</sub>). (c) Kinetic analysis showed a fourfold higher catalytic activity of P-TEFb for a substrate containing a C-terminal Ser5 phosphorylation mark compared with the reversed order. At a concentration of 0.2  $\mu$ M P-TEFb, a  $K_{cat}/K_M$  value of  $4.1 \times 10^3 \text{ M}^{-1} \text{ s}^{-1}$  was determined for pS5-C-CTD<sub>[3]</sub> compared with  $1.1 \times 10^3 \text{ M}^{-1} \text{ s}^{-1}$  for pS5-N-CTD<sub>[3]</sub> and  $2.8 \times 10^3 \text{ M}^{-1} \text{ s}^{-1}$  for the consensus CTD substrate. Data are the mean of two independent experiments; bars indicate the range.

pThr4 and pSer5) abrogate its subsequent phosphorylation by P-TEFb. Phosphorylation of Ser2 in the context of a pSer5 mark making an S<sub>2</sub>PTpS<sub>5</sub> recognition motif could indeed be challenging for any cyclin-dependent kinase as the presence of a double nega-

tively charged phospho-serine at a position where typically a basic lysine or arginine resides may lead to electrostatic repulsion. In fact, it remains to be shown that such pSer2/pSer5 double phosphorylation marks exist in a single CTD hepta-repeat.

