Multisite phosphorylation code of CDK

Mihkel Õrd, Kaidi Möll, Alissa Agerova, Rait Kivi, Ilona Faustova, Rainis Venta, Ervin Valk and Mart Loog

The quantitative model of cyclin-dependent kinase (CDK) function states that cyclins temporally order cell cycle events at different CDK activity levels, or thresholds. The model lacks a mechanistic explanation, as it is not understood how different thresholds are encoded into substrates. We show that a multisite phosphorylation code governs the phosphorylation of CDK targets and that phosphorylation clusters act as timing tags that trigger specific events at different CDK thresholds. Using phospho-degradable CDK threshold sensors with rationally encoded phosphorylation patterns, we were able to predictably program thresholds over the entire range of the Saccharomyces cerevisiae cell cycle. We defined three levels of CDK multisite phosphorylation encoding: (i) serine–threonine swapping in phosphorylation sites, (ii) patterning of phosphorylation sites, and (iii) cyclin-specific docking combined with modulation of CDK activity. Thus, CDK can signal via hundreds of differentially encoded targets at precise times to provide a temporally ordered phosphorylation pattern required for cell division.

CDKs, the master regulators of cell division, are activated by different cyclins at different cell cycle stages. Despite extensive studies, we do not have a comprehensive answer to the question of how CDKs control the temporal order of cell cycle events. It is not understood how execution timing is encoded into the hundreds of targets to ensure a flawless cell cycle progression. The main obstacle has been combinatorial complexity, because the majority of CDK targets contain intricate patterns of phosphorylation sites, docking sites, and phospho-epitopes.

The quantitative model proposed by Nurse and Stern, and later experimentally supported by a single cyclin-CDK fusion in fission yeast, states that cyclin accumulation in the cell cycle leads to rising CDK activity, reaching different activity thresholds necessary for phosphorylation of specific targets (Fig. 1a). It has been found that the phosphorylation rate of early substrates is higher than that of late substrates and that efficient phosphorylation of some early targets is dependent on cyclin-substrate interactions. However, it is still an unclear how CDK can trigger different events at different times, and we do not understand the role of different cyclins in the context of multisite phosphorylation networks.

Three main interactions drive CDK-dependent phosphorylation (Fig. 1b–d). First, the CDK active site recognizes minimal (S/T) P motif (S/TPxK/R)2,3,4. Second, cyclins can bind to specific substrates via linear motifs for substrate targeting. Third, the Cks1 subunit of the CDK complex interacts with phosphorylated TP sites and directs multisite phosphorylation.4,5

Also, cyclins modulate the intrinsic activity of CDK, with the early G1 and S cyclins conferring lower activity to Cdk1 than the late G2 and M cyclins.4,5 This could function as a mechanism that prevents premature phosphorylation of late targets, while cyclinspecific docking enables weaker G1 and S CDKs to phosphorylate early targets.

Our goal was to determine whether there is a general set of rules, termed the multisite phosphorylation code of CDK, that is used to define CDK thresholds and the execution timing of cell cycle events. Most phosphorylation networks are located in intrinsically disordered regions of CDK targets, potentially enabling linear encoding of sequence motifs along the primary structure of the protein. Previously, we found that the distances (numbers of amino acids) between phosphorylation sites and docking motifs may be crucial for CDK thresholds. The cyclin–CDK–Cks1 complex could serve as a scaffold for the disordered substrate, mediating an ordered phosphorylation process. According to our hypothesis, the linear motifs (phosphorylation and docking sites) and their patterns function as a barcode, giving a unique identity to each substrate. The cyclin–CDK–Cks1 complex can read the barcode and assign the execution of any CDK-triggered switch to a specified time point during the cell cycle.

To gain systematic understanding of threshold encoding, we took an approach that stems from the synthetic biology concept of ‘build to understand’. Using differentially encoded fluorescent threshold sensors in budding yeast, we demonstrate how multisite phosphorylation patterns can be encoded to achieve different CDK thresholds, and thus, different cell cycle timing of specific phosphorylation events. Furthermore, we define a core principle of how the CDK oscillator is separated from the bulk activities of other proline-directed kinases. We show how a CDK activity filter of this type is built using a concept of the ‘helper network’ that only CDK can use.

Results

CDK thresholds can be assigned to proteins by combining linear motifs. We aimed to design substrates that would function as CDK threshold sensors with predictable execution timing over the entire cell cycle. We used non-inhibitory Sic1(1–215), an intrinsically disordered CDK model substrate, as a chassis with phosphorylation-dependent degradation as an output signal for CDK activity. The Sic1 N terminus contains eight CDK phosphorylation sites and cyclin docking motifs (LP motif for G1 cyclins and two RxL motifs for S cyclins) (Fig. 1c). Importantly, the regions between these motifs are highly variable even within the Saccharomyces clade (Supplementary Note 1), suggesting that the regions between the motifs function mainly as a linker sequence to achieve the necessary distances between the motifs. First, we aimed to focus on the elements specific for the Cdk1–Cks1 module and excluded the cyclin docking by mutating the RxL and LP motifs (wild-type WT sensor) (Fig. 1c). Using a live-cell fluorescent microscopy protocol, we followed the timing of phosphorylation-dependent degradation of GFP-tagged sensors (Fig. 1c). The degradation was measured relative to Start, the point when 50% of the G1 transcriptional...
Repressor Whi5 has exited the nucleus (Fig. 1a,b). Ubiquitination of Sic1 is mediated by constitutively active SCF-Cdc424,25; therefore, the degradation profiles are dependent solely on phosphorylation.
Phosphorylation of two diphosphodegrons in Sic1 (sites T45–T48 and S76–S80) are a crucial requirement for Sic1 degradation\(^2\) (Fig. 1c). Importantly, site T48 of the first degron does not bear the CDK consensus motif (discussed below). To confirm the importance of the degrons as outputs, we analyzed strains that express double mutants of the degrons (Fig. 1g). Degradation was severely reduced, barely reaching the half-maximum levels by the end of mitosis, around 60–80 min from Start. These two diphosphodegrons were fixed as the output sites when varying the patterns of motifs in the sensors throughout the study.

