A CDK-mediated phosphorylation switch of disordered protein condensation

Juan Valverde, Geronimo Dubra, Henk W.P. van Den Toorn, Guido Van Mierlo, Michiel Vermeulen, Albert Heck, Carlos Elena-Real, Aurélie Fournet, Emile Al Ghoul, Dhanvantri Chahar, et al.

A CDK-mediated phosphorylation switch of disordered protein condensation

Maarten Altelaar (m.altelaar@uu.nl)
Utrecht University https://orcid.org/0000-0001-5093-5945

Juan Valverde
Utrecht University

Geronimo Dubra
CNRS

Henk W.P. Van den Toorn
Utrecht University https://orcid.org/0000-0002-0270-5763

Guido van Mierlo
EPFL https://orcid.org/0000-0001-5883-0339

Michiel Vermeulen
Radboud University Nijmegen https://orcid.org/0000-0003-0836-6894

Albert Heck
Utrecht University https://orcid.org/0000-0002-2405-4404

Carlos Elena-Real
CBS, University of Montpellier

Aurélie Fournet
CBS, University of Montpellier

Emile Al Ghoul
IGH, University of Montpellier

Dhanvantri Chahar
IGMM, University of Montpellier

Austin Haider
University of Denver

Matteo Paloni
CBS, University of Montpellier https://orcid.org/0000-0003-4841-9321

Angelos Constantinou
Institute of Human Genetics, UMR9002 CNRS-UM, 141 rue de la Cardonille, 34396 Montpellier, France.

Alessandro Barducci
Centre de Biochimie Structurale

Kingshuk Ghosh
University of Denver

Nathalie Sibille

CBS, University of Montpellier
Pau Bernadó
CBS
Puck Knipscheer
https://orcid.org/0000-0003-4198-0132
Liliana Krasinska
CNRS https://orcid.org/0000-0002-6858-0852
Daniel Fisher
French National Centre for Scientific Research https://orcid.org/0000-0002-0822-3482

Biological Sciences - Article

Keywords:

Posted Date: February 24th, 2022

DOI: https://doi.org/10.21203/rs.3.rs-1370895/v1

A CDK-mediated phosphorylation switch of disordered protein condensation

Authors: Juan Manuel Valverde[1,2†], Geronimo Dubra[3,4†], Henk van den Toorn[1,2], Guido van Mierlo[5], Michiel Vermeulen[5], Albert J.R. Heck[1,2], Carlos Elena-Real[6], Aurélie Fournet[6], Emile Al Ghoul[7], Dhanvantri Chahar[3,4], Austin Haider[8], Matteo Paloni[6], Angelos Constantinou[7], Alessandro Barducci[6], Kingshuk Ghosh[8], Nathalie Sibille[6], Pau Bernado[6], Puck Knipscheer[9], Liliana Krasinska[3,4‡], Daniel Fisher[3,4‡*], Maarten Altelaar[1,2‡*]

Affiliations:

[1]Biomolecular Mass Spectrometry and Proteomics, Bijvoet Center for Biomolecular Research and Utrecht Institute for Pharmaceutical Sciences, University of Utrecht, Utrecht, 3584 CH Utrecht, Netherlands.

[2]Netherlands Proteomics Center, Padualaan 8, 3584 CH Utrecht, Netherlands.

[3]IGMM, University of Montpellier, CNRS, Inserm, Montpellier, France.

[4]Equipe Labellisée LIGUE 2018, Ligue Nationale Contre le Cancer, Paris, France.

[5]Department of Molecular Biology, Faculty of Science, Radboud Institute for Molecular Life Sciences, Oncode Institute, Radboud University Nijmegen, 6525 GA Nijmegen, the Netherlands.

[6]CBS, University of Montpellier, INSERM, CNRS, Montpellier, France.

[7]IGH, University of Montpellier, CNRS, Montpellier, France.

[8]Department of Physics and Astronomy, and Department of Molecular and Cellular Biophysics, University of Denver, Denver, Colorado 80208, USA.

[9]Oncode Institute, Hubrecht Institute–KNAW and University Medical Center, Utrecht, 3584 CT, Netherlands.

*Correspondence to: m.altelaar@uu.nl and daniel.fisher@igmm.cnrs.fr

†‡ Equal contributions

Cell cycle transitions arise from collective changes in protein phosphorylation states triggered by cyclin-dependent kinases (CDKs), but conceptual and mechanistic explanations for the abrupt cellular reorganisation that occurs upon mitotic entry are lacking. Specific interactions between distinct CDK-cyclin complexes and sequence motifs encoded in substrates might result in highly ordered phosphorylation[1], while bistability in the mitotic CDK1 control network can trigger switch-like phosphorylation[2]. Yet the dynamics of mitotic phosphorylation has not been demonstrated *in vivo*, and the roles of most cell cycle-regulated phosphorylations are unclear. Here, we show evidence that switch-like phosphorylation of intrinsically disordered proteins (IDPs) by CDKs contributes to mitotic cellular reorganisation by controlling protein-protein interactions and phase separation. We studied protein phosphorylation in single *Xenopus* embryos throughout synchronous cell cycles, performed parallel assignment of cell cycle phases using egg extracts, and analysed dynamics of mitotic phosphorylation using quantitative targeted phosphoproteomics. This provided a high-resolution map of dynamic phosphosites from the egg to the 16-cell embryo and showed that mitotic phosphorylation occurs on entire protein complexes involved in diverse subcellular processes and is switch-like *in vivo*. Most cell cycle-regulated phosphosites occurred in CDK consensus motifs and located to intrinsically disordered regions. We found that substrates of CDKs and other cell cycle kinases are significantly more disordered than phosphoproteins in general, a principle conserved from yeast to humans, while around half are components of membraneless organelles (MLOs), whose assembly is thought to involve phase separation. Analytical modelling predicts modulation of homotypic IDP interactions by CDK-mediated phosphorylation, which was confirmed by biophysical and biochemical analysis of a model IDP, Ki-67. These results highlight the dynamic control of intrinsic disorder as a conserved hallmark of the cell cycle and suggest a mechanism for CDK-mediated mitotic cellular reorganisation.

Main

To explain behaviour of complex systems, such as the cell cycle, two general strategies have been used[3]: top-down identification of system components, such as screens which have identified hundreds of CDK substrates[4–9] and cell cycle-regulated proteins[10], and bottom-up molecular analysis of the structural effects of individual phosphorylations on single proteins[11]. Yet it has proven challenging to use studies performed at such different scales to reconcile different models of CDK-mediated phosphorylation. We reasoned that understanding how

thousands of mitotic phosphorylations[12] bring about an ordered cell cycle transition would require a multidisciplinary quantitative approach involving cell biology, biochemistry, bioinformatics and biophysics. A *sine qua non* is a time-resolved map of *in vivo* cell cycle phosphorylation in a system devoid of artifacts arising from cell synchronisation[13,14], and with temporal resolution that alternative approaches, like centrifugal elutriation[15] or FACS[16] cannot provide. Dynamic phosphorylation states cannot be determined from cell populations[17], while single-cell proteomics studies[18,19] currently have insufficient sensitivity and reproducibility for low stoichiometry and dynamic targets.

A high-resolution map of *in vivo* cell cycle phosphorylation

We took advantage of the naturally synchronous early cell cycles of *Xenopus laevis* embryos[20,21] to perform quantitative phosphoproteomics *in vivo*, using a sensitive phosphopeptide enrichment strategy[22]. We collected single embryos at 15-minute intervals while recording visual cues of cell divisions. Phosphopeptides from each embryo were purified, separated by nano-LC and analysed by mass spectrometry (Fig. 1a). We identified 4583 high-confidence phosphosites mapping to 1843 proteins (Extended data Fig. 1a; Data S1), most being phosphoserines (Extended data Fig. 1b). Individual embryo phosphorylation states strongly correlated (Extended data Fig. 1c). We thus generated a dynamic map of protein phosphorylation from an unfertilised egg to a 16-cell embryo.