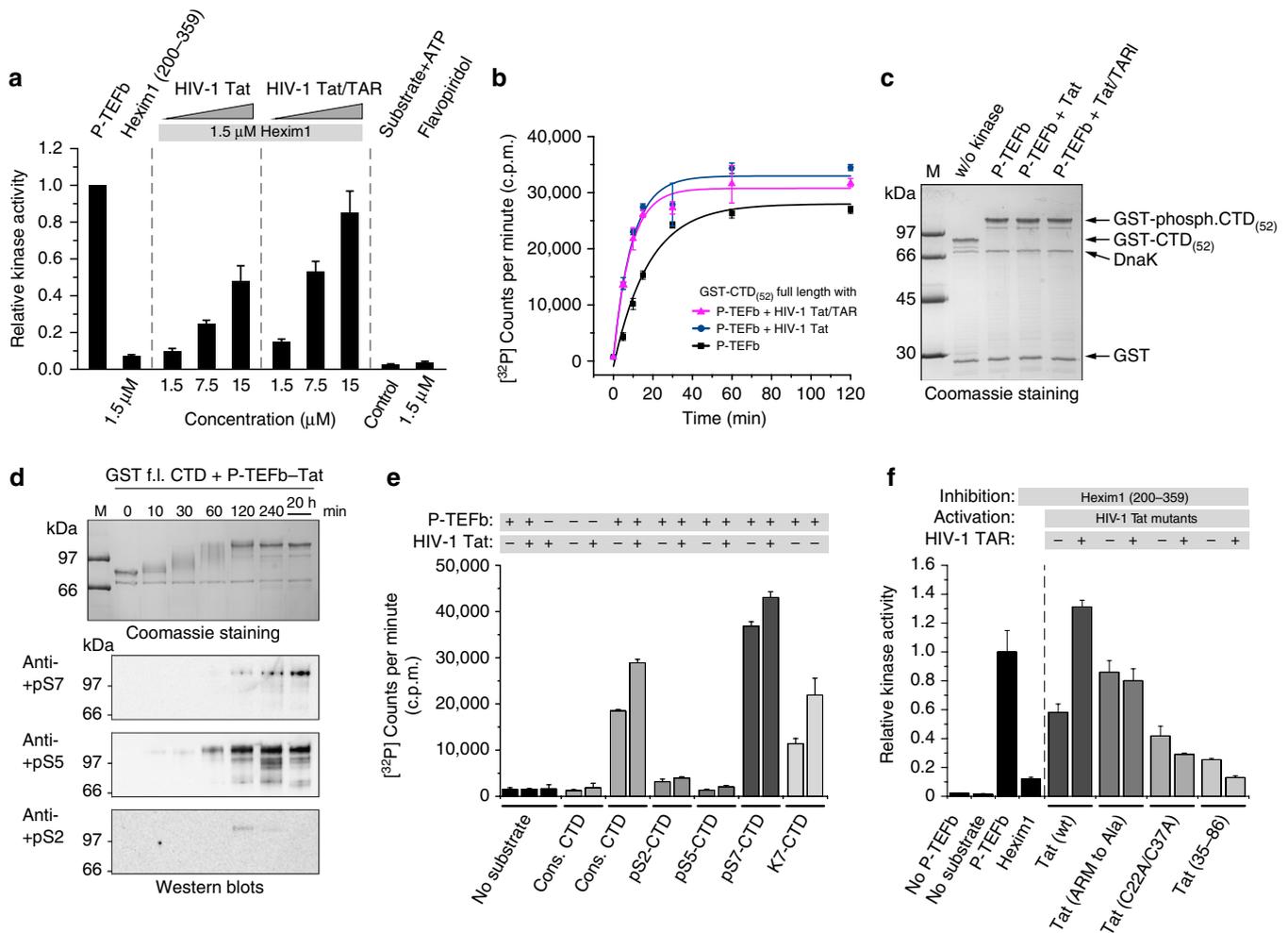


**Figure 6 | Inhibition of P-TEFb requires two distinct motifs in different domains of Hexim1.** (a) Effect of Hexim1 and 7SK snRNA on the kinase activity of P-TEFb. The inhibitory effect of full-length Hexim1 was significantly increased by 7SK, although 7SK showed no influence on its own. (b) Mapping of the functional relevant sites in Hexim1 for P-TEFb inhibition. Although the Cyclin T-binding domain (TBD) is required for P-TEFb recognition, residues between 194 and 206 are required for Cdk9 inhibition. (c) Mutation of T205G within a highly conserved PYNT motif in Hexim1 194–359 abrogated its inhibitory effect on P-TEFb activity. Mutation of Y203G or Y203W instead potentially retained the inhibitory function of Hexim1. Data in a–c are the mean  $\pm$  s.d. from three independent experiments. (d) A concentration series of Hexim1 inhibitor constructs, at 100 nM P-TEFb concentration revealed  $IC_{50}$  values of 339 nM for full-length Hexim1, 67 nM for the complex of full-length Hexim1 and 7SK snRNA, and 67 nM for N-terminally truncated Hexim1 variant 200–359. (e) Isothermal titration calorimetry measurement of P-TEFb with Hexim1 194–359 indicated a dissociation constant of 0.3  $\mu$ M.

The combinations of various phosphorylations in either one or two neighbouring hepta-repeats create the possibility for an additional layer of recognition specificity to mediate transcriptional regulation. Most simply, regions of Ser5 and Ser2 phosphorylation are spatially separated on the CTD, creating thus functional islands. Alternatively, stringent Ser5 dephosphorylation by phosphatases such as Rtr1 or Fcp1 in yeast<sup>43,44</sup> is required in order to render Ser2 phosphorylation possible. A recent study has indeed identified that the human RPA2 phosphatase specifically recognizes Ser7 phosphorylation

marks on the CTD to facilitate Ser5 dephosphorylation<sup>45</sup>. In a similar line of argument, a universal RNAPII CTD cycle along genes has been proposed, which is orchestrated by complex interplays between kinases, phosphatases and proline isomerases<sup>46</sup>. The determination of the minimal distances required between the two general phosphorylation marks Ser5 and Ser2 and their possible combinations with other modifications such as pSer7 might therefore be a key step in the understanding of CTD interacting proteins that recognize different phospho-isoforms during the transcription cycle<sup>47–50</sup>.