Because the distance between the Cks1 priming sites and the degrons is potentially a crucial parameter for degron phosphorylation\(^4\), we first tested the effect of increasing the distance between the priming site T33 and the degron T45–T48 by introducing linkers that were four and eight amino acids long (Fig. 1h). The construct with the wild-type distance of 12 amino acids was degraded at 29 min after Start, whereas the constructs with an extra four and eight amino acids were degraded much later, at 67 and 79 min, respectively (Fig. 1i–k). Next, to entirely disrupt the Cks1-mediated phosphorylation of the degron, we replaced the threonines with serines in positions 2, 5, and 33, because Cks1 only binds phosphothreonines\(^2\)\(^{,20}\) (Fig. 1c,h). Strikingly, the threshold for this sensor was shifted above the CDK peak as we observed less than 50% degradation during mitosis (Fig. 1i). These experiments confirm that both the optimal distances and Cks1 play crucial roles in threshold encoding.

Next, we showed that it is possible to bring the thresholds down by adding back the S-Cdk-specific Rxl docking motifs, thus confirming that cyclin-specific motifs can be crucial for the earlier thresholds (Fig. 1h–k) and agreeing with results from fission yeast that show that cyclin specificity enhances phosphorylation of ~50% of G1/S targets\(^6\). These results highlight an important principle in threshold encoding: the non-optimaiial phosphorylation clusters that encode higher thresholds can be compensated by cyclin docking to bring the thresholds down. The effects of different barcodes on the phosphorylation dynamics of selected sensors were also verified using Phos-tag western blotting (Supplementary Fig. 1c and Supplementary Note 2).

**Binary coding of CDK thresholds by serine—threonine swapping.** There is potential for rational binary encoding of thresholds by swapping threonines and serines in phosphorylation sites without changing the linear structure of the barcode, while providing 2\(^s\) of serine-threonine (S-T) combinations. Different combinations can encode different thresholds via at least two mechanisms. First, the phospho-docking module Cks1 exclusively binds phosphothreonines and not phospho-serines\(^2\)\(^{,20}\); therefore, the N-terminally positioned threonine sites would lower the thresholds. In contrast, when serines occur in N-terminal positions, the priming mechanism would be lost and the thresholds raised. Additionally, by using a kinase assay with peptides that contain CDK consensus motifs, we found that Cdk has about two to three times higher specificity (\(k_\text{cat}/K_\text{m}\)) toward the SP peptide compared with the TP peptide (Fig. 2a). This finding suggests that threonine residues could be phosphorylated later due to differences in CDK specificity. With respect to threshold encoding, it is important that the relative effect of pTP versus pSP with respect to Cks1 docking is much more prominent compared with the active site specificity of S-T, as exemplified using a peptide with a phosphorylated priming site and a secondary site (Fig. 2b).

To explore the possibility of binary coding, we used S-T swapping in threshold sensors with a fixed pattern containing only minimal consensus sites (including T48+1P, which in wild-type Sic1 is a non-consensus site without +1P), and no cyclin docking sites. The degradation of a substrate in which all eight N-terminal phosphorylation sites were serines (8SP) occurred 27 min after Start (Fig. 2c–e and Supplementary Fig. 2a). Next, we predicted that when the C-terminal serines (S69–S76–S80) were replaced with threonines (5SP–3TP), the threshold would be shifted up. Indeed, the degradation was delayed for 12 min, likely because the sites T69–T76–T80 were less phosphorylated, owing to the CDK active site serine preference. Because they are located at the C-terminal end of the cluster, they cannot contribute via Cks1 (Fig. 2c–e).

When the sites toward the N terminus, S45 and S48, were also mutated to threonines (3SP–5TP), a downward shift was observed in the threshold, instead of further upward shifting, leading to degradation only 6 min later than the 8SP sensor (Fig. 2c–e). This result could be predicted if the effect of both Cks1-pTP and CDK active site–SP specificity are considered. The S-T swapping at sites 45 and 48 creates a docking site for Cks1, which would compensate for the loss of specificity of the threonine sites.

Next, we tested a sensor that contains the wild-type Sic1 site combination (5TP–3SP), which, unlike the true wild type, had only minimal consensus sites (Fig. 2c). This construct yielded a threshold that occurred over 20 min earlier than the construct 5SP–3TP with the opposite S-T distribution (Fig. 2c–e). Thus, a potentiating effect from the N-terminal Cks1 docking sites combined with favorable active site specificity of the C-terminal serines is superior to the other combinations analyzed above. All threonine constructs (8TP) and 3TP–5SP are degraded relatively early, showing that the effect of C-terminal threonines is lessened in the presence of Cks1 priming sites (Fig. 2c–e and Supplementary Note 3).

Additionally, we compared the phosphorylation and degradation of 5SP–3TP and 5TP–3SP sensors via western blot, where a similar 20 min delay was detected (Supplementary Fig. 2b). Importantly, no accumulation of fully phosphorylated forms could be detected in either case, indicating that the fully phosphorylated forms are rapidly degraded in both serine- and threonine-based degrons (Supplementary Note 4).

Further, we mutated the degron S76–S80 to T76–S80 so that S80 would have preferential specificity compared with T76. The 6SP–1TP–1SP Mck1 sensor also showed a relatively strong effect from the N-terminal Cks1 site as a primer (Fig. 2i), as demonstrated with Eco1 and Cdc6\(^2\)\(^{,24}\). This could lead to GSK-Mediated phosphorylation of T76. To test this hypothesis, we deleted MCK1 in these strains and found that the degradation of the 6SP–1TP–1SP sensor was delayed by over 20 min (Fig. 2g).

