LETTER

https://doi.org/10.1038/s41586-018-0174-3

Phase-separation mechanism for C-terminal hyperphosphorylation of RNA polymerase II

Huasong Lu[1,2], Dan Yu[2], Anders S. Hansen[2], Sourav Ganguly[2], Rongdiao Liu[1], Alec Heckert[2], Xavier Darzacq[2] & Qiang Zhou[2]*

Hyperphosphorylation of the C-terminal domain (CTD) of the RPB1 subunit of human RNA polymerase (Pol) II is essential for transcriptional elongation and mRNA processing[1–3]. The CTD contains 52 heptapeptide repeats of the consensus sequence YSPTSPS. The highly repetitive nature and abundant possible phosphorylation sites of the CTD exert special constraints on the kinases that catalyse its hyperphosphorylation. Positive transcription elongation factor b (P-TEFb)—which consists of CDK9 and cyclin T1—is known to hyperphosphorylate the CTD and negative elongation factors to stimulate Pol II elongation[1,4,5]. The sequence determinant on P-TEFb that facilitates this action is currently unknown. Here we identify a histidine-rich domain in cyclin T1 that promotes the hyperphosphorylation of the CTD and stimulation of transcription by CDK9. The histidine-rich domain markedly enhances the binding of P-TEFb to the CTD and functional engagement with target genes in cells. In addition to cyclin T1, at least one other kinase—DYRK1A[6]—also uses a histidine-rich domain to target and hyperphosphorylate the CTD. As a low-complexity domain, the histidine-rich domain also promotes the formation of phase-separated liquid droplets in vitro, and the localization of P-TEFb to nuclear speckles that display dynamic liquid properties and are sensitive to the disruption of weak hydrophobic interactions. The CTD—which in isolation does not phase separate, despite being a low-complexity domain—is trapped within the cyclin T1 droplets, and this process is enhanced upon pre-phosphorylation by CDK7 of transcription initiation factor TFIIH[1–3]. By using multivalent interactions to create a phase-separated functional compartment, the histidine-rich domain in kinases targets the CTD into this environment to ensure hyperphosphorylation and efficient elongation of Pol II.

Among all transcription-related cyclins, cyclin (CYC) T—which includes T1 (CYCT1) and T2—has the longest C-terminal regions. In CYCT1, the N-terminal region is structured and contains the cyclin-box repeats required for binding and activating CDK9, whereas the C-terminal region has only a few isolated motifs and is mostly unstructured (Fig. 1a).

To determine whether the CYCT1 C-terminal region is important for regulating CDK9 activity, we performed in vitro kinase reactions to examine affinity-purified CDK9–CYCT1–Flag heterodimers that contained various truncated forms of CYCT1 with progressively shortened C termini to phosphorylate the CTD (Fig. 1a and Extended Data Fig. 1a). A mixture of glutathione S-transferase (GST)–CTD_{52} (RPB1 CTD containing all 52 repeats) and GST–CTD_9 (CTD containing 9 consensus repeats) was used in all reactions as the kinase substrates.

The various CYCT1 truncations did not substantially affect the ability of the associated CDK9 to autophosphorylate by producing the ATP-dependent mobility shift (Fig. 1a), nor did the truncations decrease the phosphorylation of CTD_9. However, upon truncation to a position at and beyond 533, CDK9 became largely unable to produce the hyperphosphorylated CTD_{52} (hereafter, IIo) as revealed by the anti-phospho-Ser5 antibody 3E8 (Fig. 1a). A similar pattern was also detected with the anti-phospho-Ser2 antibody 3E10 (Extended Data Fig. 1b). Thus, a region around position 533 in CYCT1 promoted the hyperphosphorylation by CDK9 of CTD_{52}.

To determine whether the CYCT1 C-terminal truncations affect the transcriptional activity of P-TEFb, we used the Gal4-tethering system to test fusion proteins containing the Gal4 DNA-binding domain attached to the various truncated forms of CYCT1 to activate luciferase expression from the HIV-1 promoter containing Gal4-binding sites (Fig. 1b). Correlating with the kinase results, progressive CYCT1 C-terminal truncations to and beyond position 533 markedly reduced P-TEFb transcriptional activity (Fig. 1b).

The coiled-coil domain (CCD, amino acids 379–430) and the histidine-rich domain (HRD, amino acids 480–550) exist around position 533 (Fig. 1a). In kinase reactions, only the deletion of HRD (ΔHRD) but not CCD (ΔCCD) blocked the hyperphosphorylation by CDK9 of CTD_{52} (Fig. 1c). Consistently, HRD but not CCD was required for P-TEFb transcriptional activity in the Gal4-tethering assay (Fig. 1d). Dependence on HRD for Tat/P-TEFb-activated HIV-1 elongation has previously been proposed[7], and here confirmed by our finding that wild-type CYCT1 but not CYCT1ΔHRD could effectively rescue the RNAi-knockdown of endogenous CYCT1 expression to produce long (promoter-distal) but not short (promoter-proximal) viral transcripts (Fig. 1e).

The importance of the HRD for the transcriptional activity of P-TEFb was generalized to cellular genes, as CYCT1ΔHRD produced significantly less mRNA from four representative immediate early genes—*FOS*, *JUNB*, *MYC* and *EGR1*—as well as *HSP70-1* (also known as *HSPA1A*) under both basal and heat-shock conditions (Extended Data Fig. 2), all of which require P-TEFb for optimal transcription. Co-immunoprecipitation analysis (Fig. 1f and Extended Data Fig. 2e) shows that the decreased activity of CYCT1ΔHRD was not due to any substantial change in binding to the major P-TEFb partners, including BRD4 and subunits of the super elongation complex and the 7SK small nuclear ribonucleoprotein particle that regulate P-TEFb activity[1,4,5].

What is the mechanism by which the HRD promotes the activity of P-TEFb? The first hint came from the finding that wild-type CYCT1 in HeLa nuclei was more resistant to salt extraction, as compared to CYCT1ΔHRD (Fig. 2a), which suggests that the HRD promoted the retention of CYCT1 in the nucleus. To determine more precisely the location and dynamics of this retention, we performed fluorescence recovery after photobleaching (FRAP) analysis of bindings of Halo-tagged wild-type CYCT1 and CYCT1ΔHRD to a gene array activated by reverse tetracycline-controlled transactivator (Fig. 2b). This array contains about 200 copies of an integrated transgene marked by YFP–Lac repressor bound to the *lac* operator, which enables FRAP to be performed at this spot[8]. We observed significantly faster FRAP recovery for CYCT1ΔHRD on the array than for wild-type CYCT1 (Fig. 2c). Further quantitative analysis (Extended Data Fig. 3a, b and Supplementary Information) revealed that this was mainly due to longer apparent residence time of wild-type CYCT1 (τ_{off} ~ 56 s), as compared to CYCT1ΔHRD (τ_{off} ~ 4 s).

[1]School of Pharmaceutical Sciences, Xiamen University, Xiamen, China. [2]Department of Molecular and Cell Biology, University of California, Berkeley, CA, USA. *e-mail: qzhou@berkeley.edu

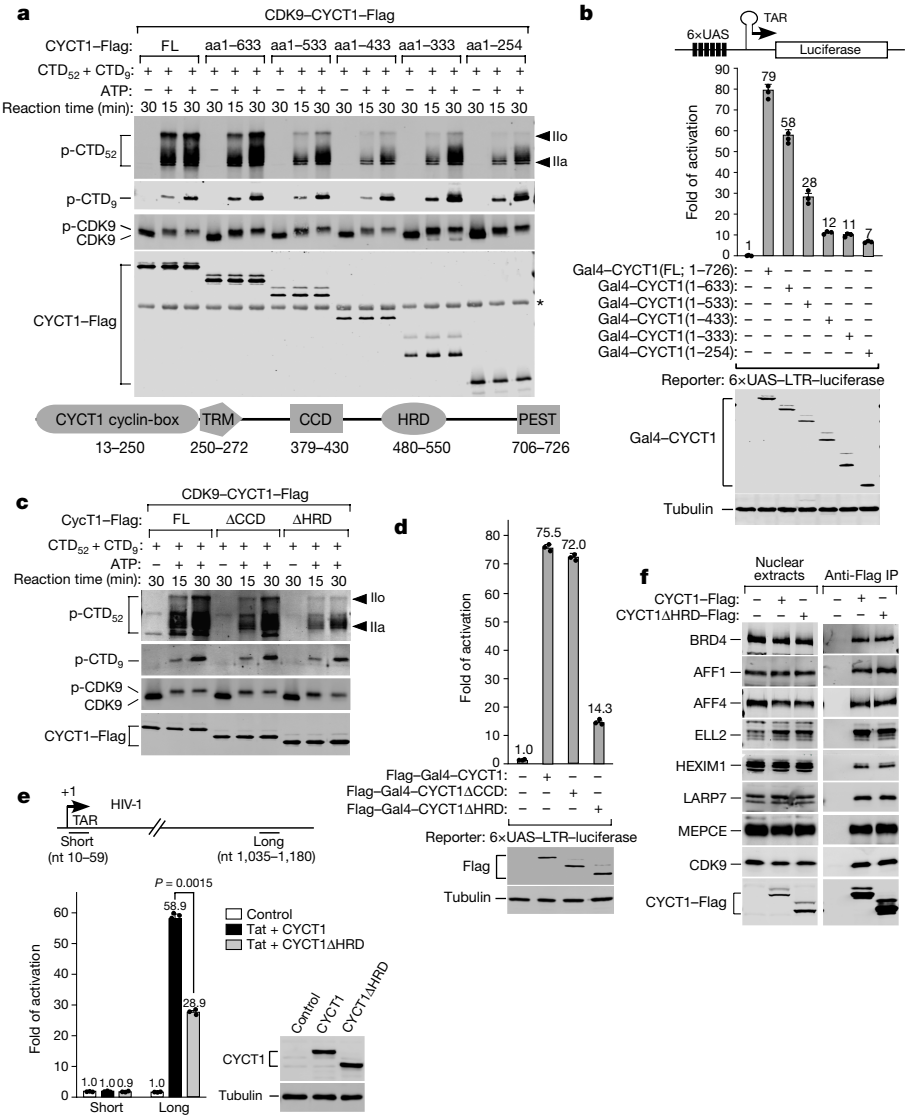


Fig. 1 | The CYCT1 HRD promotes hyperphosphorylation of Pol II CTD_{52} and activation of transcription, but not interaction with P-TEFb partners, by CDK9. a, **c**, P-TEFb containing the indicated CYCT1–Flag proteins (domain structure at bottom) were tested in kinase reactions containing GST–CTD_{52} and GST–CTD_9 as the substrates. Western blotting was performed to detect the phosphorylated products (p-CTD_{52} and p-CTD_9) and P-TEFb components. IIo and IIa, hyper- and hypo-phosphorylated CTD_{52}, respectively. *, a non-specific band. **b**, **d**, Plasmids expressing the indicated Gal4–CYCT1 fusions were co-transfected into HeLa cells with a HIV-1 LTR–luciferase reporter construct containing the Gal4 upstream activation sequences (UAS). Luciferase activities (mean ± s.d., $n = 3$) in cell extracts were measured and compared to the activity in the first lane (set to 1). Levels of indicated proteins in extracts were examined by western blotting. **e**, HeLa-based CYCT1 knockdown cells containing an integrated HIV-1 provirus were transfected with the indicated Tat + CYCT1-expressing plasmids or a control vector. The indicated proteins were detected by western blotting. HIV-1 mRNA levels at the promoter-proximal 'short' and promoter-distal 'long' positions were analysed by qRT–PCR (mean ± s.d., $n = 3$), with signals normalized to those in control cells (set to 1) and *P* value from two-tailed Student's *t*-test. **f**, Nuclear extracts of HeLa cells expressing the indicated CYCT1–Flag proteins and anti-Flag immunoprecipitates from nuclear extracts were analysed by western blotting. All western blots are representative of three independent experiments. For gel source data, see Supplementary Fig. 1. aa, amino acids; FL, full length; TRM, Tat/TAR-recognition motif; PEST, proline, glutamic acid, serine and threonine-rich sequence; nt, nucleotides.