**Figure 7 | HIV-1 Tat/TAR overcomes P-TEFb inhibition of Hexim1 but does not change its substrate specificity.** (a) The kinase activity of 0.1 μM P-TEFb using GST-CTD<sub>[8]</sub> substrate and radioactive-labelled ATP co-substrate was measured after 10 min incubation (left column). Inhibition of P-TEFb activity by Hexim1 200–359 led to a reduction of kinase activity to 6% (second column). Addition of either HIV-1 Tat or the preformed HIV-1 Tat/TAR complex in increasing concentrations recovered P-TEFb activity (central columns). As controls, substrate plus ATP without P-TEFb or with P-TEFb with flavopiridol were measured (right columns). (b) Kinetics of Tat/TAR-stimulated P-TEFb activity in the absence of Hexim1. Full-length GST-CTD<sub>[52]</sub> at 2.5 μM concentration was incubated with 2 mM ATP and 0.2 μM P-TEFb. HIV-1 Tat 1–86 or Tat/TAR was added before the reaction at concentrations of 2 μM. The specific activities  $K_{cat}/K_M$  were determined to  $4.7 \times 10^3 \text{ M}^{-1} \text{ s}^{-1}$  for P-TEFb,  $8.8 \times 10^3 \text{ M}^{-1} \text{ s}^{-1}$  for P-TEFb-Tat, and  $9.6 \times 10^3 \text{ M}^{-1} \text{ s}^{-1}$  for P-TEFb with Tat/TAR. (c) SDS-PAGE analysis of P-TEFb phosphorylated full-length CTD in the presence or absence of Tat or Tat/TAR showed no increase in the migration pattern. (d) Western blot analysis of P-TEFb-Tat-mediated CTD phosphorylation revealed prevalence for Ser5 phosphorylation. (e) The substrate specificity of P-TEFb on synthesized CTD peptides remained unchanged in the presence of HIV-1 Tat. The phosphorylation efficacy was monitored after 15 min. (f) Effects of Tat mutations on the recovery of P-TEFb activity in the presence of Hexim1. To a concentration of 0.1 μM P-TEFb, 100 μM GST-CTD and 1 mM ATP, a concentration of 1.5 μM Hexim1 was added for inhibition and 15 μM Tat or Tat/TAR for activation. Disruption of the intermolecular zinc-finger formation by mutation C22A/C37A or N-terminal truncation 35–86 significantly reduced the ability of Tat to activate P-TEFb. Data shown in a, b, e and f are the mean ± s.d. from three independent experiments.

with a calculated mass of 3,124.0 Da; a consensus CTD continuously phosphorylated at Ser5 YSPtSPSYSPtSPSYSPtSPS-[PEG]<sub>2</sub>-RR (pS5-CTD<sub>[3]</sub>, calculated mass 2,873.8 Da); a consensus CTD continuously phosphorylated at Ser7 YSPtSPpSYSPtSPpSYSPtSPpSY-[PEG]<sub>2</sub>-RR (pS7-CTD<sub>[3]</sub>, calculated mass 3,037.0 Da); a consensus CTD phosphorylated at the N-terminal Ser5 position YSPtSPSYSPtSPSYSPtSPS-[PEG]<sub>2</sub>-RR (pS5-N-CTD<sub>[3]</sub>, calculated mass 2,713.9 Da); a consensus CTD phosphorylated at the C-terminal Ser5 position SYSPTSPSYSPtSPSYSPtSPS-[PEG]<sub>2</sub>-RR (pS5-C-CTD<sub>[3]</sub>, calculated mass 2,800.9 Da); and a CTD containing lysines at position 7 as found in distal CTD repeats YSPtSPKYSPTSPKYSPTSPK-[PEG]<sub>2</sub>-RR (K7-CTD<sub>[3]</sub>, calculated mass 2,757.1 Da). Peptides were dissolved in 50 mM HEPES, pH 7.5, to a stock concentration of 5 mM. CTD peptides were marked at the C-terminus with a double arginine motif separated by a polyethylene glycol spacer for better ionization properties in the ESI-MS experiments and for increased transfer rates in the radioactive filter-binding assay.