The principle of integrating the Mck1 signal was also demonstrated by a sensor where an optimal specificity determinant +3K was added to S80 (6SP–1TP–1SP +3K). This construct was degraded very early (Fig. 2f,g), close to the wild-type sensor with cyclin docking sites (WT+RXL) and three optimal CDK motifs (Fig. 1j). This sensor also showed a relatively strongly Mck1 deletion effect compared with the 8SP version. On the other hand, the sensor with wild-type phosphorylation sites showed either no or only a mild delay in an mck1 background, indicating that when the 76–80 degron receives strong Cks1-mediated support, Mck1 input will have a relatively minor influence. The sensor 3TP–5SP with intermediate Cks1 support for 76–80 showed intermediate effect of Mck1 deletion.

Phosphatase PP2A\(^2\)\(^{,25}\) has been shown to preferentially dephosphorylate threonines\(^2\)\(^{,26}\) and could additionally contribute to binary encoding (Supplementary Note 5 and Supplementary Fig. 2d–f). Taking these findings together, we have demonstrated how substrates with differences only in S-T distribution can have different phosphorylation dynamics. When the binary encoding is combined with the distance encoding, one would potentially be able to cover the entire span of cell cycle thresholds with high temporal resolution.
Fig. 2 | Binary coding of CDK thresholds using S-T swapping. a, Steady-state kinetic analysis of the phosphorylation of H1-based model peptide substrates PKTPKKAKKL and PKSPKKAKKL using the Clb2−Cdk1−Cks1 complex. b, Similar kinetic analysis as that in a, using a primed peptide substrate based on Sic1 containing pT33−T45, using the Clb2−Cdk1 complex with and without Cks1. The plots in a and b show the mean values ± standard deviation. Data are from four and two replicate experiments, respectively. c, Diagram showing the S-T swapping schemes in the studied threshold sensors. The arrows above the barcode show direct docking-independent phosphorylation by Cdk1, and the arrows below show Cks1-mediated phosphorylation. The width of the arrows indicates the S-T specificity of CDK and pTP-pSP specificity of Cks1-dependent priming steps. d, A plot showing the dynamics of threshold sensors with binary S-T codes, as indicated in c. The graph shows the mean of a population of cells, with error bars ± s.e.m. e, f, Plots showing the 50% degradation timing values for individual cells of the indicated sensors in wild-type or mck1 strain backgrounds. The numbers above the plots show the mean value for each sensor, and the error bars are 95% confidence intervals. g, h, Plots showing the mean nuclear levels of threshold sensors of the strains presented in f. Error bars are ± s.e.m. i, A scheme showing the CDK-primed and Mck1-mediated phosphorylation of a diphosphodegron. Source data for a, b, d–h, including details on sample size are available online.
Cyclin-specific docking motifs in CDK threshold encoding. Our latest studies have revealed that aside from the RxL motifs specific for S-CDK (Rxl14,15,17) (Fig. 1c), the other three major CDK complexes (Cln2−, Clb3−, Clb2−CdK1, denoted as G1, G2, and M complexes) in budding yeast also use specific substrate docking motifs16,17,18. This suggests that the entire span of thresholds can be controlled by cyclin-specific mechanisms. We first set out to test the effect of a Cln2-specific docking motif, LP16,17,18. The degradation of the wild-type sensor was shifted approximately 10 min earlier by the addition of an LP motif (Fig. 3a,b,c). However, the effect of LP docking was less pronounced when the specificity of the T45−T48 degron was augmented by adding a proline next to T48 (T48+1P, creating a CDK consensus site). Also, a fast single diphosphodegron (LTTPPRSP, a modified cyclin E degron16,18) showed less of an enhancing effect from LP docking. The high-affinity phosphorydegron has also been previously shown to lower the threshold for Sic1 degradation19,20. These observations introduce an important principle that poor specificity of output site can be compensated by docking. On the other hand, when output sites have high specificity for CDK, the effect of docking is less pronounced.

We have recently discovered an M-CDK-specific motif, LxF, that can strongly potentiate phosphorylation by Clb2-Cdk135. We introduced this motif into sensors where the thresholds had been shifted up by reducing Cks1-dependent phosphorylation of the degrons to target the time window of M-CDK accumulation (Fig. 3a). Strikingly, this motif shifted the thresholds down by 25−35 min (Fig. 3c,d). Recently, we have also mapped a G2-CDK-specific docking motif PxF (Kövérnáki, Kivi, Puss, Ord, Loog, manuscript in preparation). The motif shifted the threshold of the T33(+4) sensor by approximately 15 min, which is considerably less than that induced by the M-CDK motif (Fig. 3e−g). By following the levels of cyclins with fluorescent tags (Fig. 3h and Supplementary Fig. 3a−c), we found that although Clb3 peaked between S-cyclin Clb5 and M-cyclin Clb23, it showed a much lower nuclear abundance. This result explains why the M-CDK motif was able to shift the threshold more than the G2-CDK motif. Importantly, we also confirmed that cyclin-specific docking of Clbs acted via the hydrophobic patch, because corresponding mutations in the cyclin pockets (hpm) shifted the thresholds back and showed prominent effects in vitro (Fig. 3i) and Supplementary Fig. 4a−q). Therefore, it is possible to encode the CDK thresholds by discriminating even between closely related B-type cyclins using cyclin-specific docking, and the resulting thresholds correlate well with the accumulation timing and abundances of cyclins (cytoplasmic CDK activity in Supplementary Note 6 and Supplementary Fig. 5). Also, the in vitro experiments are consistent with the in vivo degradation profiles, which is important evidence that the in vitro CDK specificity can be used to predict the thresholds in vivo.