We focused on 1032 sites whose variation in phosphorylation over time was statistically significant (hereafter denoted "dynamic phosphosites") which occurred on 646 proteins. Gene ontology (GO) and network analysis revealed high functional association and interconnectivity between groups of proteins involved in RNA binding and the nuclear pore complex (NPC), DNA replication and chromatin remodeling, and microtubule regulation (Fig. 1b). Hierarchical clustering uncovered four distinct groups that reflect cell cycle-regulated behaviour (Fig. 1c; Data S1). The levels of clusters A and B phosphosites were highest in eggs and post-fertilisation, and decreased during the first round of DNA replication, suggesting that dephosphorylation of these sites may prepare the zygote for upcoming cell divisions[23]. GO analysis for group A highlighted proteins involved in RNA regulation and nuclear organisation, including the NPC and nuclear transport, chromosomal structure and segregation (Extended data Fig. 1d), as also observed in a recent study on meiosis exit[24]. Cluster B phosphosites were enriched in regulators of RNA biosynthesis and stability, translation, actin, DNA replication and repair (Extended data Fig. 1d). Cluster C phosphosites progressively increased after meiotic exit, while cluster D phosphosites had a clear oscillating signature with upregulation

preceding each cell division. GO analysis of cluster C shows dominance of interphase cell cycle processes including DNA replication, RNA-related processes and chromosome organisation (Extended data Fig. 1d), and included phosphosites displaying a reciprocal oscillating trend and a lower amplitude compared to cluster D sites. Several such sites, *e.g.* S31 of the replication licensing protein MCM4, were from monophosphorylated peptides, while the multiphosphorylated forms were found in cluster D (Extended data Fig. 1e). Thus, cluster C contains the earliest phosphorylations of proteins that are highly phosphorylated at mitosis. Cluster D shows coordinated phosphorylation of multiple members of protein complexes involved in diverse processes, suggesting a common mechanism of regulation (Extended data Fig. 1f). Importantly, phosphoproteome changes were not simply a reflection of changes in abundance of the corresponding proteins (Extended data Fig. 2), which are generally negligible during *Xenopus* early development[25].

We assigned *in vivo* embryo phosphosites to different cell cycle stages by comparing with phosphorylation patterns of replicating or mitotic egg extracts (Fig. 1d). Replication was initiated by adding purified sperm chromatin to interphase egg extracts and quantified over time (Fig. 1e, top), while mitosis was triggered by adding recombinant cyclin B and verified microscopically. We also used egg extracts arrested at meiotic metaphase II (Cytostatic Factor, CSF-arrested). Overall, we identified 6937 phosphosites, which included 71% of the sites identified *in vivo* (Fig. 1f, Data S1). 1728 sites varied between S and M-phase, including 693 sites upregulated in S-phase and 1035 in mitosis (Fig. 1e, Data S1). GO analysis of interphase and mitotic sites revealed processes enriched in *in vivo* cluster C and cluster D, respectively (Extended data Fig. 3a). Several DNA-replication factors, including MCM4 and RIF1, showed multi-site phosphorylation specifically in S-phase (Extended data Fig. 3b). This phosphoproteomics dataset greatly increases the known repertoire of phosphorylation sites upregulated during S-phase[12].

We next analysed the cell cycle behaviour of dynamic phosphosites that we found *in vivo* (Extended data Fig. 3c). Most embryo cluster A sites were upregulated in both CSF-arrested meiotic extracts and mitotic extracts, highlighting the global similarities of regulation of meiotic and mitotic M-phase, despite the additional activity of the Mos/MEK/MAP kinase pathway in meiosis. Around half of embryo cluster B sites were present only in interphase, while the rest showed a minimum phosphorylation in late S-phase, confirming their dephosphorylation during the first round of DNA replication. As expected, most sites from embryo clusters C and D were part of the *in vitro* S-phase and mitotic groups, respectively.

Therefore, single embryo data can successfully identify cell cycle-dependent phosphorylation. In mitosis, as expected, monophosphorylated species are reduced because multisite phosphorylation emerges (Extended data Fig. 3d; Extended data Fig. 1e).

Predominance of CDK targets

Analysis of kinase consensus motifs showed that proline-directed (S/T-P) sites, which conform to the minimal consensus for CDKs, comprise 51% of all detected phosphosites *in vivo* and 60% of dynamic sites (Extended data Fig. 4a). Around 10% of all phosphosites matched the full CDK1-family consensus site: S/TPxK/R. Replicating and mitotic extracts displayed a similar trend (Extended data Fig. 4a). Putative CDK targets dominated all clusters, with 80% of sites in cluster D *in vivo* and mitotic clusters *in vitro* conforming to the minimal CDK motif (Fig.1g, Extended data Fig. 4b, c). Consensus sites of other kinases such as Aurora, Polo-like kinase (PLK), DBF4-dependent kinase (DDK) and Casein kinase I and II were present to a lesser extent (Extended data Fig. 4b, d). In meiotic M-phase, MAP kinases, which have the same consensus motif as CDKs, are likely responsible for sites specific to embryo cluster A or CSF extracts, but these kinases are inactivated during early embryonic cell cycles[26], suggesting that most of the other dynamic proline-directed phosphorylations are due to CDKs.

Although few direct CDK substrates have been characterised in *Xenopus*, they are likely conserved between vertebrates. We therefore manually curated a set of 654 human CDK1-subfamily targets (Data S2; see Supplementary Methods for sources). 303 of these have *Xenopus* homologues among the 1843 phosphoproteins we detected, and 149 were present among the 646 proteins with dynamic phosphosites in *Xenopus* embryos (Fig. 1h). Thus, the predominance of CDK motifs among dynamic phosphosites reflects a high proportion of *bona fide* CDK substrates. This is a conservative estimate, since we only considered proline-directed sites as CDK motifs, although we found that 10-20% of human and yeast CDK substrates (Data S2; see Supplementary Methods for sources) were non-proline-directed (Extended data Fig. 4e), confirming a recent finding[33]. These data reinforce the dominant role of CDKs in cell cycle-regulated phosphorylation.

Mitotic phosphorylation is switch-like *in vivo*

We next determined whether mitotic phosphorylation of individual phosphosites is progressive or switch-like *in vivo*. We analysed dynamics of 64 cluster D sites from diverse protein complexes in single embryos every 180-seconds using quantitative targeted phosphoproteomics[27–29] by parallel reaction monitoring[30], thereby obtaining a quantitative

description of mitotic phosphorylation *in vivo* at extremely high-time resolution (Fig. 2a). This revealed parallel and abrupt upregulation of all phosphosites preceding each cell division (Fig. 2b, c), indicating switch-like phosphorylation of diverse protein complexes at mitotic onset. This was not due to oscillation of CDK1-Y15 inhibitory phosphorylation, which was downregulated over time (Fig. 2d), as previously reported[31], consistent with lack of corresponding phosphorylation of the CDK1-Y15-regulatory enzymes, CDC25 and WEE1. In contrast, oscillating phosphorylations on NIPA and the APC/C, which regulate mitotic cyclin accumulation, as well as Greatwall kinase, which activates the PP2A inhibitors Arpp19/ENSA, were apparent (Extended data Fig. 5a). These data suggest that control of mitotic cyclin levels and PP2A activity, and therefore the overall CDK/phosphatase activity ratio[2], suffices for switch-like mitotic phosphorylation whereas regulated CDK1-Y15 phosphorylation is not essential (Extended data Fig. 5b). This is consistent with the self-sufficiency of futile cycles of opposing enzymes in generating switch-like network output in the absence of allosteric regulation[32].