**Kinase assays using recombinant P-TEFb.** Radioactive kinase reactions (typically 35 μl) were carried out with recombinant highly purified proteins, using a

standard protocol similarly as described<sup>60</sup>. P-TEFb (0.1 μM) was pre-incubated with GST-CTD (100 μM) of multiple repeats and Hexim proteins for 10 min at room temperature in kinase buffer (50 mM HEPES, pH 7.6, 34 mM KCl, 7 mM MgCl<sub>2</sub>, 2.5 mM dithiothreitol, 5 mM β-glycerol phosphate, 0.5 mM Na<sub>3</sub>VO<sub>4</sub>). Cold ATP (100 μM) and 0.5 μCi [<sup>32</sup>P]-γ-ATP were added and the reaction mixture incubated for 10 min at 30 °C at 300 r.p.m. For each reaction, two aliquots, 15 μl each, were spotted onto P81 Whatman paper squares and the reaction stopped by immediately immersing the paper in 0.75% (v/v) phosphoric acid. Paper squares were washed three times for 5 min with 0.75% (v/v) phosphoric acid, with at least 5 ml washing solution per paper. Radioactivity was counted in a Beckmann Scintillation Counter (Beckman Coulter) for 1 min. Typically, three experiment series were performed at different days with at least two independent measurements.

**Mass spectrometry analyses.** Peptide and protein masses were determined by liquid chromatography-electrospray ionization-mass spectrometry using an Agilent 1100 chromatography system and an LCQ Advantage MAX (Finnigan) mass spectrometer operating in positive ion mode. Proteins were applied onto an

Vydac RP-C4 column (Grace) at 20% buffer B (CH<sub>3</sub>CN with 0.08% trifluoroacetic acid) in buffer A (H<sub>2</sub>O plus 0.1% TFA) and eluted with a gradient from 20–80% buffer B at a flow rate of 1 ml min<sup>-1</sup>. Peptide samples were loaded onto the column at 5% buffer B and eluted with a gradient from 5–80% buffer B. Data evaluation was performed with the Xcalibur, MagTran and Bioworks software packages.

**Isothermal titration calorimetry.** Thermodynamic parameters of the Hexim1–P-TEFb and the Hexim1–CycT1 interactions were determined by ITC using either an iTC<sub>200</sub> or VP-ITC calorimeter (MicroCal). P-TEFb and CycT1 proteins were thermostated in the measurement cell and titrated with approximately tenfold higher concentrated Hexim1 proteins. Measurements with P-TEFb were carried out in two different buffers with buffer A containing 50 mM HEPES, pH 7.6, 500 mM NaCl, 10% (v/v) glycerol, 5 mM β-mercaptoethanol (high salt buffer) and buffer B containing 50 mM HEPES, pH 7.6, 250 mM NaCl, 5% (v/v) glycerol, 5 mM β-mercaptoethanol (low salt buffer). The change in heating power was observed until equilibrium was reached before the next injection was started. The data were evaluated using the manufacturer's analysis software.

**Western blot analysis.** For western blot analysis, 10 μM of human full-length GST–CTD was incubated at 30 °C with 0.04 μM P-TEFb in presence or absence of 0.4 μM Tat. Reactions were started by addition of 3 mM ATP and stopped at indicated time points by adding SDS-loading buffer to the reaction. For Coomassie staining, 1 μg of GST–CTD from the reaction mix was loaded per lane on 12% SDS–PAGE gel. For western blot analysis, 0.1 μg of GST–CTD from the reaction mix was loaded per lane on 12% SDS–PAGE gel, transferred onto nitrocellulose membrane (Whatman) and processed further by standard western-blotting protocols. For these assays, a 1:50 (α-pSer2), 1:1,000 (α-pSer5) or 1:200 (α-pSer7) dilution of primary rat IgG and 1:10,000 dilution of secondary HRP anti-rat IgG (Santa Cruz, sc-2964) antibodies were used. The anti-phospho-CTD antibodies were a kind gift from Dirk Eick, Munich.