To generalize the FRAP finding to endogenous genes (active and inactive) and cross-validate it using an orthogonal technique, we performed single-particle tracking in U2OS cells of Halo-tagged wild-type CYCT1 and CYCT1ΔHRD proteins at 95 Hz to determine whether the HRD affects CYCT1 diffusion dynamics and bound fractions[9,10] (Fig. 2d, Extended Data Fig. 3c–e, Supplementary Video 1 and Supplementary Information). Consistent with the HRD-facilitated association of CYCT1 with chromatin and the transcriptional machinery, CYCT1ΔHRD diffused faster and showed a smaller bound fraction (8.8%) than did wild-type CYCT1 (13.0%). Next, we conducted single-particle tracking using a longer exposure time (500 ms; Supplementary Video 2) to blur out fast-diffusing molecules and focus on bound molecules that were presumably located at predominantly activated genes. We observed a wide distribution of CYCT1 binding events that probably encompassed both specific and non-specific interactions (Fig. 2e). Because this distribution did not fit well with models that assumed one or two single rate-limiting steps, we could not reliably attribute a single residence time. Nevertheless, consistent with the FRAP result, wild-type CYCT1 showed binding events that were significantly more stable than those of CYCT1ΔHRD (Fig. 2e). Together, the single-particle tracking and FRAP results underscore the key role of the HRD in promoting the binding of P-TEFb to activated genes to phosphorylate CTD_{52}.

What could be the direct target of the HRD on activated genes? Consistent with a previous report that a CYCT1 mutant that lacked the HRD failed to bind the CTD[7], wild-type GST–CYCT1

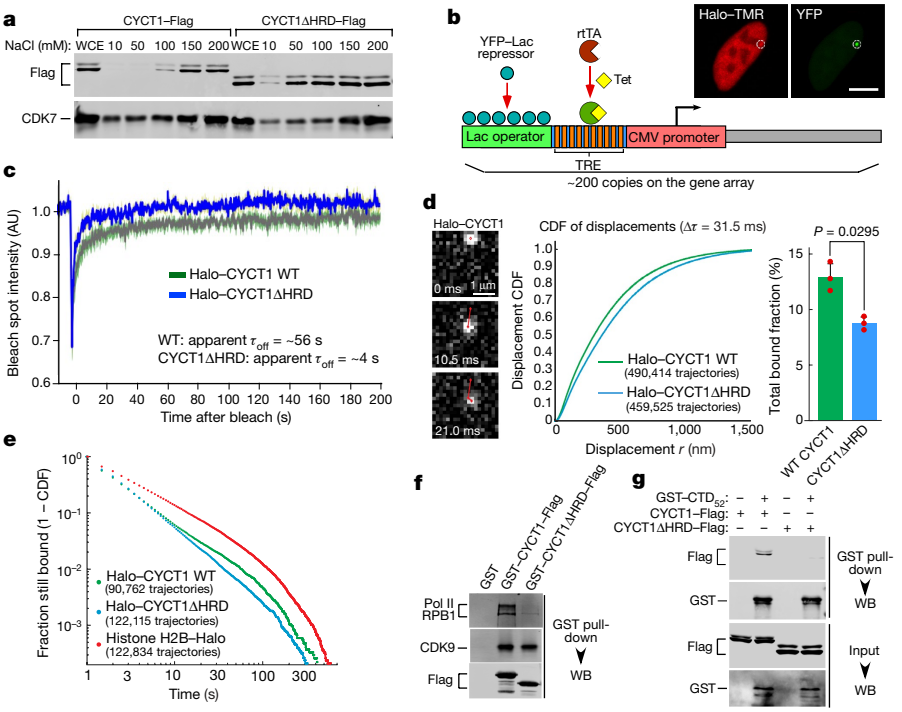


Fig. 2 | CYCT1 HRD contributes to the stable engagement of P-TEFb with activated gene array and endogenous genes, as well as direct binding to CTD by P-TEFb. a, Nuclei of HeLa cells expressing indicated CYCT1–Flag were extracted with increasing NaCl concentrations. The whole cell extract (WCE) and soluble fractions were examined by western blotting. **b**, Diagram of the gene array analysed by FRAP in **c**. Tet, tetracycline; rtTA, reverse tetracycline-controlled transactivator; TRE, tetracycline response element. The fluorescence images show the Halo–TMR and YFP emissions collected in separate channels. The small circle indicates the bleach spot. **c**, Fluorescence recovery plots of indicated Halo–CYCT1 photobleached at the spot containing the activated gene array. The intensities were normalized to pre-bleach values and are shown at various time points after the bleach. Inferred residence times are also given. AU, arbitrary unit. **d**, Single-particle tracking (1-ms excitation pulse; 95 Hz) of Halo-tagged wild-type CYCT1 (21 cells) or CYCT1ΔHRD (21 cells) labelled with photoactivatable Janelia Fluor 549 in U2OS cells. Left, examples of raw images with trajectory overlaid in red. Middle, cumulative distribution function (CDF) of displacements at $\Delta\tau = 31.5$ ms. Right, bound fraction inferred from three-state model fitting (mean ± s.e.m., $n = 3$. *P* value from two-tailed Student's *t*-test). **e**, Distribution of single-molecule binding times, uncorrected for photobleaching. Bound molecules (Janelia Fluor 646) were tracked using a long 500-ms exposure time. Plot shows 1 − CDF for trajectories captured for at least two frames. The constitutively bound histone H2B–Halo shows the photobleaching limit. Data from three independent replicates were merged and plotted. **f**, Immobilized GST or GST fusions were incubated with HeLa nuclear extract. The bound proteins were detected by western blotting (WB). **g**, Immobilized GST–CTD_{52} was incubated with the indicated affinity-purified CYCT1–Flag. The bound proteins and 2.5% of input were analysed by western blotting. All western blots are representative of three independent experiments. For gel source data, see Supplementary Fig. 1.

precipitated more Pol II from HeLa nuclear extracts than did GST–CYCT1ΔHRD (Fig. 2f). Moreover, GST–CTD pulled down P-TEFb containing wild-type CYCT1, but not CYCT1ΔHRD (Fig. 2g). Finally, a direct and HRD-dependent interaction was detected between a recombinant CYCT1 C-terminal fragment and mCherry–CTD in both the GST pulldown and glycerol gradient formats (Extended Data Fig. 8a, b). Thus, P-TEFb uses the HRD to directly target the CTD.

In addition to P-TEFb, CDK7– CYCH in TFIIH is also a CTD kinase for Pol II to clear the promoter during transcription initiation[1]. Unlike CYCT1, CYCH lacks a long C-terminal region and HRD. Compared to CDK9–CYCT1–Flag, affinity-purified CDK7–CYCH–Flag produced a markedly lower level of CTD_{52} IIo in kinase reactions (Fig. 3a), which was confirmed with recombinant CAK (CDK7–CYCH–MAT1) and P-TEFb (Millipore) in a time- and dosage-dependent manner (Extended Data Fig. 4a, b).

Appending a CYCT1 C-terminal fragment (amino acids 401–630) containing the HRD to CYCH markedly increased hyperphosphorylation by CDK7 of CTD_{52} (Fig. 3b), and partially rescued the CYCT1 knockdown to support HIV-1 transcription (Fig. 3c). However, when the HRD was deleted from the CYCH + CYCT1 chimaera, both the production of CTD_{52} IIo and the rescue of CYCT1 knockdown were mostly abolished (Fig. 3b, c), which indicates the importance of the HRD to these processes. Consistently, the chimaera, but not wild-type CYCH, was precipitated by GST–CTD (Fig. 3d).

We next investigated whether other CTD kinases also use an HRD to target Pol II for hyperphosphorylation. We noticed that the kinase DYRK1A, a CTD kinase that controls transcription of selected growth-related genes[6] and that is associated with Down syndrome, also contains an HRD (Fig. 3e). The rest of DYRK1A and CYCT1 are non-homologous. In kinase reactions, wild-type DYRK1A and DYRK1AΔHRD autophosphorylated to a similar extent, but only wild-type DYRK1A efficiently hyperphosphorylated CTD_{52} (Fig. 3e). In a co-immunoprecipitation assay employing three different anti-RPB1 antibodies, wild-type DYRK1A precipitated more Pol II than did DYRK1AΔHRD (Fig. 3f and Extended Data Fig. 4c). The precipitation of DCAF7, which binds DYRK1A N terminus[11], was unaffected by DYRK1AΔHRD. Thus, in addition to CYCT1, at least one other kinase also requires a functional HRD to efficiently bind and hyperphosphorylate CTD_{52}.

The HRD is a low-complexity domain owing to the overrepresentation of only a single amino acid. This domain, including a central cluster of multiple consecutive histidines, is highly conserved in vertebrate CYCT1 (Extended Data Fig. 5a). Using the prediction program IUPred[12] (http://iupred.enzim.hu), we found that the human CYCT1 HRD displays the highest disorder tendency in a broader intrinsically disordered region (IDR) that lacks well-defined structure[13] and overlaps with the CYCT1 C-terminal region (Extended Data Fig. 5b). Recently, the IDRs—especially those containing a low-complexity domain—have been shown to promote liquid–liquid phase separation,

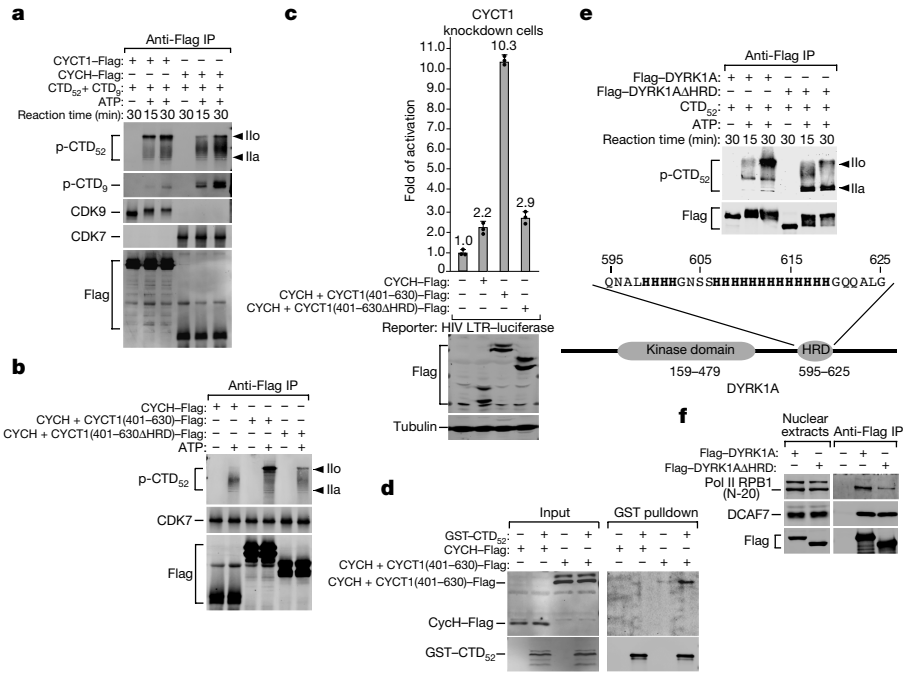


Fig. 3 | Upon fusion with CYCT1 HRD, CYCH promotes CDK7 hyperphosphorylation of Pol II CTD; DYRK1A also contains a functional HRD to target and hyperphosphorylate CTD.
a, **b**, **e**, The indicated anti-Flag immunoprecipitates (IP) were tested in kinase reactions as in Fig. 1a. **c**, The CYCT1 knockdown cells were co-transfected with HIV-1 LTR-luciferase reporter construct and plasmids expressing the indicated proteins. Luciferase activities were measured and analysed as in Fig. 1b. Data are mean ± s.d., $n = 3$. **d**, Immobilized GST–CTD_{52} was incubated with nuclear extract of HeLa cells expressing the indicated proteins. The bound proteins and 2.5% of input were analysed by western blotting. **f**, Nuclear extract of HeLa cells expressing the indicated proteins and anti-Flag immunoprecipitates from nuclear extract were analysed by western blotting. All western blots are representative of three independent experiments. For gel source data, see Supplementary Fig. 1.

which probably drives the formation of intracellular membrane-less organelles[14] for compartmentalized biochemical reactions[13].