The cell cycle phosphoproteome is intrinsically disordered

We wondered whether the diverse dynamic phosphoproteins share common structural features facilitating switch-like CDK-mediated phosphorylation. Phosphosites in general are often located in intrinsically disordered regions (IDRs) of proteins[34], which is also true for yeast and mouse CDK sites[35,8,9]. Yet previous analyses did not exclude the possibility that this is an artefact due to the enrichment of serine, threonine and proline in disordered regions, which is consistently predicted across the entire proteome of *Xenopus*, human and yeast (Extended data Fig. 6a). We corrected for this compositional bias, and found that phosphorylatable residues in IDR are indeed more highly phosphorylated than those in ordered regions (Fig. 3a-c). This enrichment was increased for proteins with at least one site displaying dynamic phosphorylation; the same was true for human CDK substrates (Fig. 3b, c). To estimate the differential phosphorylation of disordered sites globally, we calculated the ratio of dynamically phosphorylated (*Xenopus*) or CDK-phosphorylated (yeast, human) to non-phosphorylated serine and threonine in both disordered and structured regions (Extended data Fig. 6b; see Methods). This confirmed that cell cycle-regulated phosphorylation is largely skewed towards disordered regions and that CDKs preferentially phosphorylate disordered sites (Fig. 3d, Extended data Fig. 6c). We then asked whether this is also true for substrates of other protein kinases. We analysed the mitotic PLK and Aurora kinases, DYRK kinases, which promote mitotic phosphorylation of several IDPs[36], NEK kinases, which have roles in centrosome

duplication and various stages of mitosis, and MAP kinases, which share the proline-directed S/T consensus site. For each kinase, documented phosphosites were strongly enriched in IDRs (Extended data Fig. 6c, d), supporting the idea that phosphorylation of residues in IDRs is kinetically favoured[34].

To explain the dominance of CDK-mediated phosphorylation in the cell cycle, we surmised that their substrates might be more disordered than phosphoproteins in general. We therefore determined the percentage of disordered residues of proteins in our datasets, compared to the rest of their respective phosphoproteomes (Data S3). This revealed that, on average, both *Xenopus* dynamic phosphoproteins and human and yeast CDK substrates contain approximately twice the proportion of disordered amino acids as other phosphoproteins (Fig. 3e, Extended data Fig. 6e), putting them among the top quartile of proteins with the most disorder in the proteome. If this reflects the importance of disordered proteins for the cell cycle generally, then substrates of other cell cycle kinases might also be more disordered than other phosphoproteins. Indeed, targets of most cell cycle kinases are significantly more disordered than targets of MAP kinase (Fig. 3f), whose phosphosites are also proline-directed and preferentially located in IDRs (Extended data Fig. 6d).

Enrichment of MLO components among CDK substrates

We thus reasoned that phosphorylation may have been selected to regulate the functions of IDPs during the cell cycle. IDPs are key components of membrane-less organelles (MLO), many of which (*e.g.* Cajal bodies, nucleoli, nuclear pore complexes, splicing speckles) are thought to arise by phase separation (PS)[37], are disassembled in mitosis, and can be regulated by phosphorylation[36,38,39]. To corroborate our hypothesis, we analysed available data on cellular localisation for each of our curated human CDK substrates. We found that 257 (39.2%) are present in MLOs, including key IDPs such as coilin (Cajal bodies), nucleophosmin, nucleolin and Ki-67 (nucleoli), 53BP1 (53BP1 bodies), nucleoporins (NPC) and PML (PML bodies) (Fig. 3g). We then manually curated an MLO proteome from human proteomics studies (Data S4; See Supplementary Methods for sources). Homologues of 204 dynamic *Xenopus* phosphoproteins (31.6%) localise to MLOs, as do 73 of the 149 proteins (50%) that show dynamic phosphorylation in *Xenopus* and are CDK substrates in human (Fig. 3g). The vast majority of proline-directed phosphosites and confirmed CDK sites in these proteins were located in predicted IDRs (Extended data Fig. 7).

CDKs regulate IDR phase separation

Both stochastic and specific interactions between IDPs contribute to PS and MLO assembly[37,40,41]. We hypothesised that cell cycle kinase-mediated phosphorylation might modulate such interactions. We first applied a machine learning classifier[42] to predict whether cell cycle-regulated phosphoproteins have an increase in average propensity for PS (PSAP score). Indeed, we observed a sharp increase in the PSAP score, from the proteome to the phosphoproteome, and a further increase for dynamic phosphoproteins, with the highest score for mitotic cluster D (Extended data Fig. 8a). Similarly, the propensity for PS is far higher amongst targets of most cell cycle kinases (CDK, Aurora, PLK, but not NEK) and DYRK kinases than the overall phosphoproteome, but less so for MAP kinase substrates.

Next, to better understand the biochemical effects of their cell cycle-regulated phosphorylation, we analysed a selection of IDRs from CDK substrates. We applied a general heteropolymer theory that uses sequence charge decoration matrices (SCDM), based on electrostatic interactions only, to identify intra-chain interaction topology[43,44]. Since this should correlate with inter-chain interactions that promote PS, SCDMs provide indirect insights to propensity to phase separate. Of the 12 IDPs tested, 7 (nucleolin, nucleophosmin, NUP53, ELYS, MCM4, 53BP1 and the splicing factor SF3B1) had SCDM maps showing visibly decreased self-association propensity (increased red regions in Extended data Fig. 8b), implying reduced propensity to phase separate, upon CDK-site phosphorylation. Conversely, for SRRM2, CDK-mediated phosphorylation is predicted to increase intra-chain attraction (Extended data Fig. 8b) and hence PS tendency. For 4 proteins (MDC1, TICRR, COILIN, and CDT1), SCDM maps were inconclusive. To further analyse these trends, we calculated radius of gyration of several IDRs using all-atom simulation. Effects of phosphorylation on CDT1 (28.4Å to 30.3Å), TICRR (56.2Å to 57.3Å) and coilin (39 Å to 37.9 Å) were minor, while MCM4 IDR expands upon phosphorylation (21.9Å to 26.3Å), consistent with SCDM analysis. Overall, these data suggest that phosphorylation is a key regulator of homotypic interactions, an important element of PS propensity, of most IDRs.

To test this hypothesis, we focused on a model CDK substrate, Ki-67, an IDP that organises heterochromatin structure[45] and perichromosomal layer formation from nucleolar components in mitosis[46,47]. Ki-67 contains a multivalent Ki-67 repeat domain that is highly phosphorylated in mitosis by CDKs (Fig. 4a), which regulates its perichromosomal localisation[48]. SCDM analysis predicted that phosphorylation of full-length Ki-67 should promote self-interaction and thus PS, but this cannot be attributed to interactions of its repeat motif alone, since phosphorylation of the latter is predicted to reduce homotypic interactions (Fig. 4b). In