**Quantitative determination of P-TEFb activity by HPLC analysis.** For determination of P-TEFb-mediated substrate modifications, 13-mer CTD substrate peptides were designed with either Ser2 or Ser5 residues at its centre. CTD substrate peptides were incubated at 0.25 mM concentration for 5 h at 30 °C with 0.2 μM P-TEFb in presence of 3 mM ATP. Reactions were analysed by HPLC on a ProntoSIL (Bischoff Chromatography) 120–5–C18 column (250×4.6 mm<sup>2</sup>) using a Beckman Coulter System Gold consisting of a 126 Solvent Module and a 168 Detector set at a wavelength of 214 nm. Peptide samples were applied at a flow rate of 1 ml min<sup>-1</sup> at 5% buffer B (CH<sub>3</sub>CN with 0.08% TFA) in buffer A (H<sub>2</sub>O plus 0.1% TFA) and eluted in a 15 min linear gradient from 5–45% buffer B. Peak fractions were collected and masses determined by liquid chromatography-electrospray ionization-mass spectrometry using an Agilent 1100 chromatography system and an LCQ Advantage MAX (Finnigan) mass spectrometer operating in positive and negative ion mode.

## References

- Fuda, N. J., Ardehali, M. B. & Lis, J. T. Defining mechanisms that regulate RNA polymerase II transcription *in vivo*. *Nature* **461**, 186–192 (2009).
- Peterlin, B. M. & Price, D. H. Controlling the elongation phase of transcription with P-TEFb. *Mol. Cell* **23**, 297–305 (2006).
- Buratowski, S. Progression through the RNA polymerase II CTD cycle. *Mol. Cell* **36**, 541–546 (2009).
- Egloff, S. & Murphy, S. Cracking the RNA polymerase II CTD code. *Trends Genet.* **24**, 280–288 (2008).
- Chapman, R. D., Heidemann, M., Hintermair, C. & Eick, D. Molecular evolution of the RNA polymerase II CTD. *Trends Genet.* **24**, 289–296 (2008).
- Phatnani, H. P. & Greenleaf, A. L. Phosphorylation and functions of the RNA polymerase II CTD. *Genes Dev.* **20**, 2922–2936 (2006).
- Meinhart, A., Kamenski, T., Hoepfner, S., Baumli, S. & Cramer, P. A structural perspective of CTD function. *Genes Dev.* **19**, 1401–1415 (2005).
- Saunders, A., Core, L. J. & Lis, J. T. Breaking barriers to transcription elongation. *Nat. Rev. Mol. Cell Biol.* **7**, 557–567 (2006).
- Egloff, S. *et al.* Serine-7 of the RNA polymerase II CTD is specifically required for snRNA gene expression. *Science* **318**, 1777–1779 (2007).
- Chapman, R. D. *et al.* Transcribing RNA polymerase II is phosphorylated at CTD residue serine-7. *Science* **318**, 1780–1782 (2007).
- Kim, M., Suh, H., Cho, E. J. & Buratowski, S. Phosphorylation of the yeast Rpb1 C-terminal domain at serines 2, 5, and 7. *J. Biol. Chem.* **284**, 26421–26426 (2009).
- Boeing, S., Rigault, C., Heidemann, M., Eick, D. & Meisterernst, M. RNA polymerase II C-terminal heptarepeat domain Ser-7 phosphorylation is established in a mediator-dependent fashion. *J. Biol. Chem.* **285**, 188–196 (2010).
- Egloff, S. *et al.* The integrator complex recognizes a new double mark on the RNA polymerase II carboxyl-terminal domain. *J. Biol. Chem.* **285**, 20564–20569 (2010).
- Hargreaves, D. C., Horng, T. & Medzhitov, R. Control of inducible gene expression by signal-dependent transcriptional elongation. *Cell* **138**, 129–145 (2009).
- Nechaev, S. *et al.* Global analysis of short RNAs reveals widespread promoter-proximal stalling and arrest of Pol II in *Drosophila*. *Science* **327**, 335–338 (2010).
- Rahl, P. B. *et al.* c-Myc regulates transcriptional pause release. *Cell* **141**, 432–445 (2010).
- Bai, X. *et al.* TIF1γ controls erythroid cell fate by regulating transcription elongation. *Cell* **142**, 133–143 (2010).
- Tietjen, J. R. *et al.* Chemical-genomic dissection of the CTD code. *Nat. Struct. Mol. Biol.* **17**, 1154–1161 (2010).
- Kim, H. *et al.* Gene-specific RNA polymerase II phosphorylation and the CTD code. *Nat. Struct. Mol. Biol.* **17**, 1279–1286 (2010).
- Mayer, A. *et al.* Uniform transitions of the general RNA polymerase II transcription complex. *Nat. Struct. Mol. Biol.* **17**, 1272–1278 (2010).