Cyclin rearrangements change the order of thresholds. To obtain a full understanding of cyclin specificity in phosphorylation ordering, we replaced the S-cyclin gene CLB5 at its locus with the M-cyclin gene CLB214,15,17 (Fig. 4a). Because the Cdk1 inhibitory kinase Swe1 has higher specificity toward M-Cdk19,20, we also deleted the SWE1 gene to obtain a direct comparison of M- and S-CDK activities. In the resulting clb5::CLB2 strain, a threshold sensor with the M-CDK-specific motif (clb5::CLB2 T33(+4) LxF) showed an impressive 40 min earlier degradation compared to the strain with the wild-type cyclin order expressing a construct without docking sites (Fig. 4b−d). Compared to the wild-type strain expressing the sensor with the docking site (T33(+4) LxF), the shift was over 20 min. This clearly illustrates that expression of cyclins is correlated with the phosphorylation of specific targets, and that cyclin rearrangement will change the order of thresholds due to cyclin-specific phosphorylation.

Furthermore, in the clb5::CLB2 strain, the substrate without cyclin docking sites was degraded approximately 25 min earlier, revealing that M-CDK has higher docking-independent intrinsic activity compared with S-CDK (Fig. 4c,d). This result presents in vivo evidence for the increasing intrinsic activity principle we proposed previously14,17 (Fig. 4e,f). Additionally, the sensor bearing the S-CDK motif RxL did not change its threshold upon cyclin replacement, providing further evidence of this principle, because in vitro data predicted that M-CDK, while having only minor specificity for the RxL, can compensate the loss of docking by its higher intrinsic activity16,17. This also explains the previously published result that Clb2 can compensate the loss of Clb5 in the clb5::CLB2 swe1Δ strain21. Furthermore, specifically targeting M-CDK activity to a late sensor triggers an earlier degradation compared with S-CDK, providing additional evidence for the higher intrinsic activity of M-CDK (Supplementary Fig. 6a,b and Supplementary Note 7).

Increasing CDK activity profile is also important to gain better resolution for higher thresholds. If a uniform CDK activity is accumulating, the fold-change in the activity is considerably reduced at higher cyclin levels compared with the early cell cycle. For example, in a simplified cell cycle model in which four cyclins are expressed sequentially at ten units each (Fig. 4g), a sharp increase in CDK activity occurs in the early stages of cell cycle. However, in the final stages (progression from 0.75 to 1), only a slight increase of 1.25 times occurs (Fig. 4g). Logarithmically increasing intrinsic activity, however, helps to keep the fold-change, and, thus, the threshold resolution, high during the entire cell cycle (Fig. 4h,i). To illustrate the importance of this quality, we modelled the phosphorylation of four substrates with different CDK specificities22. Using the simplified cell cycle model (Fig. 4g), we analyzed the phosphorylation of the substrates with CDK activity increasing according to the uniform or logarithmic model. In the uniform model, the early substrates are phosphorylated in a switch-like manner, but the substrates with lower kinase specificity are phosphorylated to a much lower extent and have no temporal switch-like resolution (Fig. 4h). However, with CDK activity increasing logarithmically, each substrate is phosphorylated rapidly at the expression of subsequent cyclins (Fig. 4i).

Non-consensus CDK sites as specificity filters: separating CDK from the other proline-directed kinases. The C-terminal site of the T45−T48 degron, the S90TTKS, is a non-consensus CDK site, without the necessary +1P19. We hypothesized that a non-proline site could be a filter to prevent other proline-directed kinases from triggering Sic1 degradation. Most of the kinome can be divided into basophilic, acidophilic, and proline-directed kinases with relatively overlapping site specificity40, which poses an unsolved enigma of how signaling specificity is achieved. According to our hypothesis, although the S/TP sites may have overlapping specificity with other kinases, the non-proline site acts as a filter that transmits only the CDK signal, owing to the unique Cks1-mediated phosphorylation.

To test this, we measured the levels of the threshold sensors during pheromone treatment, which leads to activation of MAPK Fus3 and inhibition of Cdk1 activity41,42. Like CDKs, MAPKs are proline-directed kinases. However, Fus3 and CDK have diametrically opposed signaling functions. Whereas CDK activity leads to the cell cycle, Fus3 activity causes cell cycle arrest in G1 and induces the alternative cell fate program, mating. The system of two opposing kinase signaling pathways has been considered as the most basic model system in cellular decision-making between proliferation and differentiation. Paradoxically, however, the opposing decisions are signaled via kinases with overlapping site specificity.

In pheromone-treated cells, Whi5 was localized to the nucleus, which indicates cell cycle arrest in G1 (Fig. 5a). During a 2-h period of the arrest, the levels of wild-type sensor remained stable and showed only a mild decrease. However, a construct with an added proline after T48 (T48+1P) was steadily degraded in
arrested cells, leading to the loss of more than half of the protein during this period (Fig. 5b). This result confirms that when Cdk1 is inhibited, other proline-directed kinases can phosphorylate the S/TP-based diphosphodegrons but are unable to phosphorylate the degron containing a non-proline site. This non-proline site alone would represent a high threshold for any proline-directed MAPK-Fus3 MAPK-Fus3

MAPK-Fus3 MAPK-Fus3

GFP

CDK

20 aa

T33(+8)

T33(+8) LxF

T33S T33S LxF

T33S T33S PxF

T33(+4) +LxF or PxF

16 aa

MAPK-Fus3 MAPK-Fus3

GFP

CDK

12 aa

T5S T33S

T5S T33S PxF

T33(+4) +PxF

T33(+8) +PxF

T5S T33S +PxF

16 aa

MAPK-Fus3 MAPK-Fus3

GFP

CDK

12 aa

T5S T33S

T5S T33S PxF

T33(+4) +PxF

T33(+8) +PxF

T5S T33S +PxF

T33S T33S PxF

WT +LP

WT LP

T48+1P +LP

G1-CDK

WT

T48+1P

LLTPPRSP

LLTPPRSP LP

0

20

40

60

80

Timing (min)

0

5

10

15

20

25

30

35

40

Fluorescence intensity (a.u.)