agreement, coarse-grained (CG) molecular dynamics (MD) simulations (Extended data Movies 1 and 2) showed that the radius of gyration of full-length Ki-67 decreased upon phosphorylation (Fig. 4c, left) while that of a single consensus repeat motif increased (Fig. 4c, right). MD simulations also showed that PS is enhanced by increasing repeat valency and counteracted by phosphorylation (Fig. 4d), consistent with SCDM analysis. To test these predictions experimentally, we first used the optogenetic Cry2 "optodroplet" system[49] with full length Ki-67 or a series of deletion mutants. Full-length Ki-67 localised to the nucleolus, as expected, but exposure to blue light caused rapid appearance of small round foci in the nucleoplasm, which was dependent on the level of induced Ki-67 expression, consistent with PS (Extended data Fig. 9a). Importantly, promoting CDK-mediated phosphorylation by inhibiting PP2A with okadaic acid[2] led to foci formation in the absence of blue light, while pan-CDK inhibition with purvalanol A prevented induction of foci upon light (Fig. 4e, f). These results indicate that, as predicted by SCDM and MD, phosphorylation of full-length Ki-67 promotes PS. Results were similar for constructs lacking the C-terminal LR domain, that binds chromatin, or the N-terminal domain, which is required for the nucleolar localisation of Ki-67 (Extended data Fig. 9b). Finally, we purified a consensus repeat polypeptide (Extended data Fig. 10a) and phosphorylated it *in vitro* with recombinant CDK complexes. Nuclear Magnetic Resonance spectroscopy showed a reduced amide proton spectral dispersion typical for an IDP, and confirmed appearance of 7 phosphorylated residues upon incubation with purified CDKs and ATP (Fig. 4g). We mapped phosphorylation sites and intensity by phosphoproteomics and Phos-Tag-SDS-PAGE, indicating stoichiometric phosphorylation (Extended data Fig. 10b, c). Purified GFP-tagged Ki-67 repeat motif could phase-separate *in vitro*, and, as predicted, this was abolished upon full phosphorylation by CDK (Fig. 4h). Taken together, these results confirm that CDK-mediated phosphorylation is able to both promote or inhibit homotypic interactions that contribute to PS, and suggest that Ki-67 may have several competing modes of PS that are differentially regulated by phosphorylation. Our data suggest a mechanism for Ki-67-mediated mitotic targeting of nucleolar components to the perichromosomal layer[45,46] via CDK-mediated phosphorylation, which reduces PS of several major nucleolar IDPs, thus triggering nucleolar disassembly, while simultaneously promoting PS of Ki-67 bound to chromatin to recruit nucleolar components.

In conclusion, this work reveals *in vivo* that CDK-dependent mitotic phosphorylation occurs in a switch-like manner on diverse proteins whose common denominators are a high level of disorder and localisation to MLOs. Furthermore, our data show that CDK-mediated

phosphorylation regulates homotypic interactions between IDPs, which may coordinate diverse cellular processes during the cell cycle. While this is not incompatible with models in which high-affinity interactions contribute to MLO formation by PS[50,51], it suggests that cell cycle control may be less specific than previously thought.

Acknowledgments: We thank Merlijn Witte for technical assistance with the *Xenopus laevis* egg fertilization experiments, Ariane Abrieu for a gift of CSF egg extracts, and Markus Raschle from the Technical University of Kaiserslautern for providing the *Xenopus laevis* protein database. **Funding:** AJRH and MA acknowledge support from the Horizon 2020 program INFRAIA project Epic-XS (Project 823839) and the NWO funded Netherlands Proteomics Centre through the National Road Map for Large-scale Infrastructures program X-Omics (Project 184.034.019) of the Netherlands Proteomics Centre. JMV is supported by scholarships from the Ministry of Science and Technology of Costa Rica (MICITT) and the University of Costa Rica (UCR). PK and MV are funded by the Oncode Institute which is financed by the Dutch Cancer Society and by the gravitation program CancerGenomiCs.nl from the Netherlands Organisation for Scientific Research (NWO). DF and LK are Inserm employees. GD is funded by the Institut National de Cancer, France (INCa) PRT-K programme (PRT-K17 n° 2018-023). The Fisher lab is funded by the Ligue Nationale Contre le Cancer, France (EL2018.LNCC/DF) and INCa (PLBIO18-094). The CBS is a member of France-BioImaging (FBI) and the French Infrastructure for Integrated Structural Biology (FRISBI), supported by the French National Research Agency (ANR-10-INBS-04-01 and ANR-10-INBS-05).

Author contributions: MA and DF conceived and supervised the project. JMV, PK and LK designed and interpreted experiments. JMV, HT, AH, GvM, LK and GD performed experiments and interpreted the data. MV supervised GvM. JMV, LK, GD, DF and MA wrote the paper.

Competing interests: Authors declare no competing interests.

Data and materials availability: All data is available in the main text or the supplementary materials. All data, code, and materials are available on request.

Supplementary Materials:

Materials and Methods

Extended data figures 1-10

Data S1-S4

Movies S1, S2

References

1. Örd, M. *et al.* Multisite phosphorylation code of CDK. *Nat Struct Mol Biol* **26**, 649–658 (2019).
2. Krasinska, L. *et al.* Protein Phosphatase 2A Controls the Order and Dynamics of Cell-Cycle Transitions. *Mol Cell* **44**, 437–50 (2011).
3. Hyman, A. A. Whither systems biology. *Philos Trans R Soc Lond B Biol Sci* **366**, 3635–3637 (2011).
4. Errico, A., Deshmukh, K., Tanaka, Y., Pozniakovsky, A. & Hunt, T. Identification of substrates for cyclin dependent kinases. *Adv. Enzyme Regul.* **50**, 375–399 (2010).
5. Ubersax, J. A. *et al.* Targets of the cyclin-dependent kinase Cdk1. *Nature* **425**, 859–64 (2003).
6. Chi, Y. *et al.* Identification of CDK2 substrates in human cell lysates. *Genome Biol* **9**, R149 (2008).
7. Blethrow, J. D., Glavy, J. S., Morgan, D. O. & Shokat, K. M. Covalent capture of kinase-specific phosphopeptides reveals Cdk1-cyclin B substrates. *Proc Natl Acad Sci U S A* **105**, 1442–7 (2008).
8. Holt, L. J. *et al.* Global analysis of Cdk1 substrate phosphorylation sites provides insights into evolution. *Science* **325**, 1682–6 (2009).
9. Michowski, W. *et al.* Cdk1 Controls Global Epigenetic Landscape in Embryonic Stem Cells. *Mol Cell* **78**, 459-476.e13 (2020).
10. Mahdessian, D. *et al.* Spatiotemporal dissection of the cell cycle with single-cell proteogenomics. *Nature* **590**, 649–654 (2021).
11. Orlicky, S., Tang, X., Willems, A., Tyers, M. & Sicheri, F. Structural basis for phosphodependent substrate selection and orientation by the SCFCdc4 ubiquitin ligase. *Cell* **112**, 243–256 (2003).
12. Olsen, J. V. *et al.* Quantitative phosphoproteomics reveals widespread full phosphorylation site occupancy during mitosis. *Sci Signal* **3**, ra3 (2010).
13. Cooper, S. The synchronization manifesto: a critique of whole-culture synchronization. *FEBS J* **286**, 4650–4656 (2019).
14. Ly, T., Endo, A. & Lamond, A. I. Proteomic analysis of the response to cell cycle arrests in human myeloid leukemia cells. *eLife* **4**, e04534 (2015).
15. Ly, T. *et al.* A proteomic chronology of gene expression through the cell cycle in human myeloid leukemia cells. *Elife* **3**, e01630 (2014).
16. Ly, T. *et al.* Proteomic analysis of cell cycle progression in asynchronous cultures, including mitotic subphases, using PRIMMUS. *eLife* **6**, e27574 (2017).