- Sims, R. J. III *et al.* The C-terminal domain of RNA polymerase II is modified by site-specific methylation. *Science* **332**, 99–103 (2011).
- Brown, N. R., Noble, M. E., Endicott, J. A. & Johnson, L. N. The structural basis for specificity of substrate and recruitment peptides for cyclin-dependent kinases. *Nat. Cell Biol.* **1**, 438–443 (1999).
- Holt, L. J. *et al.* Global analysis of Cdk1 substrate phosphorylation sites provides insights into evolution. *Science* **325**, 1682–1686 (2009).
- Baskaran, R., Dahmus, M. E. & Wang, J. Y. Tyrosine phosphorylation of mammalian RNA polymerase II carboxyl-terminal domain. *Proc. Natl Acad. Sci. USA* **90**, 11167–11171 (1993).
- Hsin, J. P., Sheth, A. & Manley, J. L. RNAP II CTD phosphorylated on threonine-4 is required for histone mRNA 3' end processing. *Science* **334**, 683–686 (2011).
- Egloff, S., van Herreweghe, E. & Kiss, T. Regulation of polymerase II transcription by 7SK snRNA: two distinct RNA elements direct P-TEFb and Hexim1 binding. *Mol. Cell Biol.* **26**, 630–642 (2006).
- Dames, S. A. *et al.* Structure of the Cyclin T binding domain of Hexim1 and molecular basis for its recognition of P-TEFb. *Proc. Natl Acad. Sci. USA* **104**, 14312–14317 (2007).
- Michels, A. A. *et al.* Binding of the 7SK snRNA turns the HEXIM1 protein into a P-TEFb (CDK9/cyclin T) inhibitor. *EMBO J.* **23**, 2608–2619 (2004).
- Barboric, M. *et al.* Interplay between 7SK snRNA and oppositely charged regions in HEXIM1 direct the inhibition of P-TEFb. *EMBO J.* **24**, 4291–4303 (2005).
- Schulte, A. *et al.* Identification of a Cyclin T-binding domain in Hexim1 and biochemical analysis of its binding competition with HIV-1 Tat. *J. Biol. Chem.* **280**, 24968–24977 (2005).
- Barboric, M. *et al.* Tat competes with HEXIM1 to increase the active pool of P-TEFb for HIV-1 transcription. *Nucleic Acids Res.* **35**, 2003–2012 (2007).
- Zhou, M. *et al.* Tat modifies the activity of CDK9 to phosphorylate serine 5 of the RNA polymerase II carboxyl-terminal domain during human immunodeficiency virus type 1 transcription. *Mol. Cell Biol.* **20**, 5077–5086 (2000).
- Tahirov, T. H. *et al.* Crystal structure of HIV-1 Tat complexed with human P-TEFb. *Nature* **465**, 747–751 (2010).
- Ott, M., Geyer, M. & Zhou, Q. The control of HIV transcription: keeping RNA polymerase II on track. *Cell Host Microbe* **10**, 426–435 (2011).
- He, N. *et al.* HIV-1 Tat and host AFF4 recruit two transcription elongation factors into a bifunctional complex for coordinated activation of HIV-1 transcription. *Mol. Cell* **38**, 428–438 (2010).
- Sobhian, B. *et al.* HIV-1 Tat assembles a multifunctional transcription elongation complex and stably associates with the 7SK snRNP. *Mol. Cell* **38**, 439–451 (2010).
- Lin, C. *et al.* AFF4, a component of the ELL/P-TEFb elongation complex and a shared subunit of MLL chimeras, can link transcription elongation to leukemia. *Mol. Cell* **37**, 429–437 (2010).
- Byun, J. S. *et al.* ELL facilitates RNA polymerase II pause site entry and release. *Nat. Commun.* **3**, 633 (2012).
- He, N. *et al.* Human polymerase-associated factor complex (PAF<sub>c</sub>) connects the super elongation complex (SEC) to RNA polymerase II on chromatin. *Proc. Natl Acad. Sci. USA* **108**, 636–645 (2011).
- Biswas, D. *et al.* Function of leukemogenic mixed lineage leukemia 1 (MLL) fusion proteins through distinct partner protein complexes. *Proc. Natl Acad. Sci. USA* **108**, 15751–15756 (2011).
- Bartkowiak, B. *et al.* CDK12 is a transcription elongation-associated CTD kinase, the metazoan ortholog of yeast Ctk1. *Genes Dev.* **24**, 2303–2316 (2010).
- Blazek, D. *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).
- Mosley, A. L. *et al.* Rtr1 is a CTD phosphatase that regulates RNA polymerase II during the transition from serine 5 to serine 2 phosphorylation. *Mol. Cell* **34**, 168–178 (2009).