0

0.2

0.4

0.6

0.8

1

Fluorescence intensity (a.u.)

0

0.2

0.4

0.6

0.8

1

Fluorescence intensity (a.u.)

0

0.2

0.4

0.6

0.8

1

WT LP

T48+1P LP

LLTPPRSP LP

67 42 79 47 94 60

67 51 79 65 94 86

29 19 16 14 12 7.5

Timing (min)

67 42 79 47 94 60

67 51 79 65 94 86

29 19 16 14 12 7.5

Timing (min)

Clb5

Clb4

Clb3

Clb2

Clb1

0

20

40

60

80

Time from Whi5 nuclear exit (min)

0

20

40

60

80

Time from Whi5 nuclear exit (min)

0

20

40

60

80

Time from Whi5 nuclear exit (min)

Fig. 3 | Cyclin-specific encoding of CDK thresholds. a, Diagrams of the threshold sensors used in b–g. (b–d) The dynamics of threshold sensors with cyclin-specific docking motifs for G1-CDK (LP motif), M-CDK (LxF motif), and G2-CDK (PxF motif). The plots show mean nuclear fluorescence intensities of the sensors ± s.e.m. e–g, Plots showing the 50% degradation timing values for individual cells of the strains presented in b–d. The numbers above the plot are the mean values for each strain, the error bars are 95% CIs of the mean. h, Nuclear fluorescence intensities of cyclins fused to Citrine of cells synchronized at the time of 50% of Whi5 nuclear export in late G1. Plot shows the mean values ± s.e.m. i, Autoradiographs showing in vitro phosphorylation of purified threshold sensors with indicated linear docking motif using different wild-type or cyclin-substrate docking mutant CDK complexes. hpm, hydrophobic patch mutant; lpd, LP docking mutant. j, Similar reactions with wild-type CDK complexes were separated using Phos-tag SDS-PAGE to analyze the multisite phosphorylation pattern. Uncropped gel images are shown in Supplementary Data Set 1. Source data for b–h including details on sample size are available online.
kinase, which would even exceed the peak value of CDK activity. Adding different ‘helper networks’ of priming and cyclin docking sites will make it possible to encode any threshold downward, for CDK exclusively.

The filter principle and the number of phosphorylation sites. To demonstrate the importance of the filter principle on the CDK multisite phosphorylation, we designed a substrate construct in which all eight N-terminal CDK sites of Sic1, including the non-proline

Fig. 4 | Cyclin rearrangements change the order of the thresholds. Live-cell microscopy was used to measure the degradation of threshold sensors in swi1Δ and cib5::CLB2 swi1Δ strains. a, A plot showing the mean±s.e.m. intensities of S and M cyclins (as in Fig. 3h). In cib5::CLB2 strain, CLB2 is placed into CLB5 locus and expressed from CLB5 promoter. b, Diagrams of the threshold sensors used in c, d, c. Plot showing mean nuclear levels of the sensors in swi1Δ or cib5::CLB2 swi1Δ strain background. The error bars are ±s.e.m. d, Plots showing degradation timing values for the indicated sensors in single cells. The numbers above the plot show the mean for each sensor, the error bars are 95% CI. e, f, Diagrams showing the principle of changing intrinsic activity of CDK. The phosphorylation specificity of CDK complexes toward a consensus phosphorylation motif rises in the order of expression of cyclins in the cell cycle. ‘Increase in activity’ shows the rise in CDK activity in the period between the dotted lines.

h, i, Steady-state accumulation of maximally phosphorylated substrate with six phosphorylation sites with an ordered distributive phosphorylation mechanism was simulated using the formula y=(k × u)⁴/(1 + k × u+(k × u)² + (k × u)³ + (k × u)⁴ + (k × u)⁵), where k is the rate constant of the kinase divided by the rate constant of phosphatase, and u is the relative concentration of the kinase compared to the phosphatase. In h, the CDK activity follows the uniform model, whereas in i, the CDK activity follows the logarithmic model presented in g. Source data for a, c, d, h, i, including details on sample size, are available online.
site T48, were changed to minimal consensus SP sites (8SP, also used in Fig. 2d). As expected, this substrate was also leaking during pheromone arrest (Fig. 5c). Surprisingly, however, in cycling cells, the 8SP substrate was degraded with the same kinetics as the WT substrate (with a non-CDK output site), despite the lack of Cks1-priming sites (Fig. 6a,b).

These results shape an important mechanistic principle of the CDK control system. According to the initial models, the main emerging property of multisite phosphorylation is the ultrasensitivity of the switch (high Hill coefficient) based on distributive-accumulative multisite phosphorylation\(^1,4,5\). According to this view, a substrate with eight minimal consensus sites and at least a four-site output signal (two diphosphodegrons) should, in theory, create a substrate with eight minimal consensus sites and at least a four-site phosphorylation level providing a mechanistic basis for the CDK quantitative model. Our data clearly show that there can be many distinct thresholds, providing fine-grained control over the order of events from a small number of CDK complexes (Fig. 6c and Supplementary Fig. 6c–e). Most CDK targets contain multiple phosphorylation sites clustered in disordered regions\(^1\). The serine-threonine distribution in phosphorylation sites and the pattern of these sites, combined with positioning of cyclin docking motifs, creates a unique barcode on each substrate (Supplementary Fig. 6g). On the basis of the CDK multisite phosphorylation parameters, different Cks1-mediated and cyclin-dependent connections can be predicted for each target (Supplementary Fig. 6g). As shown here, these connections can determine the phosphorylation dynamics of different targets. Importantly, we have shown that there is a high correlation between the in vitro phosphorylation specificities and in vivo phosphorylation timings in case of the Sic1-based substrates. Additionally, there is significant in vitro evidence of the impact of Cks1 and cyclin docking on phosphorylation of several CDK targets\(^1,4,6,7,13,14,15\), suggesting that the multisite phosphorylation code is relevant for the temporal ordering of CDK phosphorylation in general. However, depending on the substrate and especially the nature of the output phosphorylation sites, the multisite phosphorylation parameters and the linker regions between the motifs can have different importance in determining the phosphorylation timing.