17. Purvis, J. E. & Lahav, G. Encoding and decoding cellular information through signaling dynamics. *Cell* **152**, 945–956 (2013).
18. Budnik, B., Levy, E., Harmange, G. & Slavov, N. SCoPE-MS: mass spectrometry of single mammalian cells quantifies proteome heterogeneity during cell differentiation. *Genome Biol.* **19**, 161 (2018).
19. Lombard-Banek, C., Moody, S. A., Manzini, M. C. & Nemes, P. Microsampling Capillary Electrophoresis Mass Spectrometry Enables Single-Cell Proteomics in Complex Tissues: Developing Cell Clones in Live Xenopus laevis and Zebrafish Embryos. *Anal Chem* **91**, 4797–4805 (2019).
20. Newport, J. & Kirschner, M. A major developmental transition in early Xenopus embryos: I. characterization and timing of cellular changes at the midblastula stage. *Cell* **30**, 675–686 (1982).
21. Newport, J. W. & Kirschner, M. W. Regulation of the cell cycle during early Xenopus development. *Cell* **37**, 731–42 (1984).
22. Post, H. *et al.* Robust, Sensitive, and Automated Phosphopeptide Enrichment Optimized for Low Sample Amounts Applied to Primary Hippocampal Neurons. *J. Proteome Res.* **16**, 728–737 (2017).
23. Clift, D. & Schuh, M. Restarting life: fertilization and the transition from meiosis to mitosis. *Nat Rev Mol Cell Biol* **14**, 549–562 (2013).
24. Presler, M. *et al.* Proteomics of phosphorylation and protein dynamics during fertilization and meiotic exit in the Xenopus egg. *Proc. Natl. Acad. Sci. U.S.A.* **114**, E10838–E10847 (2017).
25. Peuchen, E. H. *et al.* Phosphorylation Dynamics Dominate the Regulated Proteome during Early Xenopus Development. *Sci Rep* **7**, 15647 (2017).
26. Ferrell, J. E., Jr., Wu, M., Gerhart, J. C. & Martin, G. S. Cell cycle tyrosine phosphorylation of p34cdc2 and a microtubule-associated protein kinase homolog in Xenopus oocytes and eggs. *Mol Cell Biol* **11**, 1965–71 (1991).
27. Lawrence, R. T., Searle, B. C., Llovet, A. & Villén, J. Plug-and-play analysis of the human phosphoproteome by targeted high-resolution mass spectrometry. *Nat Methods* **13**, 431–434 (2016).
28. Schmidlin, T. *et al.* Assessment of SRM, MRM3, and DIA for the targeted analysis of phosphorylation dynamics in non-small cell lung cancer. *PROTEOMICS* **16**, 2193–2205 (2016).
29. Schmidlin, T. *et al.* High-Throughput Assessment of Kinome-wide Activation States. *Cell Systems* **9**, 366-374.e5 (2019).

30. Peterson, A. C., Russell, J. D., Bailey, D. J., Westphall, M. S. & Coon, J. J. Parallel Reaction Monitoring for High Resolution and High Mass Accuracy Quantitative, Targeted Proteomics *. *Molecular & Cellular Proteomics* **11**, 1475–1488 (2012).
31. Tsai, T. Y.-C., Theriot, J. A. & Jr, J. E. F. Changes in Oscillatory Dynamics in the Cell Cycle of Early Xenopus laevis Embryos. *PLOS Biology* **12**, e1001788 (2014).
32. Goldbeter, A. & Koshland, D. E., Jr. An amplified sensitivity arising from covalent modification in biological systems. *Proc Natl Acad Sci U S A* **78**, 6840–4 (1981).
33. Suzuki, K. *et al.* Identification of non-Ser/Thr-Pro consensus motifs for Cdk1 and their roles in mitotic regulation of C2H2 zinc finger proteins and Ect2. *Scientific Reports* **5**, 7929 (2015).
34. Iakoucheva, L. M. *et al.* The importance of intrinsic disorder for protein phosphorylation. *Nucleic Acids Res.* **32**, 1037–1049 (2004).
35. Moses, A. M., Hériché, J.-K. & Durbin, R. Clustering of phosphorylation site recognition motifs can be exploited to predict the targets of cyclin-dependent kinase. *Genome Biol.* **8**, R23 (2007).
36. Rai, A. K., Chen, J.-X., Selbach, M. & Pelkmans, L. Kinase-controlled phase transition of membraneless organelles in mitosis. *Nature* **559**, 211–216 (2018).
37. Shimobayashi, S. F., Ronceray, P., Sanders, D. W., Haataja, M. P. & Brangwynne, C. P. Nucleation landscape of biomolecular condensates. *Nature* **599**, 503–506 (2021).
38. Berchtold, D., Battich, N. & Pelkmans, L. A Systems-Level Study Reveals Regulators of Membrane-less Organelles in Human Cells. *Mol. Cell* **72**, 1035-1049.e5 (2018).
39. Hur, W. *et al.* CDK-Regulated Phase Separation Seeded by Histone Genes Ensures Precise Growth and Function of Histone Locus Bodies. *Dev Cell* **54**, 379-394.e6 (2020).
40. Lin, Y.-H., Wu, H., Jia, B., Zhang, M. & Chan, H. S. Assembly of model postsynaptic densities involves interactions auxiliary to stoichiometric binding. *Biophys J* **121**, 157–171 (2022).
41. Ghosh, K. Stoichiometric versus stochastic interaction in models of liquid-liquid phase separation. *Biophys J* **121**, 4–6 (2022).
42. van Mierlo, G. *et al.* Predicting protein condensate formation using machine learning. *Cell Rep* **34**, 108705 (2021).
43. Huihui, J. & Ghosh, K. An analytical theory to describe sequence-specific inter-residue distance profiles for polyampholytes and intrinsically disordered proteins. *J Chem Phys* **152**, 161102 (2020).

44. Huihui, J. & Ghosh, K. Intrachain interaction topology can identify functionally similar intrinsically disordered proteins. *Biophys J* **120**, 1860–1868 (2021).
45. Sobecki, M. *et al.* The cell proliferation antigen Ki-67 organises heterochromatin. *Elife* **5**, e13722 (2016).
46. Booth, D. G. *et al.* Ki-67 is a PP1-interacting protein that organises the mitotic chromosome periphery. *Elife* **3**, e01641 (2014).
47. Hayashi, Y., Kato, K. & Kimura, K. The hierarchical structure of the perichromosomal layer comprises Ki67, ribosomal RNAs, and nucleolar proteins. *Biochemical and Biophysical Research Communications* **493**, 1043–1049 (2017).
48. Hégarat, N. *et al.* Cyclin A triggers Mitosis either via the Greatwall kinase pathway or Cyclin B. *The EMBO Journal* **39**, e104419 (2020).
49. Shin, Y. *et al.* Spatiotemporal Control of Intracellular Phase Transitions Using Light-Activated optoDroplets. *Cell* **168**, 159-171.e14 (2017).
50. Feng, Z., Jia, B. & Zhang, M. Liquid–Liquid Phase Separation in Biology: Specific Stoichiometric Molecular Interactions vs Promiscuous Interactions Mediated by Disordered Sequences. *Biochemistry* **60**, 2397–2406 (2021).
51. Musacchio, A. On the role of phase separation in the biogenesis of membraneless compartments. *EMBO J* e109952 (2022) doi:10.15252/embj.2021109952.