In vitro, phase separation is reversible[13] and influenced by several parameters (for example, temperature, ionic strength, post-translational modifications and so on), which enable liquid droplets to form upon reaching a certain threshold and to quickly dissemble when pushed in the opposite direction[15]. To determine whether T1-IDR (amino acids 462–654)—the longest, HRD-containing CYCT1 IDR—could phase separate in an HRD-dependent manner, we purified the GFP–T1-IDR and GFP–T1-IDRΔHRD fusions from *Escherichia coli* (Extended Data Fig. 5c). At 150 mM NaCl, the protein solutions remained translucent. When lowered to 37.5 mM, a typical concentration used to induce phase separation in vitro[16,17], the wild-type GFP–T1-IDR solution immediately turned opaque, whereas GFP–T1-IDRΔHRD showed no change (Extended Data Fig. 5e).

Under a microscope, wild-type GFP–T1-IDR spontaneously formed micrometre-sized, spherical droplets, whereas GFP–T1-IDRΔHRD produced only a low level of irregular aggregates (Fig. 4a). Similarly, the GFP fusion that contained the longest IDR (amino acids 491–686) of DYRK1A also formed droplets in an HRD-dependent manner (Extended Data Fig. 6). Notably, the GFP–T1-IDR droplets quickly disappeared once NaCl returned to 150 mM (Extended Data Fig. 5f). Furthermore, 1,6-hexanediol, a compound that is known to perturb weak hydrophobic interactions to disassemble structures that exhibit liquid-like properties[18,19], completely blocked droplet formation (Extended Data Fig. 5g). Finally, when nine histidines within the histidine cluster in CYCT1 HRD were changed to alanines (Extended Data Fig. 7a), the resulting GFP–T1-IDR(9A) mutant formed only tiny droplets (Fig. 4b), which indicates that the histidines are essential for phase separation. Functionally, the mutation decreased phosphorylation by P-TEFb of CTD_{52}, and the transcriptional activity of P-TEFb (Extended Data Fig. 7c, d).

In cells, both ectopically expressed CYCT1–Flag or eGFP–CYCT1 and endogenous CYCT1 displayed a punctuated staining pattern inside the nuclei (Fig. 4c–e and Extended Data Fig. 5h). This has previously been attributed to the localization of P-TEFb to the nuclear speckles enriched with transcription and splicing factors[20,21]. Importantly, HRD and its histidine cluster were needed to target CYCT1 to the speckles (Fig. 4c and Extended Data Figs. 5h, 7e), and the same has also been reported for DYRK1A[22]. Mirroring its disruption of the GFP–T1-IDR droplets in vitro, 1,6-hexanediol also quickly dissembled the CYCT1 nuclear speckles (Fig. 4d). Finally, when performing time-lapse phase-contrast imaging of cells expressing eGFP–CYCT1, we observed multiple cells in which CYCT1 speckles displayed at least one fusion event within the period of a few minutes, demonstrating their dynamic and liquid-like properties (Fig. 4e).

Following previous findings[23,24], we asked whether the HRD-containing IDR could recruit the CTD into phase-separated droplets, given their demonstrated direct interaction. Indeed, although the CTD itself is a low-complexity sequence, recombinant mCherry–CTD[23] (Extended Data Fig. 5d) alone did not phase separate, but it was readily incorporated into droplets when incubated together with GFP–T1-IDR (Fig. 4f).

During transcription, the CTD is phosphorylated first by CAK in TFIIH and then by P-TEFb[1,4,5]. Pre-phosphorylation by CAK (Extended Data Fig. 8c) not only enhanced the incorporation of mCherry–CTD into the GFP–T1-IDR droplets, but also promoted phase separation overall by producing bigger and brighter droplets (Fig. 4f). Consistently, the pre-phosphorylated GST–CTD precipitated more GFP–T1-IDR (Extended Data Fig. 8d). Underscoring the physiological relevance of these observations, hyperphosphorylated Pol IIo is known to preferentially localize in the nuclear speckles[25].

Although both CTD hyperphosphorylation by CDK9 and HRD-mediated droplet formation were largely inhibited by 2.5% 1,6-hexanediol, the HRD–CTD binding (and CDK9–CYCT1 interaction) was not substantially inhibited until 1,6-hexanedil was at 10% (Fig. 4g–i). This important difference suggests that phase separation, which is caused by weak, multivalent interactions among the HRDs and is easily disrupted by 1,6-hexanediol, is critical for the hyperphosphorylation of the CTD.

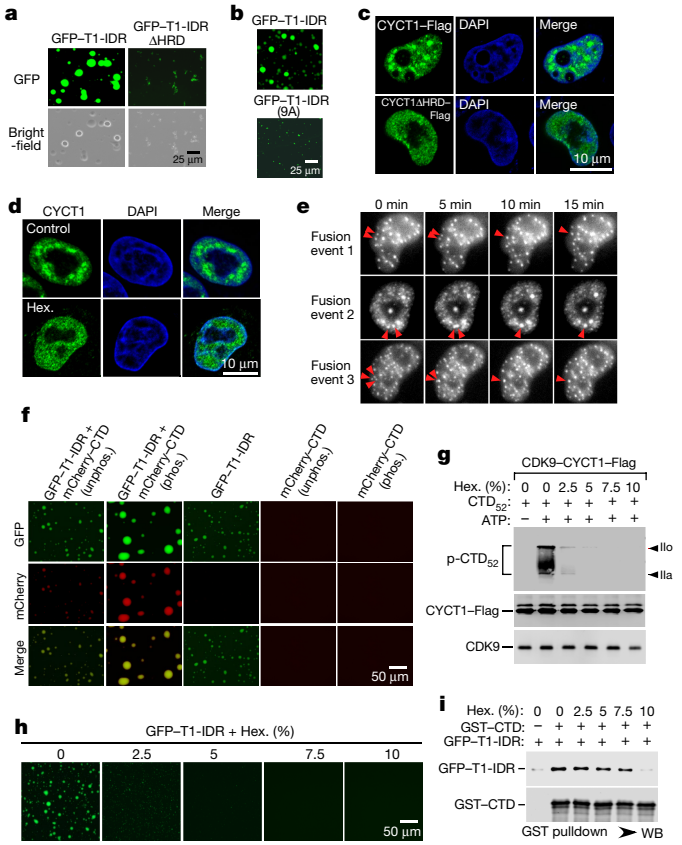


Fig. 4 | The hyperphosphorylation of CTD by P-TEFb is promoted by CYCT1 IDR, which forms phase-separated droplets and/or speckles in an HRD-dependent manner and recruits the CTD into these compartments. a, **b**, Solutions containing the indicated proteins at 6 mg ml^{-1} (**a**) or 2 mg ml^{-1} (**b**) were examined with a microscope for fluorescence (GFP) or under white light (bright field). **c**, HeLa cells expressing wild-type CYCT1–Flag or CYCT1ΔHRD–Flag were examined by indirect immunofluorescence staining with anti-Flag monoclonal antibody. DNA was counterstained using DAPI. **d**, HeLa cells were treated with 10% 1,6-hexanediol (Hex.) or without1,6-hexandiol (control) for 1 min and then analysed by immunofluorescence staining with anti-CYCT1 antibody. **e**, HeLa cells expressing eGFP–CYCT1 were examined by time-lapse phase-contrast imaging. Three separate cell nuclei are shown, in which CYCT1 speckles underwent spontaneous fusions as indicated by the arrows. **f**, Solutions containing GFP–T1-IDR at 3 mg ml^{-1} and/or phosphorylated (phos.) or unphosphorylated (unphos.) mCherry–CTD at 1.2 mg ml^{-1} were examined under a fluorescence microscope as in **a**. **g**, Kinase reactions as in Fig. 1a were performed in the presence of the indicated concentrations of 1,6-hexandiol. p-CTD$_{52}$ was detected by western blotting. Levels of CYCT1–Flag and its bound CDK9 in each reaction are also shown. **h**, 1,6-Hexanediol was present at the indicated concentrations in solutions containing 6 mg ml^{-1} wild-type GFP–T1-IDR and 37.5 mM NaCl, which were examined under a fluorescence microscope. **i**, Immobilized GST–CTD was incubated with recombinant GFP–T1-IDR in reactions containing the indicated concentrations of 1,6-hexandiol. The bait and bound proteins were analysed by western blotting. All western blots are representative of three independent experiments. For gel source data, see Supplementary Fig. 1.

On the other hand, the relatively drug-resistant HRD–CTD binding is probably key to the recruitment of the CTD to CDK9 in droplets, but is insufficient to establish the optimal environment for hyperphosphorylation (Extended Data Fig. 9).

In summary, our data indicate that in at least two kinases—CYCT1 of P-TEFb and DYRK1A—the HRD, a low-complexity domain of previously unknown function, promotes the hyperphosphorylation of the CTD by targeting the CTD as well as inducing phase separation in vitro and in cells. The phase-separated droplets and speckles compartmentalize the kinase and substrate to enable highly efficient reactions, which results in the hyperphosphorylation of the CTD and robust transcriptional elongation and RNA processing.

FUS and TAF15, which are two proteins that contain low-complexity domain and that are active in transcription initiation, have previously been shown to form phase-separated hydrogels that trap the CTD[23,24]. Additionally, the bidirectionally transcribed enhancers and resulting antisense transcripts have been proposed to control initiation in part through phase separation[26]. What was unknown until now is whether any transcription factors and the CTD are involved in droplet formation after initiation[1]. Our finding that phase separation is induced by CYCT1 of P-TEFb (a well-defined transcription elongation factor) and DYRK1A (a probable gene-specific elongation factor[6]) has expanded the regulatory roles of phase separation to the next stage of the transcription cycle. Furthermore, these studies show that some key initiation and elongation factors that phase separate should no longer be viewed as passive passengers waiting to be picked up by the CTD. Rather, they have active roles in recruiting Pol II through multivalent interactions to their droplets and/or speckles that function as hubs where much of transcription and RNA processing is dynamically controlled.