A striking feature of the CDK multisite phosphorylation code is the possibility of binary coding by serine-threonine swapping. At least three different mechanisms, when combined, can encode a wide range of thresholds without changing the overall pattern of the network. The number of serine-threonine combinations increase with the number of sites in the network\(^2\), presenting a simple way to tune cell cycle progression during evolution. In fact, two mechanisms of the binary coding—Cdk1 active site serine preference and threonine specificity of PP2A—have also been observed in metazoans\(^3,4\).

The main principle of encoding the CDK thresholds over the span of the cell cycle is based on the concept of limiting phosphorylation steps that are crucial for output signals. The limiting step in threshold sensors used in this study was the non-consensus site T48 in a diphosphodegron. In principle, such a limiting site can also be poorly accessible owing to structural constraints. However,
the relative importance of non-consensus sites as limiting steps in the CDK targets seems to be wide, as more experimental evidence supports the previously unappreciated fact that a large fraction of CDK signaling could be transmitted via non-consensus sites \(^{13,19}\). Our studies and those of Suzuki et al. indicate that aside from the absence of +1P, the other less-critical CDK consensus site requirements, +3K/R and \(-2P^{31,32}\), are often present in non-consensus motifs, which provides a minimal degree of CDK specificity and distinguishes the sites from random serines and threonines. With the rate-limiting site alone being above the scale of the thresholds, the so-called ‘helper network’ becomes the core concept of CDK multisite phosphorylation encoding. These networks, which consist of Cks1 priming sites and cyclin docking modules, can bring the thresholds to any predicted lower level, down to the early stages of the cell cycle (Fig. 6d).

This study also provides multiple lines of in vivo evidence for the logarithmically changing CDK intrinsic activity principle\(^{13,19,46}\). Importantly, aside from being a safeguard mechanism to prevent premature triggering of later switches, the logarithmic principle is also important to gain better resolution of the switches at higher thresholds (Fig. 4g-i). Additional mechanisms, such as inhibition of CDK or inactivation of phosphatases, also contribute to the modulation of CDK activity during the cell cycle and can improve the resolution of CDK thresholds. While having a lesser role in budding yeast, the Wee1 kinase inhibits the activity of mitotic CDKs in fission yeast prior to mitosis, where the inactivation of Wee1 at mitotic entry creates a sharp increase in CDK activity, necessary for reaching the mitotic thresholds\(^{37-49}\).

In conclusion, although several proteomics screens and targeted studies on particular CDK-controlled switches have yielded a considerable amount of knowledge of CDK-controlled processes, we still lack a mechanistic understanding of how the most crucial element of the CDK system, the temporal ordering of the hundreds of switches, is achieved. There is empirical evidence of CDK thresholds and cyclin specificity; however, it has been difficult to draw a generalized mechanism for threshold encoding, because the targets have different individual structures and downstream functions. In the current study, we took the opposite approach, which stems from the synthetic biology concept of ‘build to understand’. By using a set of threshold sensors, we were able to rationally vary the network parameters and obtain a full set of rules and a core mechanistic understanding of the CDK multisite phosphorylation code. Because CDK is not an ordinary kinase with simple on/off switching but a system operating via multiple thresholds, gaining full control over it as a programmable system would certainly be useful for synthetic biology applications. Even though cells are using the phosphorylation networks ubiquitously, our knowledge has so far been limited for the de novo design of synthetic multisite phosphorylation circuits.

**Online content**

Any methods, additional references, Nature Research reporting summaries, source data, statements of data availability and
Methods

Yeast strains. Yeast strains were MATa haploids of the W303 strain and are described in Supplementary Table 3. Gene deletions, promoter substitutions, and epitope tagging were performed using standard methods based on PCR-based homologous recombination14,21. For C-terminal tagging with yeastCitrine, GFP in pFA6a-GFP-kanMX61 was replaced with yeastCitrine. The transformants were selected for single-copy integration by fluorescence intensity.

Sic1-based threshold sensors. The threshold sensors were based on a non-inhibitory domain of Sic1 (positions 1–215)22,23 and were expressed under the native Sic1 promoter. To restrain the localization of the threshold sensor to either the nucleus or the cytoplasm, an SV40 NLS (PKKKRKV) or a NES sequence (NELALKAGDLINK), respectively, was added to the N terminus of the protein. Additionally, Rl and L motifs in Mut were replaced with the sites LLTPPRSPR. (T2A T5A T33A T45A S69A) and the C-terminal degron in positions 74–82 was introduced at the docking site mutations described above, the phosphorylation site mutations described previously52,53, except lysates were prepared using Mixer Mill MM 400 (Retch). For purification of Cln2 and Cln3, the sequence from Sic1 positions 85–94 and 110–119 was replaced with a PfX motif with sequence PPGPKFPYK (Käswigi, Kivi, Puss, Örd, Loog, manuscript in preparation). To introduce the LxF motifs, the sequence from Sic1 positions 88–94 and 112–118 was substituted with the sequence PEKLYFO.