Figures

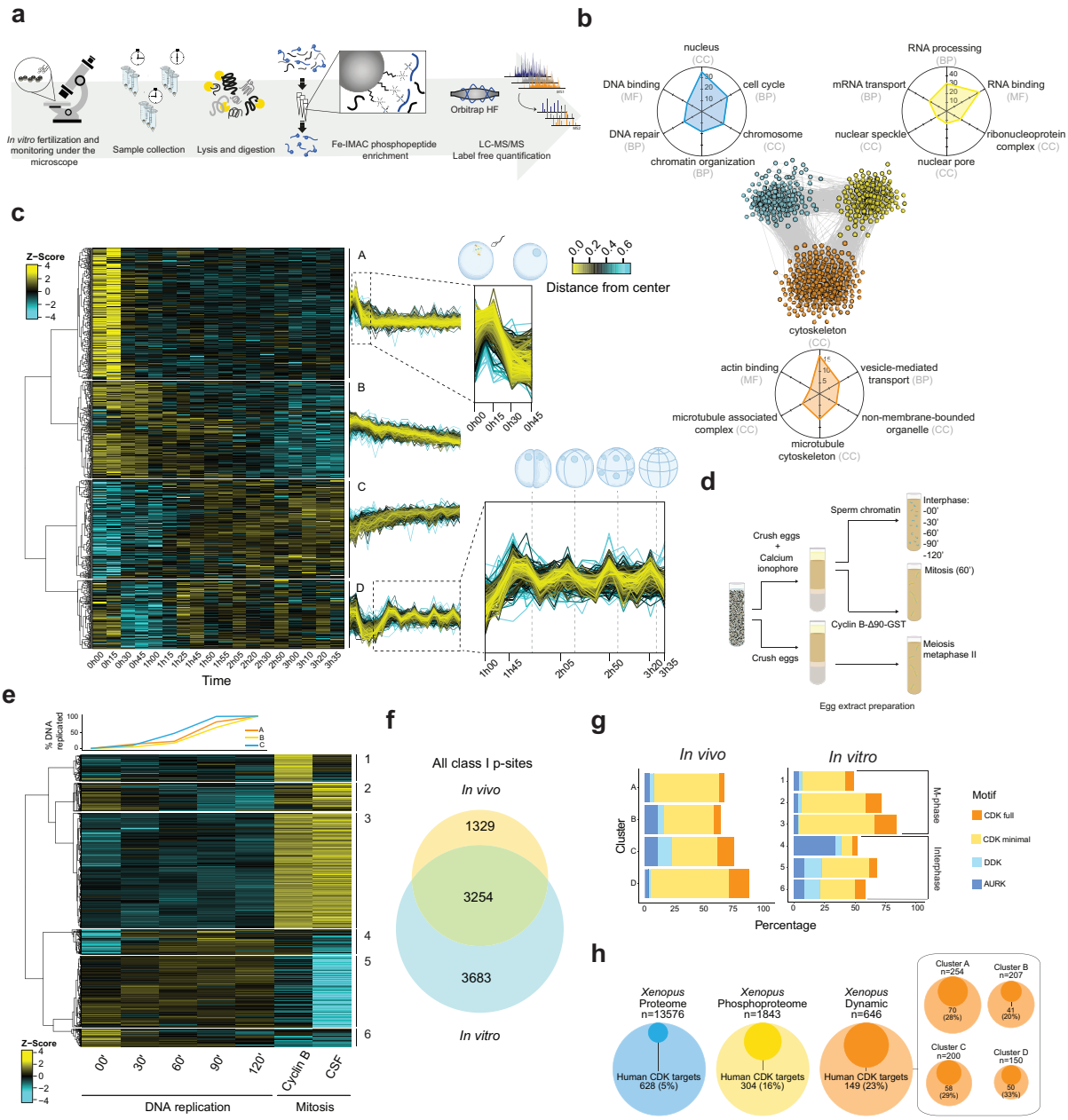


Figure 1. The time-resolved phosphoproteome from a single-cell to a 16-cell embryo and its cell cycle assignment. (a) Schematic representation of the workflow. Single *Xenopus* eggs and embryos were collected followed by cell lysis, protein digestion, phosphopeptide enrichment and high-resolution proteomics analysis. (b) STRING network of functionally associated proteins undergoing dynamic phosphorylation (each node represents a protein). Vicinity clustering reveals three main groups (yellow, blue and orange) with a high degree of association. Radar plots show the corresponding GO terms (adjusted p value <0.05) for each group (axes show -Log_{10}(adj p value) for each GO term). (c) Hierarchical clustering of

significantly changing phosphosites (ANOVA, Benjamini-Hochberg correction, FDR 0.05), reveals 4 clusters with distinct regulation (A-D). Dashed boxes in clusters A and D are zoomed-in to highlight dynamic phosphorylation patterns (dashed lines depict the time points of cell division). (d) Scheme of the experiment in the *Xenopus* egg extract. (e) Top: quantification of DNA replication in each biological replicate. Below: Hierarchical clustering of dynamic phosphosites (ANOVA, Benjamini-Hochberg correction, FDR 0.05) reveals differential regulation of phosphosites during S-phase and mitosis. (f) Overlap between *in vivo* (embryo) and *in vitro* (egg extract) phosphoproteomics. (g) Proportion of phosphosites according to their potential upstream kinase for each cluster in the *in vivo* (top) and *in vitro* (bottom) experiments. (h) Circle plots presenting enrichment of homologues of human CDK substrates among *Xenopus* phosphoproteins detected *in vivo* and those with dynamic phosphosites.