Data availability

Uncropped scans of all western blots are provided in Supplementary Figure 1. The raw slowSPT and spaSPT data are freely available in Spot-On readable CSV and Matlab formats in the form of single-molecule trajectories at Zenodo (https://zenodo.org/record/1215836). The Spot-On Matlab code is available, together with a step-by-step guide, at Gitlab (https://gitlab.com/tjian-darzacq-lab/spot-on-matlab). All other data are available from the corresponding author on reasonable request.

Online content

Any Methods, including any statements of data availability and Nature Research reporting summaries, along with any additional references and Source Data files, are available in the online version of the paper at https://doi.org/10.1038/s41586-018-0174-3.

Received: 11 November 2017; Accepted: 2 May 2018;
Published online 30 May 2018.

1. Harlen, K. M. & Churchman, L. S. The code and beyond: transcription regulation by the RNA polymerase II carboxy-terminal domain. *Nat. Rev. Mol. Cell Biol.* **18**, 263–273 (2017).
2. Eick, D. & Geyer, M. The RNA polymerase II carboxy-terminal domain (CTD) code. *Chem. Rev.* **113**, 8456–8490 (2013).
3. Zaborowska, J., Egloff, S. & Murphy, S. The pol II CTD: new twists in the tail. *Nat. Struct. Mol. Biol.* **23**, 771–777 (2016).
4. Zhou, Q., Li, T. & Price, D. H. RNA polymerase II elongation control. *Annu. Rev. Biochem.* **81**, 119–143 (2012).
5. Kwak, H. & Lis, J. T. Control of transcriptional elongation. *Annu. Rev. Genet.* **47**, 483–508 (2013).
6. Di Vona, C. et al. Chromatin-wide profiling of DYRK1A reveals a role as a gene-specific RNA polymerase II CTD kinase. *Mol. Cell* **57**, 506–520 (2015).
7. Taube, R., Lin, X., Irwin, D., Fujinaga, K. & Peterlin, B. M. Interaction between P-TEFb and the C-terminal domain of RNA polymerase II activates transcriptional elongation from sites upstream or downstream of target genes. *Mol. Cell. Biol.* **22**, 321–331 (2002).
8. Darzacq, X. et al. *In vivo* dynamics of RNA polymerase II transcription. *Nat. Struct. Mol. Biol.* **14**, 796–806 (2007).
9. Hansen, A. S., Pustova, I., Cattoglio, C., Tjian, R. & Darzacq, X. CTCF and cohesin regulate chromatin loop stability with distinct dynamics. *eLife* **6**, e25776 (2017).
10. Hansen, A. S. et al. Robust model-based analysis of single-particle tracking experiments with Spot-On. *eLife* **7**, e33125 (2018).
11. Xiang, J. et al. DYRK1A regulates Hap1–Dcaf7/WDR68 binding with implication for delayed growth in Down syndrome. *Proc. Natl Acad. Sci. USA* **114**, E1224–E1233 (2017).
12. Dosztányi, Z., Csizmok, V., Tompa, P. & Simon, I. IUPred: web server for the prediction of intrinsically unstructured regions of proteins based on estimated energy content. *Bioinformatics* **21**, 3433–3434 (2005).
13. Mitrea, D. M. & Kriwacki, R. W. Phase separation in biology; functional organization of a higher order. *Cell Commun. Signal.* **14**, 1 (2016).
14. Courchaine, E. M., Lu, A. & Neugebauer, K. M. Droplet organelles? *EMBO J.* **35**, 1603–1612 (2016).
15. Banani, S. F., Lee, H. O., Hyman, A. A. & Rosen, M. K. Biomolecular condensates: organizers of cellular biochemistry. *Nat. Rev. Mol. Cell Biol.* **18**, 285–298 (2017).

16. Lin, Y., Protter, D. S., Rosen, M. K. & Parker, R. Formation and maturation of phase-separated liquid droplets by RNA-binding proteins. *Mol. Cell* **60**, 208–219 (2015).
17. Ying, Y. et al. Splicing activation by Rbfox requires self-aggregation through its tyrosine-rich domain. *Cell* **170**, 312–323.e310 (2017).
18. Strom, A. R. et al. Phase separation drives heterochromatin domain formation. *Nature* **547**, 241–245 (2017).
19. Molliex, A. et al. Phase separation by low complexity domains promotes stress granule assembly and drives pathological fibrillization. *Cell* **163**, 123–133 (2015).
20. Herrmann, C. H. & Mancini, M. A. The Cdk9 and cyclin T subunits of TAK/P-TEFb localize to splicing factor-rich nuclear speckle regions. *J. Cell Sci.* **114**, 1491–1503 (2001).
21. Marcello, A. et al. Recruitment of human cyclin T1 to nuclear bodies through direct interaction with the PML protein. *EMBO J.* **22**, 2156–2166 (2003).
22. Salichs, E., Ledda, A., Mularoni, L., Albà, M. M. & de la Luna, S. Genome-wide analysis of histidine repeats reveals their role in the localization of human proteins to the nuclear speckles compartment. *PLoS Genet.* **5**, e1000397 (2009).
23. Kato, M. et al. Cell-free formation of RNA granules: low complexity sequence domains form dynamic fibers within hydrogels. *Cell* **149**, 753–767 (2012).
24. Kwon, I. et al. Phosphorylation-regulated binding of RNA polymerase II to fibrous polymers of low-complexity domains. *Cell* **155**, 1049–1060 (2013).
25. Mortillaro, M. J. et al. A hyperphosphorylated form of the large subunit of RNA polymerase II is associated with splicing complexes and the nuclear matrix. *Proc. Natl Acad. Sci. USA* **93**, 8253–8257 (1996).
26. Hnisz, D., Shrinivas, K., Young, R. A., Chakraborty, A. K. & Sharp, P. A. A phase separation model for transcriptional control. *Cell* **169**, 13–23 (2017).

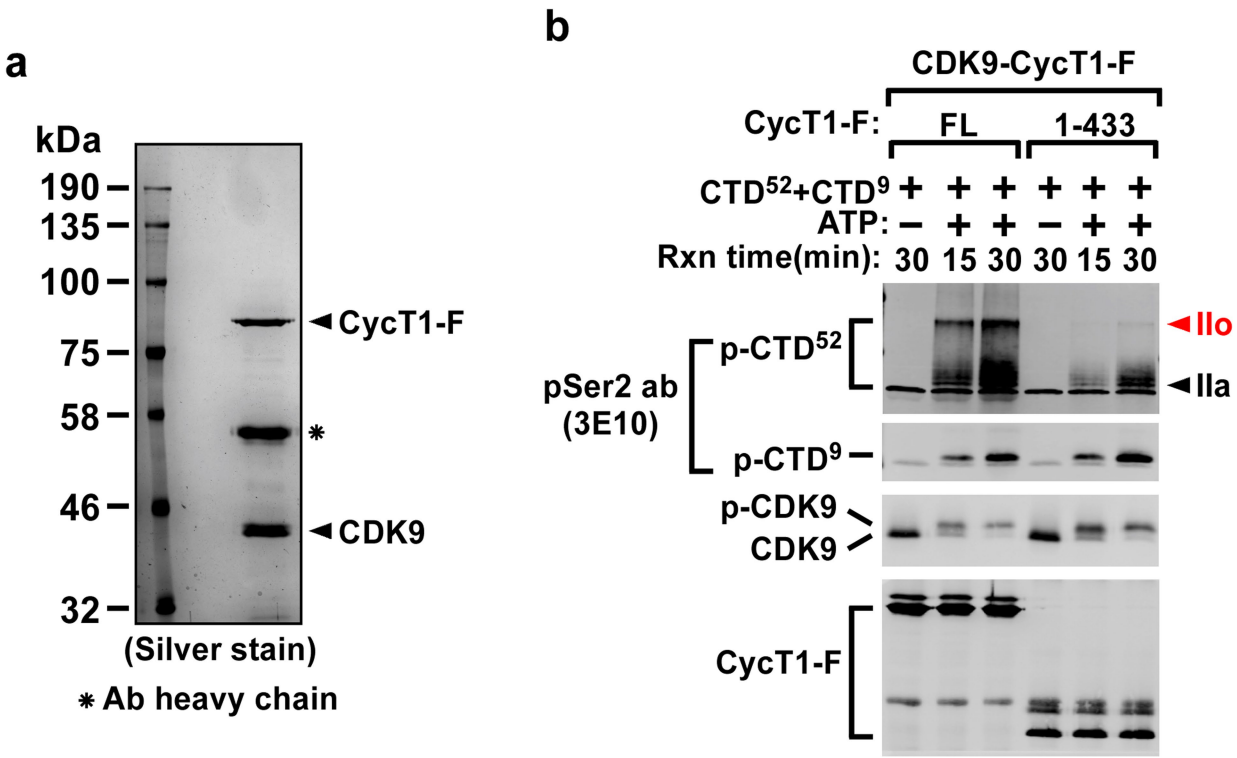
Acknowledgements We thank S. McKnight, M. Geyer, J. Hurley and their colleagues for providing the various expression plasmids, and U. Schulze-Gahmen for technical help. This work was supported by the National Institutes of Health grant R01AI041757 to Q.Z. and the California Institute of Regenerative Medicine grant LA1-08013 to X.D.

Reviewer information *Nature* thanks J. Lis, D. Taatjes and the other anonymous reviewer(s) for their contribution to the peer review of this work.

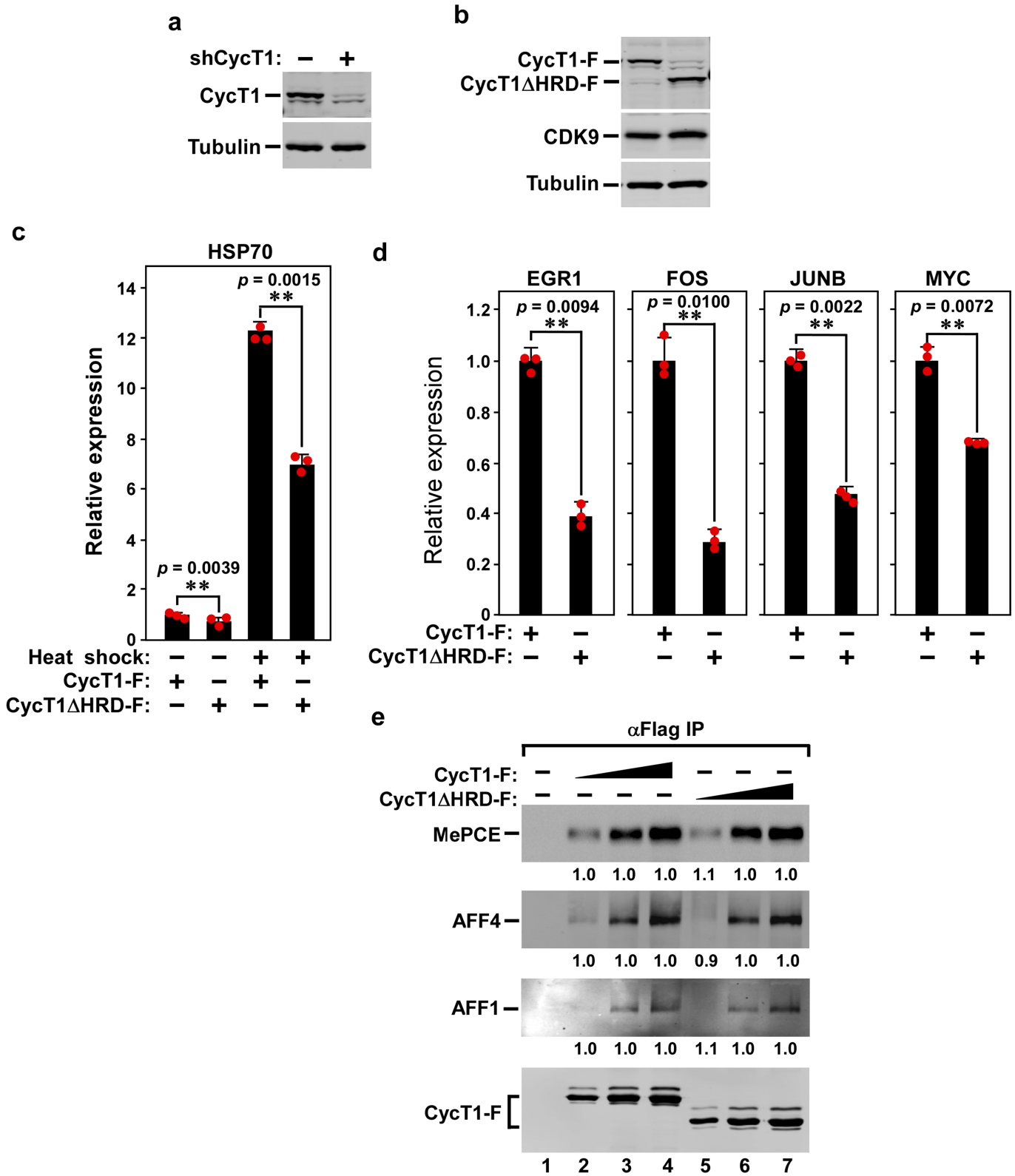
Author contributions H.L., X.D. and Q.Z. conceived the studies. H.L., D.Y. and R.L. performed kinase reaction assays, cell culture, immunofluorescence staining and droplet formation experiments and analyses. A.S.H. performed and analysed the single-particle tracking experiment. S.G. performed and analysed the FRAP assay. H.L. and A.H. performed and analysed the time-lapse phase-contrast imaging experiment. H.L., A.S.H. and Q.Z. wrote the manuscript, and all authors contributed ideas and reviewed the manuscript.