In threshold sensors containing RlX motifs, two RlX motifs of Sic1, RRLF and RRLFIRE, were retained as in wild-type protein. In targets with LP motifs, the Lp motif in Sic1 VLLIP was left unmodified, and an additional LP motif from Ste20 (sequence NNVYSLDPQFPQRSTVSSY) was introduced to positions 146–162. The PfX and LxF motifs were introduced to the same positions as RlX motifs. In case of PfX motifs, the sequence from Sic1 positions 85–94 and 110–119 was replaced with a PfX motif with sequence PPGPKFPYK (Käswigi, Kivi, Puss, Örd, Loog, manuscript in preparation). To introduce the LxF motifs, the sequence from Sic1 positions 88–94 and 112–118 was substituted with the sequence PEKLYFO.

Protein purification. Clb5–, Clb3–, and Clb2–Cdk1 complexes were purified from yeast cells with the TAP method using C-terminally tagged cyclins as described previously22, except lysates were prepared using Mixer Mill MM 400 (Retch). For purification of Cln2–Cdk1 complexes, an N-terminally tagged 3HA-Cln2 fusion protein, was overexpressed in yeast cells and purified by immobilized metal affinity chromatography with antibody against the HA epitope as described22 (rabbit polyclonal anti-HA was purchased from Labas, Estonia). Cks1 was purified as described previously22.

Sic1-based threshold sensors were purified using N-terminal 6His-tag. 6His-tagged proteins were expressed in Escherichia coli BL21RP cells at 37°C using 1 mM IPTG to induce expression. His-tagged proteins were purified using standard cobalt affinity chromatography with 200 mM imidazole for elution.

Kinase assays. The general composition of the assay mixture contained 50 mM HEPES, pH 7.4, 180 mM NaCl, 5 mM MgCl2, 20 mM imidazole, 2% glycerol, 0.2 mg/ml BSA, 500 mM Cs1, and 500 μM ATP ((with added [γ-32P]-ATP) (Hartmann Analytische). Substrate protein concentrations were 1 μM (in the linear S versus v0 range, several fold below the estimated Km value). The concentrations of kinase complexes were around 0.2 nM. The kinase assays were performed under conditions below 10% of initial substrate turnover. Reactions were carried out at room temperature and were stopped at two time points (7 and 14 min) using SDS-PAGE sample buffer. For separation of differentially phosphorylated forms of the substrate proteins, we used 10% SDS-PAGE supplemented with 100 μM Phos-tag (Wako Chemical Industries). Electrophoresis of Phos-tag SDS-PAGE gels was carried out at 15 mA for 3 h. For histone H1 peptide kinetics, the peptides with sequences PTKPPKAKKL and PKSPKPAAK (Inbio) were phosphorylated by Clb2–Cdk1 at eight peptide concentrations ranging from 1 μM to 3,000 μM. H1 peptides were visualized by high-performance capillary electrophoresis, and phosphorylation intensities were quantified. The biotinylated Sic1pT33-T45 peptide (biotin-MQGQK(p) TPKQPQSNLVPVPTSTKSFK) (Storko) was used at six concentrations ranging from 1 μM to 300 μM, and the phosphorylation reactions were pipetted onto SAM2 Biotin Capture Membrane (Promega Corporation). γ32P phosphorylation signals were detected using an Amersham Typhoon 5 Bioimaging System (Imaging Sciences). Signals were quantified using ImageQuant TL (Amersham Biosciences), and GraphPad Prism 5.0 was used for data analysis. All kinase assays were performed in at least two replicate experiments. The number of replicates and the quantified signals for each replicate are available in the source data files online.

Time-lapse fluorescence microscopy. Yeast cultures were grown at 30°C in synthetic complete media lacking uracil with 2% glucose (SC-URA) to OD 0.2–0.6. Cells were pipetted onto 0.8 mm cover glass slips and covered with a 1-mm thick 1.5% agarose pad made with SC. Before the start of the experiment, cells were held on the slide under the agarose pad for 60 min. Imaging was performed at 30°C using a Zeiss Observer Z1 microscope with a 63×/1.4NA oil immersion objective and an AxioCam 506 mono camera (Zeiss), using 3×3 binning for imaging. The temperature of the agarose pad was held at 30°C using a Tempcontrol37–2 digital from PeCon. Images were taken every 3 min, and imaging sessions were 8h long. Up to 12 positions were imaged in one experiment using an automated stage and ZEN software. Definite Focus was used to keep the cells in focus during the experiment. Strains expressing threshold sensor—EGFP fusion proteins were exposed for 15 ms using a Colibri 470 LED module. Imaging of cyclins fused with yeastCitrine was performed using a Colibri 505 LED module for 500 ms. Whis-mCherry was imaged using a Colibri 540–580 LED module for 750 ms. All Colibri modules were used at 25% power. Image segmentation, cell tracking, and quantification of fluorescence signals was performed using MATLAB (The MathWorks, Inc.) as described in ref. 22. For every threshold sensor, we analyzed data from at least two repeats with different transformants. Only daughter cells were used for analysis, and at least 60 cells from each strain were analyzed. The exact number of cells used in the analysis of each strain can be found in Supplementary Table 3 and from the source data files available online.

Western blotting. The threshold sensors with either nuclear localization signal or nuclear export signal and a C-terminal 3HA tag were expressed from a pRS315 vector under the ADHI promoter. Cells were grown in synthetic complete media lacking uracine (SC LEU), with 2% glucose at 30°C to OD 0.3. Cells were then treated for 2.5 h with 1 μg/ml α-factor, washed thoroughly, and released into fresh SC LEU medium. Five milliliters of cells was centrifuged every 10 min and frozen in liquid nitrogen. Cells were lysed by bead beating in lysis buffer containing urea. Proteins were separated using Phos-tag SDS-PAGE with 50 μM Phos-tag and 8% acrylamide. Blotting of Phos-tag SDS-PAGE gels was performed using a dry system iBlot (Invitrogen). Purified anti-HA 11 epitope tag antibody (1:500) (clone 16B12, Biolegend Cat. No. 901501) and HRP-conjugated anti-mouse antibody (1:7,500) from Labas, Estonia were used to detect the 3HA-tagged proteins.