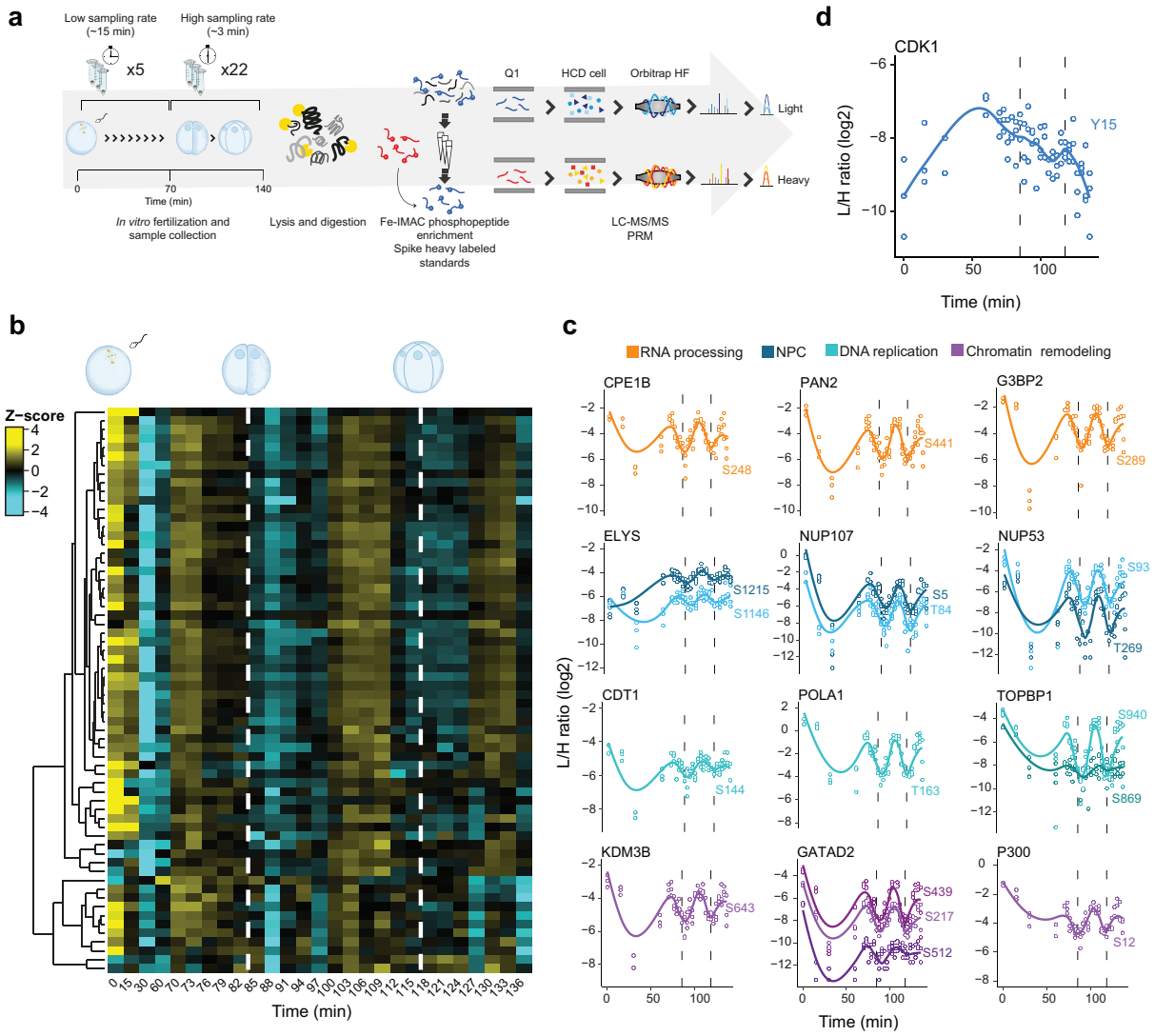
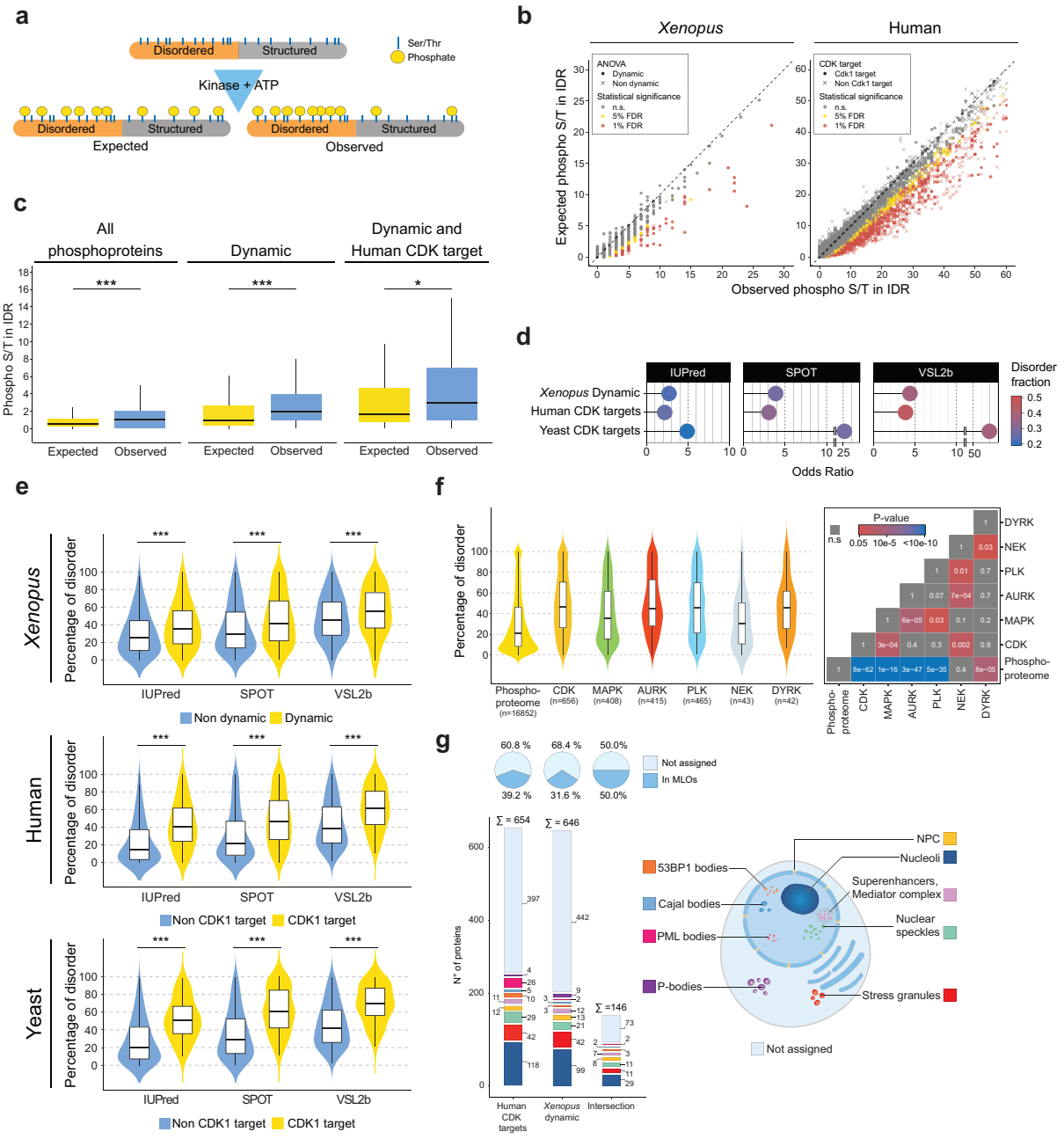


Figure 2. Switch-like mitotic phosphorylation *in vivo*. (a) Schematic representation of the workflow. Samples were collected over two cell divisions and enriched phosphopeptides were subjected to targeted proteomics analysis. (b) Heat map shows a highly synchronous wave of phosphorylation preceding each of the two cell divisions. Dashed lines depict times when cell divisions were recorded. (c) Single phosphosite plots from selected proteins. Each dot represents a biological replicate (n=3). Dashed lines depict times when cell divisions were recorded. (d) Single phosphosite plot of CDK1 inhibitory phosphorylation (Y15).



459

460 **Fig. 3. The cell cycle phosphoproteome is characterised by intrinsic disorder and MLO**
461 **components.** (a) Scheme illustrating hypothetical enrichment of phosphorylation in disordered
462 regions when taking into account amino acid compositional bias. (b) Scatter plot of expected
463 vs observed phosphorylated Ser/Thr for each protein of human and *Xenopus* phosphoprotein
464 datasets. FDR thresholds of 5% and 1% are marked in yellow and red respectively. Circles:
465 proteins with at least one dynamic phosphorylation in *Xenopus*, or human CDK1 subfamily
466 substrates, respectively. (c) Boxplots showing expected vs observed phosphorylated Ser/Thr
467 among all phosphoproteins detected (left), phosphoproteins with at least one dynamic
468 phosphosite (middle), and dynamic phosphoproteins also detected as CDK1 subfamily targets
469 in humans (right). Distributions were compared with the Wilcoxon signed-rank test. *p<0.05,

**p<0.01, ***p<0.001. (d) Plots showing the common Odds Ratio of Ser/Thr phosphorylation in structured and ordered regions calculated with the Fisher's test (see Extended data Fig. 6b, c). For all organisms, the disordered regions were calculated with three different disorder predictors. The disordered fraction is presented in a colour scale. (e) Violin plots of the distribution of disordered residues per protein for CDK targets vs the rest of the phosphoproteome for human and yeast, and dynamic phosphoproteins vs the rest of the phosphoproteome for *Xenopus*. Intrinsic disorder was calculated with three different predictors (IUPred, SPOT, and VSL2b). Statistical significance was evaluated with the Wilcoxon–Mann–Whitney test; ***p<0.001. (f) Violin plot (left) showing the distribution of disordered residues per protein for CDK, MAPK, Aurora, PLK, NEK and DYRK kinase targets vs the rest of the phosphoproteome for human targets. Statistical significance was assessed by Kruskal-Wallis ANOVA, and pairwise comparisons were performed with Dunn's post-hoc tests. The adjusted p-values (Benjamini-Hochberg) are shown in a tile plot (right). (g) Human CDK1 subfamily targets, *Xenopus* dynamic phosphoproteins, and the intersection of both sets, that are present in our manually curated proteome of membraneless organelles.