Competing interests The authors declare no competing interests.

Additional information
Extended data is available for this paper at https://doi.org/10.1038/s41586-018-0174-3.
Supplementary information is available for this paper at https://doi.org/10.1038/s41586-018-0174-3.
Reprints and permissions information is available at http://www.nature.com/reprints.
Correspondence and requests for materials should be addressed to Q.Z.
Publisher's note: Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Extended Data Fig. 1 | The CYCT1 HRD-dependent hyperphosphorylation of Pol II CTD_{52} by affinity-purified CDK9–CYCT1–Flag dimer can also be detected with anti-phospho-Ser2 antibody. a, Examination of the affinity-purified CDK9–CYCT1–Flag dimer containing wild-type CYCT1–Flag by SDS–PAGE and silver staining. The anti-Flag affinity-purification was performed under high salt plus detergent (1 M KCl + 1% NP-40) conditions to strip away all the P-TEFb-associated factors but keep the CDK9–CYCT1 interaction intact. **b**, Affinity-purified CDK9–CYCT1–Flag heterodimers containing the indicated CYCT1–Flag proteins were tested in kinase reactions containing a mixture of GST-fused CTD_{52} and CTD_9 as the substrates. The phosphorylated p-CTD_{52} and p-CTD_9 were detected by western blotting with the anti-phospho-Ser2 antibody 3E10. Although a very similar pattern of CTD phosphorylation was detected with both anti-phospho-Ser2 and anti-phospho-Ser5 antibodies, the unphosphorylated CTD present in the ATP(–) lanes was only detected by the former antibody, making the phospho-Ser5 antibody a preferred choice for detecting CTD phosphorylation in these kinase reactions.



Extended Data Fig. 2 | See next page for caption.

Extended Data Fig. 2 | The CYCT1 HRD is required for efficient transcription of human immediate early genes as well as the *HSP70-1* gene under both basal and heat-shock conditions. a, Confirmation by western blotting of the knockdown of endogenous CYCT1 expression in HEK293T cells expressing the *CYCT1* (also known as *CCNT1*)-specific shRNA (shCYCT1). **b**, Anti-Flag western blotting analysis of the expression of either wild-type CYCT1–Flag or CYCT1ΔHRD–Flag from the shCYCT1-resistant plasmid introduced into the knockdown cells. The α-tubulin and CDK9 levels were used as controls. **c**, The HRD-dependent transcription of four cellular immediate early genes. The mRNA levels of the indicated immediate early genes in the knockdown cells expressing wild-type CYCT1 or CYCT1ΔHRD–Flag (analysed in **b**) were examined by qRT–PCR, normalized to that of *GAPDH* and shown. The activity in the first column of each group was set to 1. Data are mean ± s.d., $n = 3$, and P values from two-tailed Student's t-test. **d**, The CYCT1 HRD is required for optimal *HSP70-1* transcription under both basal and heat-shock conditions. The HSP70 mRNA levels in the knockdown cells expressing wild-type CYCT1 or CYCT1ΔHRD were examined under heat-shock or non-heat-shock conditions, as in **c**. **e**, HeLa cells were transfected in twofold increments with the plasmid expressing either wild-type CYCT1–Flag or CYCT1ΔHRD–Flag. Anti-Flag immunoprecipitates (IP) from nuclear extracts were examined by western blotting for the proteins labelled on the left. The levels of the co-precipitated MEPCE, AFF4 and AFF1 were first quantified and then normalized to that of their corresponding CYCT1–Flag bait. The levels of MEPCE, AFF4 and AFF1 bound to low, middle and high levels of CYCT1ΔHRD–Flag were then divided by those of MEPCE, AFF4 and AFF1 bound to the corresponding levels of wild-type CYCT1–Flag and shown in lanes 5–7, with the numbers in lanes 2–4 set to 1 as a reference.

a

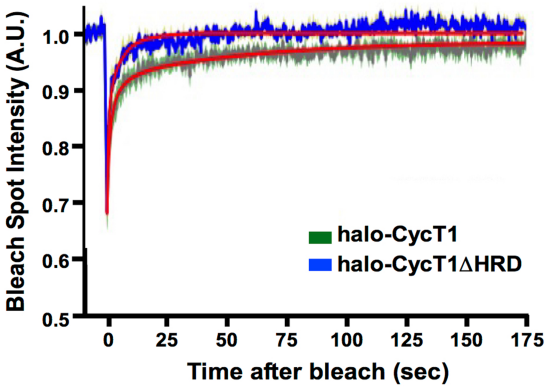
Estimates of parameters extracted from fitting of FRAP data (with ± standard errors) wherever applicable

	D_{app} ($\mu m^2 sec^{-1}$)	k^*_{on} (sec^{-1})	k_{off} (sec^{-1})	*Sum of residuals*
halo-CycT1 WT	1.66 ± 0.08	0.0044 ± 0.0007	0.0176 ± 0.0019	0.00173
halo-CycT1ΔHRD	4.15 ± 0.9	0.0657 ± 0.04	0.231 ± 0.07	0.00871

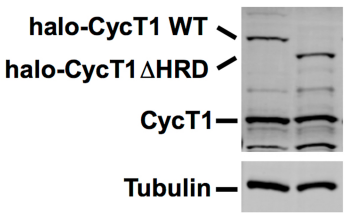
Parameters evaluated from the fitted FRAP data above with the corresponding ranges

	D_{app} ($\mu m^2 sec^{-1}$)	C_{eq}	τ_{on} (*sec*)	τ_{off} (*sec*)
halo-CycT1 WT	1.66 ± 0.08	~ 0.2	226 (*196-270*)	56 (*51-64*)
halo-CycT1ΔHRD	4.15 ± 0.9	~ 0.2	15 (*9-38*)	4 (*3-6*)