Bioinformatics. The homologs of Sic1 from budding yeasts were found using the ProViz tool22. Multiple sequence alignment and the positional conservation scoring was performed using lalign22.

SLiMsearch459 was used to search for potential non-S/TP phosphorylation sites from the disordered regions of the yeast proteome. The disorder score (IUPRED) cutoff was set at 0.3. We searched for two motifs, based on different non-S/TP consensus motifs: TP[10,30][ST][P][KR][KR] and TP[8,28][P][ST] [+P][0,1][KR].

Reporting Summary. Further information on research design is available in the Nature Research Reporting Summary linked to this article.

Data availability

Source data for Fig. 1g,i–k, 2a,b,d–h, 3b–h, 4a,c,d,h,i, 5a–c, 6b,c and Supplementary Figs. 2a–e, 3b,c, 4a–q, 5a–c, 6b,c and Supplementary Tables 1–9 are available in the source data files online. All other data are available upon request.

References

50. Longtime, M. S. et al. Additional modules for versatile and economical PCR-based gene deletion and modification in Saccharomyces cerevisiae. Yeast 14, 953–961 (1998).
51. Janke, C. et al. A versatile toolbox for PCR-based tagging of yeast genes: new fluorescent proteins, more markers and promoter substitution cassettes. Yeast 21, 947–962 (2004).
52. Puig, O. et al. The tandem affinity purification (TAP) Method: A general procedure of protein complex purification. Methods 24, 218–229 (2001).
53. Ubersax, J. A. et al. Targets of the cyclin-dependent kinase Cdk1. Mol. Cell. Biol. 218–229 (2001).
54. McCusker, D. et al. Cdk1 coordinates cell-surface growth with the cell cycle. Mol. Cell. Biol. 14, 953–961 (1998).
55. McCusker, D. et al. Cdk1 coordinates cell-surface growth with the cell cycle. Nat. Cell Biol. 9, 506–515 (2007).
56. Reed, G. J., Reynolds, D. W., Verma, R. & Deshaies, R. J. Cdk1 is required for G1 cyclin-cyclin-dependent kinase activity in budding yeast. Mol. Cell. Biol. 20, 5858–5864 (2000).
56. Doncic, A., Eser, U., Atay, O. & Skotheim, J. M. An algorithm to automate yeast segmentation and tracking. *PLoS One* **8**, e57970 (2013).

57. Jehl, P., Manguy, J., Shields, D. C., Higgins, D. G. & Davey, N. E. ProViz—a web-based visualization tool to investigate the functional and evolutionary features of protein sequences. *Nucleic Acids Res.* **44**, W11–W15 (2016).

58. Waterhouse, A. M., Procter, J. B., Martin, D. M. A., Clamp, M. & Barton, G. J. Jalview Version 2—a multiple sequence alignment editor and analysis workbench. *Bioinformatics* **25**, 1189–1191 (2009).

59. Krystkowiak, I. & Davey, N. E. SLiMSearch: a framework for proteome-wide discovery and annotation of functional modules in intrinsically disordered regions. *Nucleic Acids Res.* **45**, W464–W469 (2017).
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.

Statistics

For all statistical analyses, confirm that the following items are present in the figure legend, table legend, main text, or Methods section.

- □ Confirmed
- □ The exact sample size (n) for each experimental group/condition, given as a discrete number and unit of measurement
- □ A statement on 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 statistical parameters 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

Our web collection on statistics for biologists contains articles on many of the points above.

Software and code

Policy information about availability of computer code

Data collection Microscopy data was collected using ZEISS software (Zeiss), kinase assays were quantified using ImageQuant TL (Amersham Biosciences).

Data analysis Microscopy data was analysed using custom-made scripts for image segmentation, cell tracking and quantification in MATLAB (Mathworks). The scripts are available upon request.

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. 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

All data will be made available upon request from corresponding author Mart Loog (mart.loog@ut.ee).

Field-specific reporting

Please select the one below that is 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
Life sciences study design

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

| Sample size | No sample-size pre-calculations were made. In microscopy experiments, 50 cells were considered to be sufficient for the analysis, because replicate experiments with such sample size showed very little variation. |
|-------------|---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| Data exclusions | No data was excluded from the analysis.                                                                                                           |
| Replication | All data presented is from at least two replicate experiments, all replication events were successful and are included in the data.                                                                 |
| Randomization | Samples were only grouped by yeast strains, no further allocations were made.                                                                                                                             |
| Blinding | Investigators were not blinded to the group allocation, however, automated analysis was used whenever possible.                                                                                           |

Reporting for specific materials, systems and methods

We require information from authors about some types of materials, experimental systems and methods used in many studies. Here, indicate whether each material, system or method listed is relevant to your study. If you are not sure if a list item applies to your research, read the appropriate section before selecting a response.

| Materials & experimental systems | Methods |
|----------------------------------|---------|
| n/a                              | n/a     |
| □ ☒ Antibodies                  | □ ☒ ChIP-seq |
| □ ☒ Eukaryotic cell lines       | □ ☒ Flow cytometry |
| □ ☒ Palaeontology               | □ ☒ MRI-based neuroimaging |
| □ ☒ Animals and other organisms |         |
| □ ☒ Human research participants |         |
| □ ☒ Clinical data               |         |

Antibodies

| Antibodies used | Purified anti-HA.11 epitope tag antibody (clone 16812, Biolegend) and HRP-conjugated anti-mouse antibody from Labas, Estonia were used to detect the 3HA-tagged proteins. |
| Validation      | Each lot of anti-HA.11 epitope tag antibody is quality control tested by Western blotting by the manufacturer (Biolegend). |