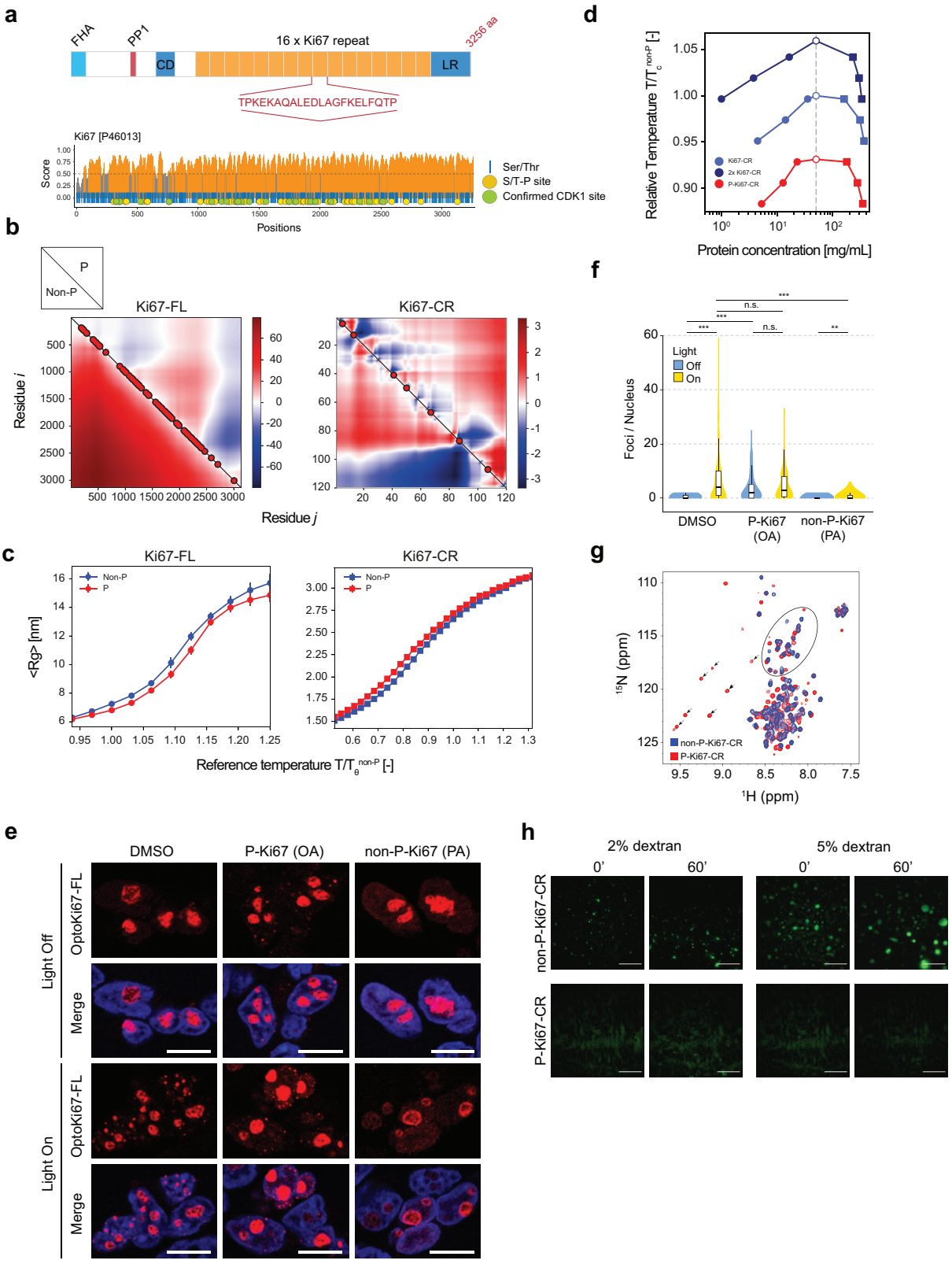


Fig. 4. CDK-mediated phosphorylation regulates phase separation of a model IDP. (a) Top, scheme of the human Ki-67 protein (FHA, forkhead-associated domain; PP1, PP1 phosphatase-binding domain; CD, conserved domain; LR, leucine arginine-rich domain).

Highlighted, Ki-67 repeat consensus motif. Bottom, diagram of IUPred score over the length of human Ki-67. Regions with scores >0.5 (orange) are considered to be disordered, and <0.5 (grey) structured. Blue vertical lines indicate Ser and Thr residues; yellow circles, known Ser/Thr-Pro phosphosites; green circles, confirmed CDK1 subfamily phosphorylations. (b) Sequence Charge Decoration Matrix (SCDM) maps for full length Ki-67 (FL, left) and Ki-67 consensus repeat (CR, right), depicting the contribution of electrostatic interaction dictating the distance between two amino acid residues i and j (shown in x and y axes). The values of SCDM for different residue pairs (i,j) are shown using colour schemes with red and blue denoting positive (repulsive) and negative (attractive) values, respectively. The lower and upper triangles indicate SCDM map for the unphosphorylated (non-P) and phosphorylated (P) sequences, respectively. Confirmed and putative (Ser/Thr-Pro) CDK phosphorylation sites are indicated with red circles. (c) Dependency of the radius of gyration (Rg) on the simulation temperature in single-chain MD simulations for full chain Ki-67 (left) and consensus repeat (right). The reference temperature is the θ temperature of the non-phosphorylated molecule for full chain and consensus repeat, respectively. Reported error bars are obtained by block analysis over 10 blocks. (d) Binodal curves from phase coexistence simulations of the Ki-67 consensus repeat sequence. For each temperature, filled circles indicate the dilute phase density and squares indicate the coexisting dense phase density. Empty circles indicate the fitted critical temperature (T_c) of each system. The T_c of the non-phosphorylated monomer (light blue empty circle) was the reference for the normalisation of the temperature values. The light gray dashed line indicates the total concentration used in the simulations. The reference temperature is the θ temperature of the non-phosphorylated molecule for full chain and consensus repeat, respectively. Reported error bars are obtained by block analysis over 10 blocks. (e) Representative fluorescent images of HEK-293 cells expressing opto-Ki-67 (FL) construct before (Light Off) and after (Light On) exposure to blue light. Cells were pretreated for 1h with either vehicle (DMSO), 0.5 μM okadaic acid (OA), to inhibit protein phosphatase 2, or 5 μM purvalanol A (PA), to inhibit CDKs. DNA was stained with Hoechst 33258; scale bars, 10μm. (f) Violin plot presenting quantification of results from (e); the number of foci per nucleus was counted. Statistical significance was assessed by one-way ANOVA on ranks (Kruskal–Wallis test) and pairwise *post-hoc* comparisons using the Mann–Whitney test. P-values were adjusted by the Benjamini-Hochberg method. (g) Overlaid NMR ^{1}H-^{15}N HSQC of unphosphorylated (blue) and CDK-phosphorylated (red) GFP-tagged Ki-67 consensus repeat. Each cross-peak corresponds to one residue. The seven new deshielded cross peaks (highlighted by a black flag)

appearing above 8.5 ppm in ^{1}H correspond to phosphorylated serines or threonines (^{1}H downfield chemical shift perturbation on phosphorylated Ser/Thr residues due to phosphate electronegativity). Non phosphorylated Ser/Thr residues are surrounded by a black oval. (h) Representative fluorescence images of *in vitro* phase separation assay with purified GFP-tagged Ki-67 consensus repeat (CR), non-phosphorylated (non-P) or *in vitro* phosphorylated with recombinant CDK1-cyclin B-CKS1 (P), at indicated dextran concentrations and time points; scale bars, 10μm.

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- SupplementaryInformation.pdf
- DATAS1Phosphoproteomicsdata.xlsx
- DATAS2CDKphosphorylationdata.xlsx
- DATAS3Proteindisorderprediction.xlsx
- DATAS4HumanproteinsinMLOs.xlsx
- SupplementarymovieS1.mpg
- SupplementarymovieS2A.mpg
- SupplementarymovieS2B.mpg