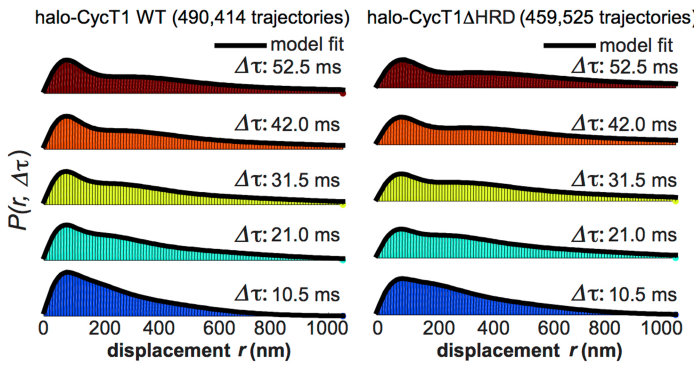
b



c



d

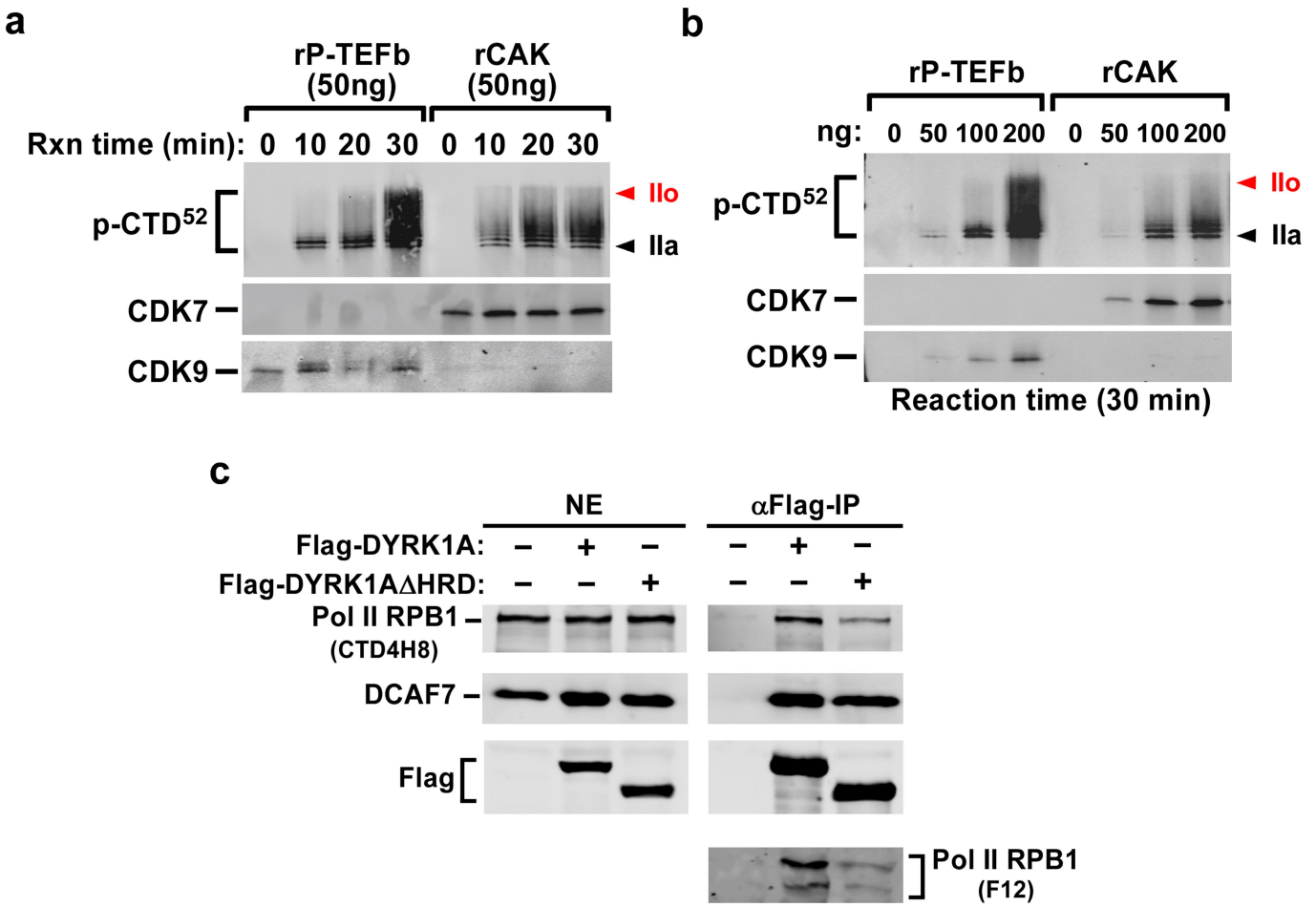


e

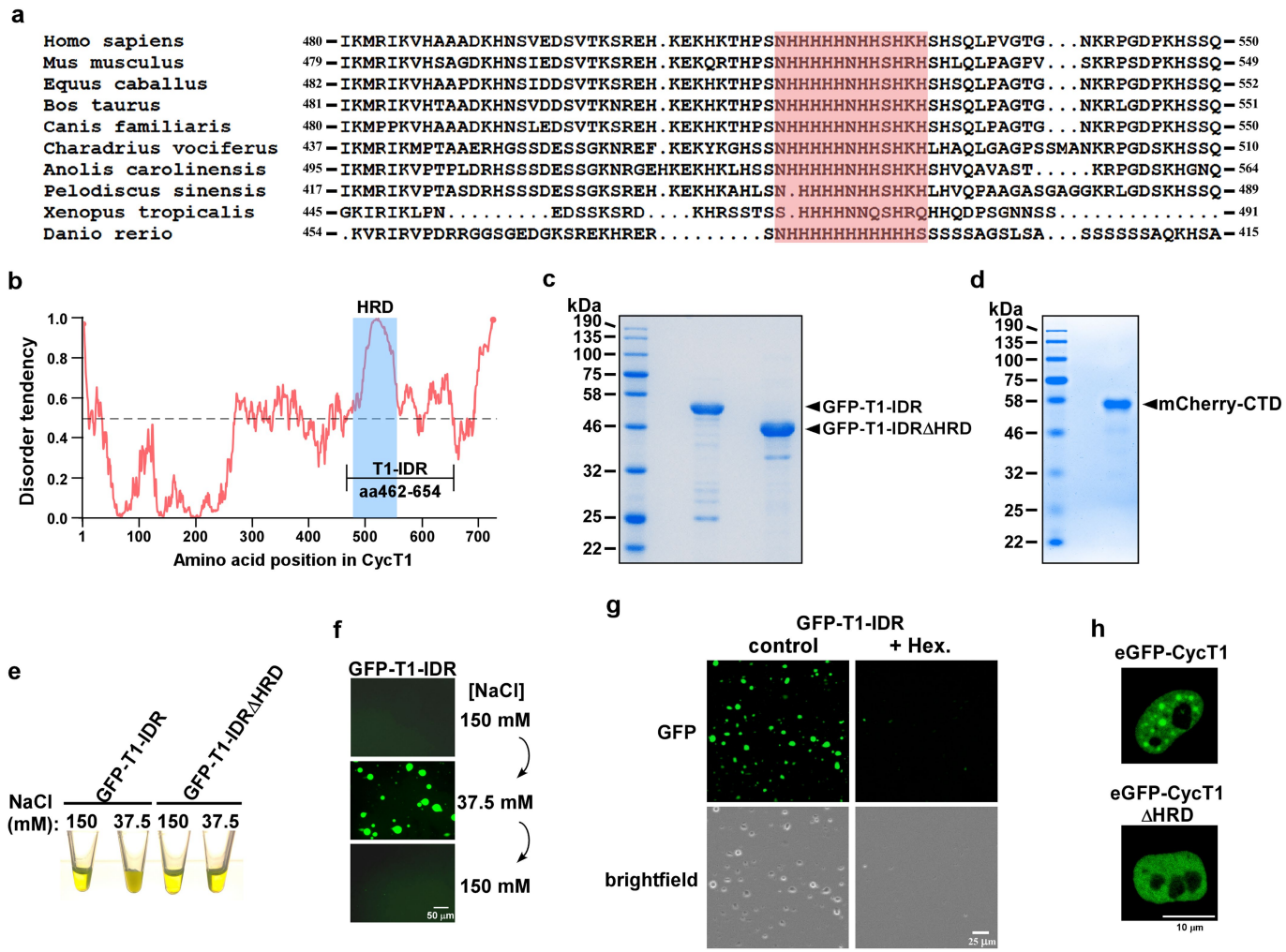
3-state Spot-On model fit (1 bound, 2 free states)

U2OS halo-CycT1	WT	ΔHRD
F_{BOUND}	13.0% ± 2.0%	8.9% ± 0.9%
D_{BOUND} (μm²/s)	0.04 ± 0.0	0.04 ± 0.0
F_{SLOW}	50.2% ± 4.3%	41.5% ± 9.5%
D_{SLOW} (μm²/s)	1.18 ± 0.10	1.39 ± 0.16
F_{FAST}	39.8% ± 4.0%	49.6% ± 10.4%
D_{FAST} (μm²/s)	5.70 ± 0.09	6.83 ± 1.01

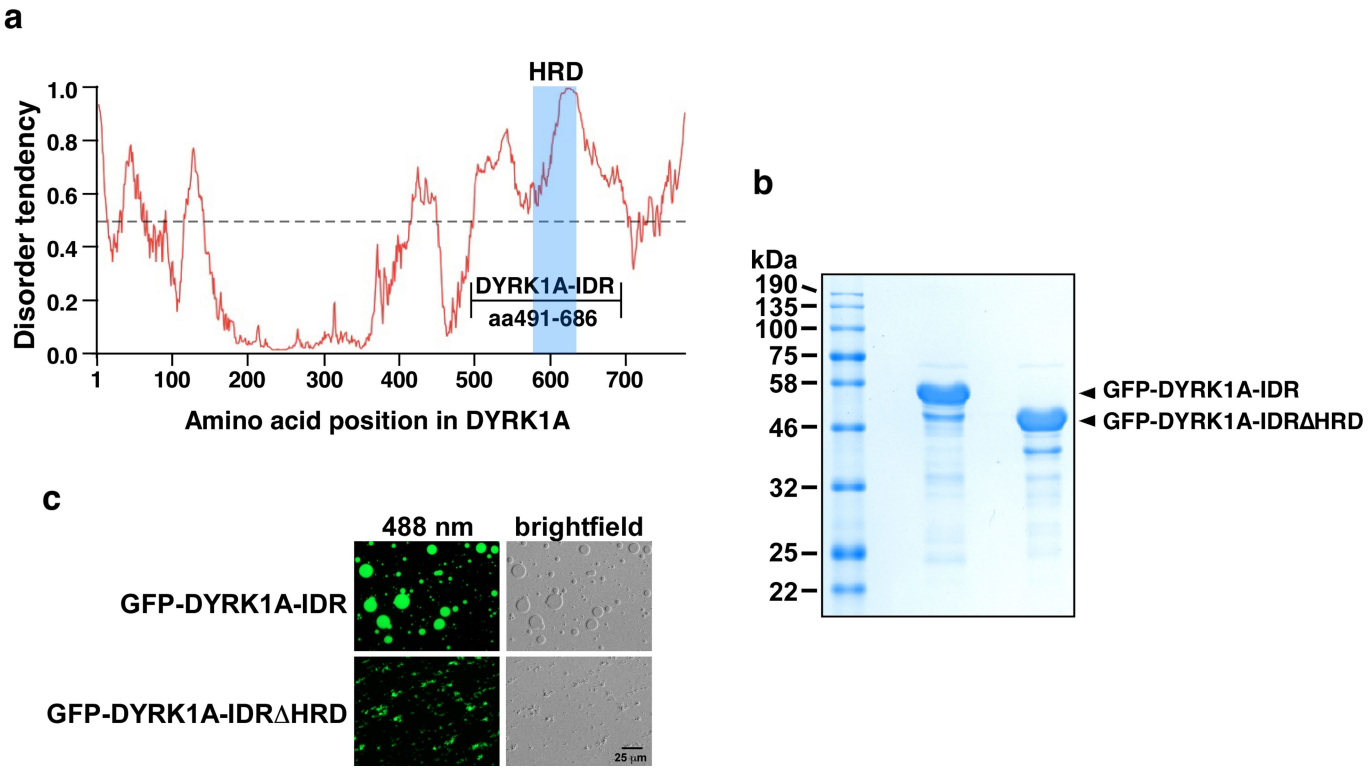
Extended Data Fig. 3 | The HRD promotes the binding of CYCT1 to activated gene expression array, endogenous genes and the RNA Pol II CTD. a, Estimates of parameters extracted from fitting of FRAP data (top) and parameters calculated from the fitted FRAP data (bottom). **b**, Model fit overlaid on raw FRAP data. A full description of how the reaction–diffusion FRAP model was fitted to the data is provided in Supplementary Information. **c**, Anti-CYCT1 western blotting analysis of the levels of endogenous CYCT1, and the stably expressed Halo–CYCT1 and Halo–CYCT1ΔHRD, in lysates of the engineered U2OS cell lines. The α-tubulin levels provided an internal control. **d**, Histograms of displacements at the indicated $\Delta\tau$ with three-state model fit overlaid. The three-state model is described in Supplementary Information. **e**, Table of best-fit parameters from fitting Spot-On to the raw displacements for three independent replicates (about 5–10 cells per replicate). Values in the table are mean ± s.d., $n = 3$.