44. Cho, E. J., Kobor, M. S., Kim, M., Greenblatt, J. & Buratowski, S. 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).
45. Egloff, S., Zaborowska, J., Laitem, C., Kiss, T. & Murphy, S. Ser7 phosphorylation of the CTD recruits the RPAP2 Ser5 phosphatase to snRNA genes. *Mol. Cell* **45**, 111–122 (2012).
46. Bataille, A. R. *et al.* A universal RNA polymerase II CTD cycle is orchestrated by complex interplays between kinase, phosphatase, and isomerase enzymes along genes. *Mol. Cell* **45**, 158–170 (2012).
47. Li, M. *et al.* Solution structure of the Set2-Rpb1 interacting domain of human Set2 and its interaction with the hyperphosphorylated C-terminal domain of Rpb1. *Proc. Natl Acad. Sci. USA* **102**, 17636–17641 (2005).
48. Vojnic, E., Simon, B., Strahl, B. D., Sattler, M. & Cramer, P. Structure and carboxyl-terminal domain (CTD) binding of the Set2 SRI domain that couples histone H3 Lys36 methylation to transcription. *J. Biol. Chem.* **281**, 13–15 (2006).
49. Lunde, B. M. *et al.* Cooperative interaction of transcription termination factors with the RNA polymerase II C-terminal domain. *Nat. Struct. Mol. Biol.* **17**, 1195–1201 (2010).
50. Becker, R., Loll, B. & Meinhart, A. Snapshots of the RNA processing factor SCAF8 bound to different phosphorylated forms of the carboxyl-terminal domain of RNA polymerase II. *J. Biol. Chem.* **283**, 22659–22669 (2008).
51. Chapman, R. D., Palancade, B., Lang, A., Bensaude, O. & Eick, D. The last CTD repeat of the mammalian RNA polymerase II large subunit is important for its stability. *Nucleic Acids Res.* **32**, 35–44 (2004).
52. Loog, M. & Morgan, D. O. Cyclin specificity in the phosphorylation of cyclin-dependent kinase substrates. *Nature* **434**, 104–108 (2005).
53. Anand, K., Schulte, A., Vogel-Bachmayr, K., Scheffzek, K. & Geyer, M. Structural insights into the cyclin T1-Tat-TAR RNA transcription activation complex from EIAV. *Nat. Struct. Mol. Biol.* **15**, 1287–1292 (2008).
54. Perales, R. & Bentley, D. ‘Cotranscriptionality’: the transcription elongation complex as a nexus for nuclear transactions. *Mol. Cell* **36**, 178–191 (2009).
55. Viladevall, L. *et al.* TFIIF and P-TEFb coordinate transcription with capping enzyme recruitment at specific genes in fission yeast. *Mol. Cell* **33**, 738–751 (2009).
56. Qiu, H., Hu, C. & Hinnebusch, A. G. Phosphorylation of the Pol II CTD by KIN28 enhances BUR1/BUR2 recruitment and Ser2 CTD phosphorylation near promoters. *Mol. Cell* **33**, 752–762 (2009).
57. Bigalke, J. M. *et al.* Formation of Tat-TAR containing ribonucleoprotein complexes for biochemical and structural analyses. *Methods* **53**, 78–84 (2011).
58. Czudnochowski, N., Vollmuth, F., Baumann, S., Vogel-Bachmayr, K. & Geyer, M. Specificity of Hexim1 and Hexim2 complex formation with cyclin T1/T2, importin alpha and 7SK snRNA. *J. Mol. Biol.* **395**, 28–41 (2010).
59. Schönichen, A. *et al.* A flexible bipartite coiled coil structure is required for the interaction of Hexim1 with the P-TEFb subunit cyclin T1. *Biochemistry* **49**, 3083–3091 (2010).
60. Hastie, C. J., McLauchlan, H. J. & Cohen, P. Assay of protein kinases using radiolabeled ATP: a protocol. *Nat. Protoc.* **1**, 968–971 (2006).