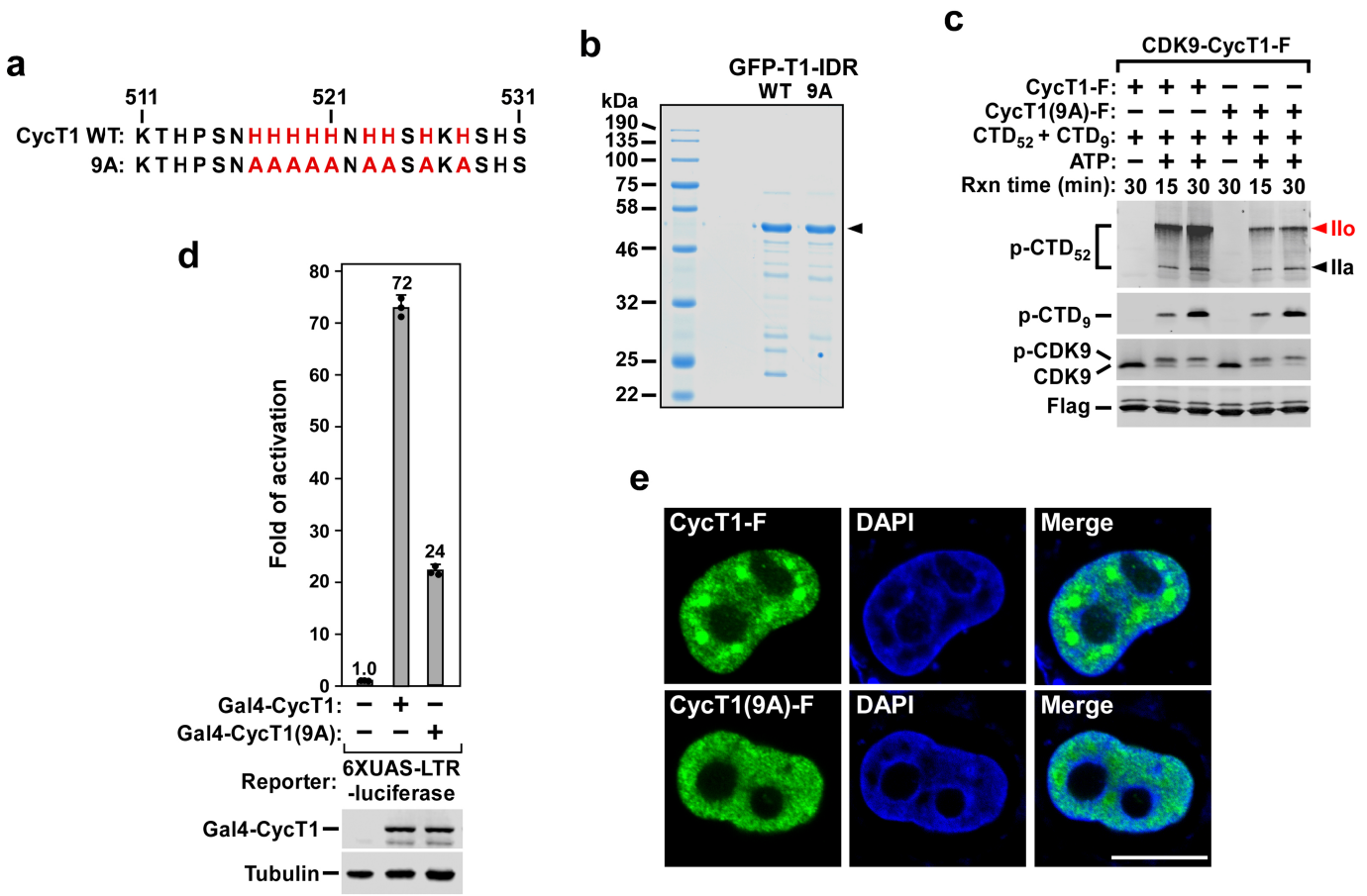
Extended Data Fig. 4 | Examination of recombinant P-TEFb and CAK for their phosphorylation of the Pol II CTD in kinase reactions and contribution of the DYRK1A HRD to DYRK1A–Pol II interaction. **a**, **b**, Compared to CDK9 in recombinant P-TEFb, CDK7 in recombinant CAK (CDK7–CYCH–MAT1) shows decreased ability to hyperphosphorylate CTD_{52} in a time- and dosage-dependent manner. The indicated amounts of baculovirus-produced recombinant P-TEFb or CAK (Millipore) were added to in vitro kinase reactions that also contained GST–CTD_{52} as the substrate. The reactions were performed for the indicated periods of time. The products were analysed by western blotting with the phospho-Ser5 antibody. CDK7 in recombinant CAK and CDK9 in recombinant P-TEFb were also examined by western blotting. **c**, The deletion of the HRD causes DYRK1A to decrease interaction with RNA Pol II but not DCAF7. Nuclear extract (NE) of HeLa cells expressing the indicated proteins and anti-Flag immunoprecipitates derived from NE were analysed by western blotting with the various antibodies as labelled.



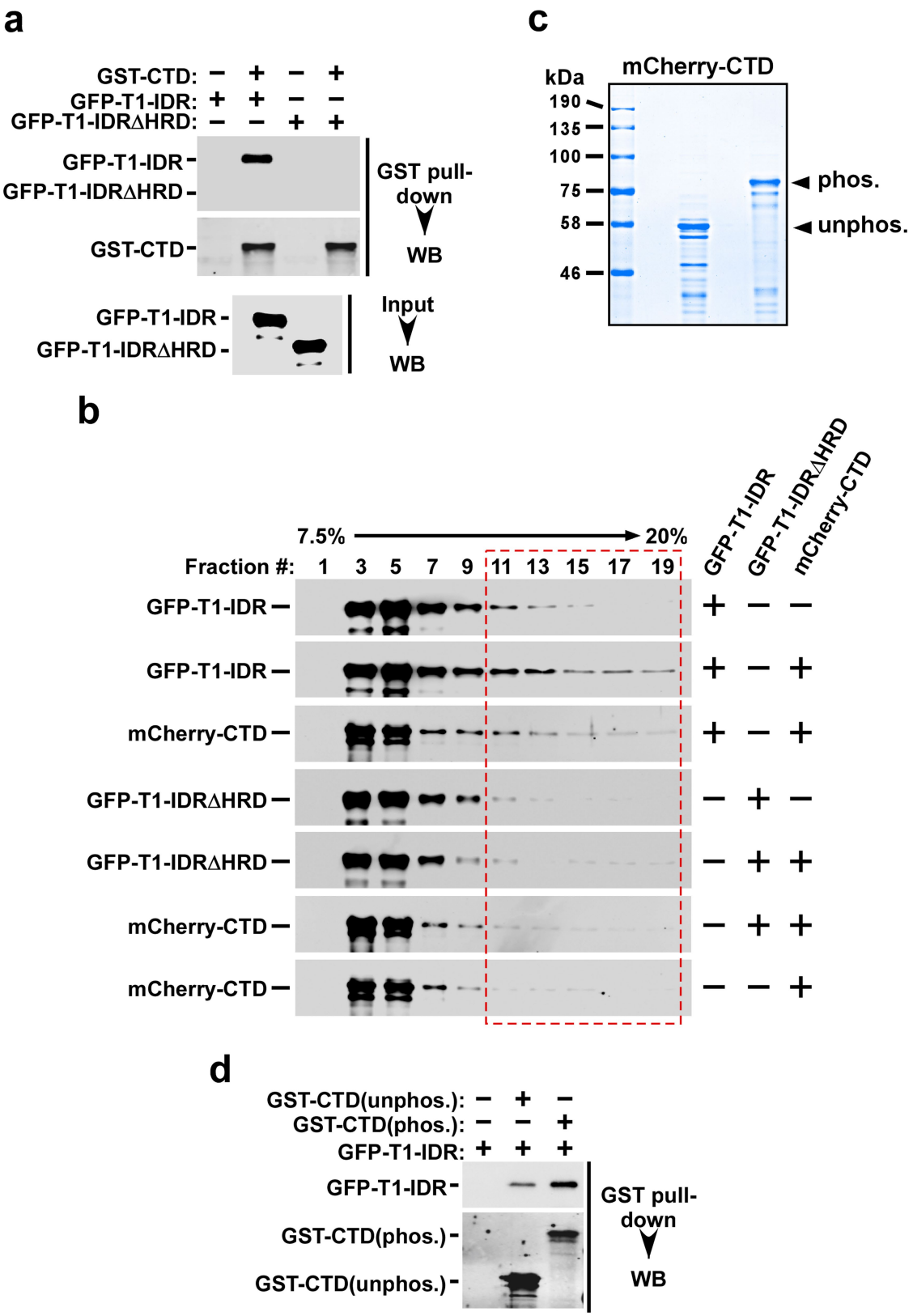
Extended Data Fig. 5 | GFP–T1-IDR purified from recombinant *E. coli* forms phase-separated liquid droplets that are HRD-dependent and highly sensitive to elevated salt concentration and exposure to 1,6-hexanediol. a, Alignment of CYCT1 amino acid sequences in the HRD regions among the indicated vertebrate species, with the central histidine cluster shaded in red. **b**, Intrinsic disorder tendency was predicted by IUPred across the entire length of CYCT1. The scores are assigned between 0 and 1, and a score above 0.5 indicates disorder. The HRD region is shaded in blue. The longest stretch of IDR is labelled as T1-IDR, and its boundaries are marked. **c**, **d**, The C-terminally Strep-tagged GFP–T1-IDR fusions (**c**) and N-terminally His-tagged mCherry–CTD fusion (**d**) were purified from recombinant *E. coli* BL21 cells (Supplementary Information). Ten micrograms of each of the fusion proteins was examined by SDS–PAGE followed by Coomassie blue staining. **e**, Protein solutions containing either wild-type GFP–T1-IDR or GFP–T1-IDRΔHRD at 6 mg ml^{-1} were adjusted to the indicated salt concentrations and their appearances in Eppendorf tubes are shown. **f**, NaCl concentrations in the wild-type GFP–T1-IDR solution were changed in the indicated sequential order and then examined under a fluorescence microscope. **g**, Protein solutions containing wild-type GFP–T1-IDR at 6 mg ml^{-1} were adjusted to 37.5 mM NaCl with or without 10% 1,6-hexanediol, and then examined under a fluorescence microscope. **h**, Live-cell images of a HeLa cell expressing either wild-type eGFP–CYCT1 or eGFP–CYCT1ΔHRD at levels similar to that of endogenous CYCT1.



Extended Data Fig. 6 | The longest stretch of IDR in DYRK1A promotes formation of phase-separated droplets in an HRD-dependent manner. **a**, Intrinsic disorder tendency was predicted by IUPred across the entire length of DYRK1A. The scores are assigned between 0 and 1, and a score above 0.5 indicates disorder. The HRD region is shaded in blue. The longest stretch of IDR is labelled as DYRK1A-IDR and its boundaries are marked. **b**, The indicated C-terminally Strep-tagged GFP fusions were purified from recombinant *E. coli* BL21 cells. Two micrograms each of the fusion proteins was examined by SDS–PAGE followed by Coomassie blue staining. **c**, Solutions containing 5 mg ml^{-1} of the indicated fusion proteins, 37.5 mM NaCl and 10% PEG8000 were trapped between coverslips and examined with a microscope under either fluorescent (488 nm) or normal white light (brightfield).

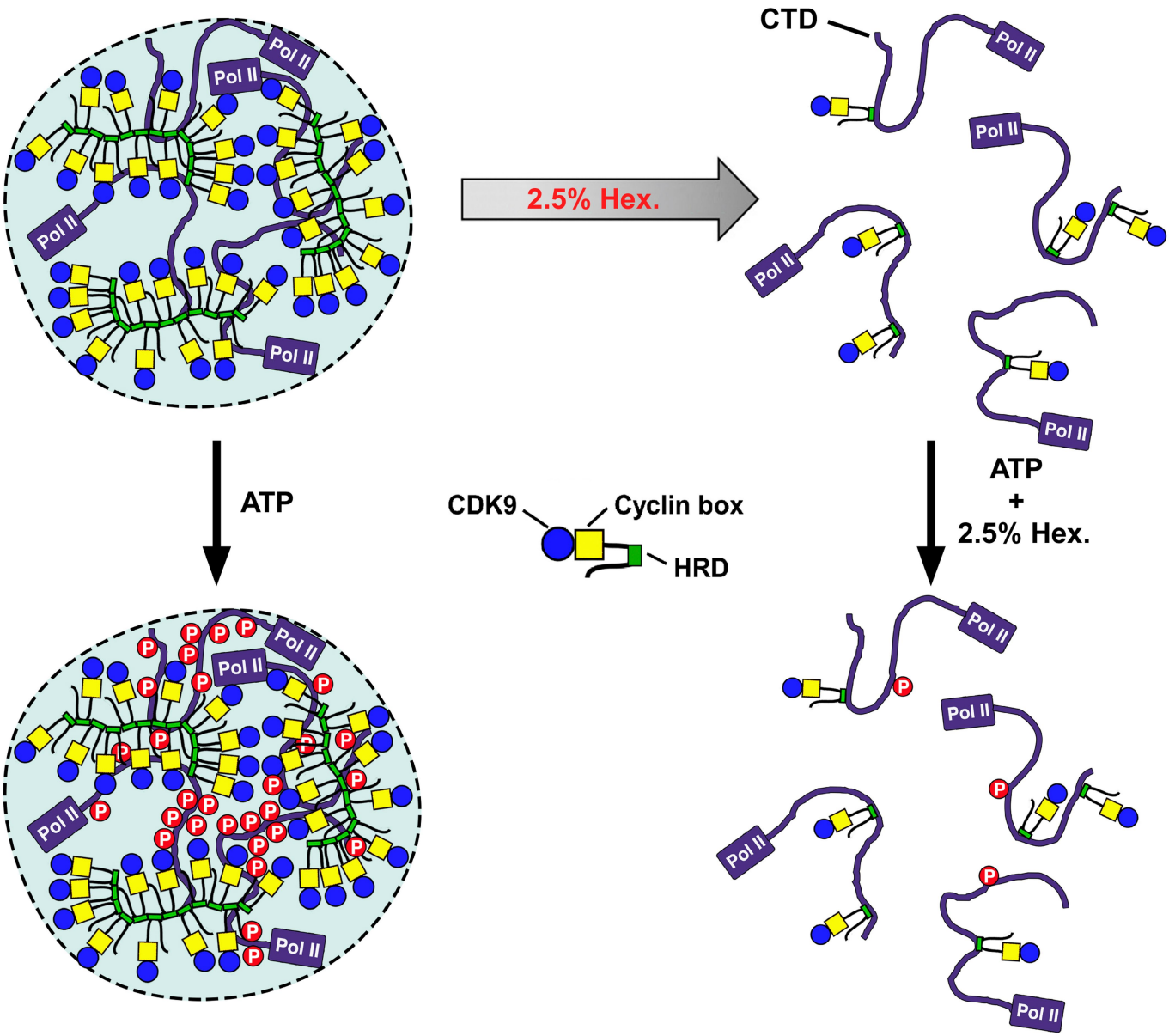


Extended Data Fig. 7 | The central histidine cluster within the CYCT1 HRD is essential for promotion of phase separation by CYCT1 IDR in vitro and in cells, and for P-TEFb to hyperphosphorylate the Pol II CTD and activate HIV transcription. a, The nine histidines in the central histidine cluster within the CYCT1 HRD are highlighted in red and changed to alanines in the 9A mutant. **b**, The C-terminally Strep-tagged GFP–T1-IDR fusions containing either wild-type GFP–T1-IDR or the GFP–T1-IDR(9A) mutant sequence were purified from *E. coli* BL21 cells and examined by SDS–PAGE followed by Coomassie blue staining. **c**, CDK9–CYCT1–Flag heterodimers containing the indicated CYCT1–Flag proteins were affinity-purified from HeLa cells and tested in kinase reactions, with the reaction products analysed as in Fig. 1a. **d**, Plasmids expressing the Gal4 DNA binding domain fused to the indicated CYCT1 proteins were co-transfected into HeLa cell with a HIV-1 LTR–luciferase reporter construct containing the UAS for Gal4. Luciferase activities in cell extracts were measured and analysed as in Fig. 1b. **e**, Fixed and permeabilized HeLa cells expressing wild-type CYCT1–Flag or CYCT1(9A)–Flag were examined by indirect immunofluorescence with the mouse anti-Flag monoclonal antibody and Alexa Fluor 488-conjugated goat anti-mouse secondary antibody. DNA was counterstained using DAPI.



Extended Data Fig. 8 | See next page for caption.

Extended Data Fig. 8 | CYCT1 binds directly to the Pol II CTD in an HRD-dependent manner and the binding is enhanced after the CTD is phosphorylated by CAK (CDK7–CYCH–MAT1). a, Immobilized GST–CTD was incubated with recombinant GFP–T1-IDR or GFP–T1-IDRΔHRD. The input (2.5%) and the bound proteins were analysed by western blotting. **b**, Binding reactions containing the purified recombinant fusion proteins indicated on the right were analysed in a 7.5 to 20% glycerol gradient containing 500 mM NaCl plus 0.5% NP-40, which was centrifuged at 55,000 r.p.m. and 4 °C for 13 h. The indicated fractions were analysed by western blotting to detect the distributions of proteins marked on the left. The entire length of the CTD could be bound by varying numbers of IDRs, resulting in the formation of a series of complexes with broad distributions in the gradient. **c**, mCherry–CTD was incubated with or without immobilized CAK for 6 h in kinase reactions and then analysed by SDS–PAGE and Coomassie blue staining. **d**, Immobilized GST–CTD was incubated with (phos.) or without (unphos.) CAK for 6 h in kinase reactions. After washing, the GST–CTD beads were incubated with GFP–T1-IDR. The indicated proteins were eluted off the beads and analysed by western blotting.



Extended Data Fig. 9 | A model depicting how P-TEFb uses the CYCT1 HRD to target and recruit the Pol II CTD into a phase-separated compartment—which is formed by weak, multivalent homotypic interactions among the HRDs—to enable highly efficient phosphorylation of CTD by P-TEFb in the presence of ATP. At 2.5% 1,6-hexanediol, the HRD-mediated phase separation but not the direct HRD–CTD binding is disrupted, making it impossible for P-TEFb to hyperphosphorylate the CTD.

nature research

Corresponding author(s): Qiang Zhou

Reporting Summary

Nature Research wishes to improve the reproducibility of the work that we publish. This form provides structure for consistency and transparency in reporting. For further information on Nature Research policies, see Authors & Referees and the Editorial Policy Checklist.

Statistical parameters

When statistical analyses are reported, confirm that the following items are present in the relevant location (e.g. figure legend, table legend, main text, or Methods section).

n/a	Confirmed	
☐	☒	The exact sample size (*n*) for each experimental group/condition, given as a discrete number and unit of measurement
☐	☒	An indication of whether measurements were taken from distinct samples or whether the same sample was measured repeatedly
☐	☒	The statistical test(s) used AND whether they are one- or two-sided *Only common tests should be described solely by name; describe more complex techniques in the Methods section.*
☒	☐	A description of all covariates tested
☒	☐	A description of any assumptions or corrections, such as tests of normality and adjustment for multiple comparisons
☐	☒	A full description of the statistics including central tendency (e.g. means) or other basic estimates (e.g. regression coefficient) AND variation (e.g. standard deviation) or associated estimates of uncertainty (e.g. confidence intervals)
☒	☐	For null hypothesis testing, the test statistic (e.g. *F*, *t*, *r*) with confidence intervals, effect sizes, degrees of freedom and *P* value noted *Give P values as exact values whenever suitable.*
☒	☐	For Bayesian analysis, information on the choice of priors and Markov chain Monte Carlo settings
☒	☐	For hierarchical and complex designs, identification of the appropriate level for tests and full reporting of outcomes
☒	☐	Estimates of effect sizes (e.g. Cohen's *d*, Pearson's *r*), indicating how they were calculated
☐	☒	Clearly defined error bars *State explicitly what error bars represent (e.g. SD, SE, CI)*

Our web collection on statistics for biologists may be useful.

Software and code

Policy information about availability of computer code

Data collection	FiJi(ImageJ); Matlab version 2016b; Matlab version of Spot-On (v1.0; GitLab tag 92cdf210) (https://gitlab.com/tjian-darzacq-lab/spot-on-matlab)
Data analysis	FiJi(ImageJ); Matlab version 2016b; Matlab version of Spot-On (v1.0; GitLab tag 92cdf210) (https://gitlab.com/tjian-darzacq-lab/spot-on-matlab)

For manuscripts utilizing custom algorithms or software that are central to the research but not yet described in published literature, software must be made available to editors/reviewers upon request. We strongly encourage code deposition in a community repository (e.g. GitHub). See the Nature Research guidelines for submitting code & software for further information.

Data

Policy information about availability of data

All manuscripts must include a data availability statement. This statement should provide the following information, where applicable:

- Accession codes, unique identifiers, or web links for publicly available datasets
- A list of figures that have associated raw data
- A description of any restrictions on data availability

Uncropped scans for all western blots are provided in Supplementary Figure 1. The raw slowSPT and spaSPT data are freely available in Spot-On readable CSV and

Matlab formats in the form of single-molecule trajectories at Zenodo: https://zenodo.org/record/1215836. The Spot-On Matlab code is available together with a step-by-step guide at Gitlab: https://gitlab.com/tjian-darzacq-lab/spot-on-matlab. All other data are available from the corresponding author on reasonable request.

Field-specific reporting

Please select the best fit for your research. If you are not sure, read the appropriate sections before making your selection.

☒ Life sciences ☐ Behavioural & social sciences ☐ Ecological, evolutionary & environmental sciences

For a reference copy of the document with all sections, see nature.com/authors/policies/ReportingSummary-flat.pdf

Life sciences study design

All studies must disclose on these points even when the disclosure is negative.

Sample size	The number of trajectories used for SPT analysis was described in Fig. 2, and found to give statistically reliable findings. No sample size determination needed for rest.
Data exclusions	No data exclusions occurred in this study.
Replication	All experimental findings have been reliably reproduced at least three times and sometimes even more (e.g. SPT data). Any use of statistical methods have been described in relevant figure legends.
Randomization	This study does not involve randomization of samples/organisms/participants.
Blinding	This study does not require investigators to be blinded to group allocation during data collection and/or analysis.

Reporting for specific materials, systems and methods

Materials & experimental systems

n/a	Involved in the study
☒	☐ Unique biological materials
☐	☒ Antibodies
☐	☒ Eukaryotic cell lines
☒	☐ Palaeontology
☒	☐ Animals and other organisms
☒	☐ Human research participants

Methods

n/a	Involved in the study
☒	☐ ChIP-seq
☒	☐ Flow cytometry
☒	☐ MRI-based neuroimaging

Antibodies

Antibodies used	AFF1: Bethylaboratories, Cat. #A302-344A, validated by WB, 1:1000 dilution AFF4: Abcam, Cat. # ab57077, validated by WB, 1:1000 dilution ELL2: Bethyl Laboratories, 238 Cat. # A302-505A, validated by WB, 1:1000 dilution MePCE: Bethyl Laboratories, Cat. # A304-184A, validated by WB, 1:1000 dilution RNA Pol II phospho-Ser2(3E10): Millipore, Cat. # 04-1571, validated by WB, 1:1000 dilution RNA Pol II phospho-Ser5(3E8): Millipore, Cat. # 04-1572, validated by WB, 1:1000 dilution Total RNA Pol II: Santa Cruz, Cat. # sc-56767, validated by WB, 1:1000 dilution CDK7: Sigma, Cat. # C7089-.2ML, validated by WB, 1:3000 dilution GST: CST, Cat. #2622, validated by WB, 1:1000 dilution Tubulin: EMD CHEMICALS, Cat. #CP06, validated by WB, 1:2000 dilution CycT1: Santa Cruz, Cat. # sc-10750, validated by WB (1:1000 dilution) and IF (1:200 dilution) Anti-Flag: Sigma, Cat. # F3165, validated by WB and IF, 1:500 dilution CDK9, LARP7, HEXIM1 and Brd4 abs: Homemade, validated by WB, 1:1000 dilution
Validation	All commercial antibodies were also validated by the manufactures as indicated on their web sites.

Eukaryotic cell lines

Policy information about cell lines

Cell line source(s)	ATCC
Authentication	UC Berkeley Cell Culture Facility
Mycoplasma contamination	No contamination as verified by UC Berkeley Cell Culture Facility
Commonly misidentified lines (See ICLAC register)	N/A