## Acknowledgements

We thank Karin Vogel-Bachmayr, Diana Ludwig and Sascha Gentz for excellent technical assistance and Roger S. Goody, Ali Kinkhabwala and Robert P. Fisher for stimulating discussions. We thank Dirk Eick for a kind gift of the anti-phospho-CTD antibodies and Friederike Itzen for initial work on AFF4. Flavopiridol was a kind gift of the Aids Reagent Program. N.C. was supported by a grant from the Fond der Chemischen Industrie (FCI). C.A.B. is a fellow of the IMPRS in Chemical Biology, Dortmund. This work was supported by a grant from the Deutsche Forschungsgemeinschaft to M.G. (GE-976/5).

## Author contributions

N.C. and C.A.B. prepared proteins, performed P-TEFb activity experiments and contributed with kinetic measurements. M.G. designed the study and wrote the paper. All authors discussed the results and commented on the manuscript.

## Additional information

**Supplementary Information** accompanies this paper at <http://www.nature.com/naturecommunications>

**Competing financial interests:** The authors declare no competing financial interests.

**Reprints and permission** information is available online at <http://npg.nature.com/reprintsandpermissions/>

**How to cite this article:** Czudnochowski, N. *et al.* Serine-7 but not serine-5 phosphorylation primes RNA polymerase II CTD for P-TEFb recognition. *Nat. Commun.* **3**:842 doi: 10.1038/ncomms1846 (2012).

**License:** This work is licensed under a Creative Commons Attribution-NonCommercial-Share Alike 3.0 Unported License. To view a copy of this license, visit <http://creativecommons.org/licenses/by-nc-sa/3.0/>