

## S1 Appendix

### Reusable rule-based cell cycle model explains compartment-resolved dynamics of 16 observables in RPE-1 cells

Paul F. Lang<sup>1,+,\*</sup>, David R. Penas<sup>2</sup>, Julio R. Banga<sup>2</sup>, Daniel Weindl<sup>3</sup>, and Bela Novak<sup>1,\*\*</sup>

<sup>1</sup>Department of Biochemistry, University of Oxford, Oxford, United Kingdom

<sup>2</sup>Computational Biology Lab, MBG-CSIC (Spanish National Research Council),  
Pontevedra, Spain

<sup>3</sup>Computational Health Center, Helmholtz Zentrum München Deutsches  
Forschungszentrum für Gesundheit und Umwelt (GmbH), Neuherberg, Germany

<sup>+</sup>Current Address: Biosim, Boston, Massachusetts, United States of America

\*plang@biosim.ai

\*\*bela.novak@bioch.ox.ac.uk

**Table A** — Variables of the core model.

ODE variable (Unit <sup>1</sup> )	Description	BNGL equivalent
t (min)	time	time
CycE (AU <sub>E</sub> )	cyclin E:Cdk2 complex	CCNE()
tE2f (AU <sub>E2f</sub> )	total transcription factor E2f	E2F(DBD!?,RB1!?,Ser332)
E2f (AU <sub>E2f</sub> )	free, unphosphorylated E2f	E2F(DBD,RB1,Ser332~u)
pE2f (AU <sub>E2f</sub> )	free, phosphorylated E2f	E2F(DBD,RB1,Ser332~p)
Rb (AU <sub>E2f</sub> )	unphosphorylated retinoblastoma protein (incl. (p)E2f:Rb)	RB1(E2F!?,Ser807_Ser811~u)
pRb (AU <sub>E2f</sub> )	phosphorylated retinoblastoma protein (always free)	RB1(E2F,Ser807_Ser811~p)
fRb (AU <sub>E2f</sub> )	free, unphosphorylated Rb	RB1(E2F,Ser807_Ser811~u)
pE2f:Rb (AU <sub>E2f</sub> )	phosphorylated E2f bound to retinoblastoma protein	E2F(DBD,RB1!1,Ser332~p). RB1(E2F!1,Ser807_Ser811~u)
E2f:PX (AU <sub>Px</sub> )	E2f bound to the promoter of gene X (CycA, CycE, E2f, FoxM1)	E2F(DBD!1,RB1,Ser332~u). *_promoter(E2F!1)
E2f:Rb (AU <sub>E2F</sub> )	unphosphorylated E2f bound to retinoblastoma protein	E2F(DBD,RB1!1,Ser332~u). RB1(E2F!1,Ser807_Ser811~u)
CycA (AU <sub>A</sub> )	cyclin A:Cdk1/2 complex	CCNA()
tEmi1 (AU <sub>Cdh1</sub> )	total early mitotic inhibitor 1 (Emi1)	FBX05(APC!?,FZR1,Ser182)
Emi1 (AU <sub>Cdh1</sub> )	free, unphosphorylated Emi1	FBX05(APC,FZR1,Ser182~u)
pEmi1 (AU <sub>Cdh1</sub> )	phosphorylated Emi1 (always free)	FBX05(APC,FZR1,Ser182~p)
Apc (AU <sub>Cdh1</sub> )	unphosphorylated anaphase promoting complex (APC/C) without substrate adaptor subunit	APC(FZR1_CDC20,FBX05,Ser355~u)
pApc (AU <sub>Cdh1</sub> )	free, phosphorylated APC/C	APC(FZR1_CDC20,FBX05,Ser355~p)
Cdh1 (AU <sub>Cdh1</sub> )	free, unphosphorylated APC/C substrate adaptor subunit Cdh1	FZR1(APC,FBX05,nTerm~u)
pCdh1 (AU <sub>Cdh1</sub> )	phosphorylated Cdh1 (always free)	FZR1(APC,FBX05,nTerm~p)
Apc:Cdh1 (AU <sub>Cdh1</sub> )	APC/C bound to Cdh1	APC(FZR1_CDC20!1,FBX05,Ser355~u). FZR1(APC!1,FBX05,nTerm~u)
pApc:Cdh1 (AU <sub>Cdh1</sub> )	phosphorylated APC/C bound to Cdh1	APC(FZR1_CDC20!1,FBX05,Ser355~p). FZR1(APC!1,FBX05,nTerm~u)
Apc:Cdh1:Emi1 (AU <sub>Cdh1</sub> )	APC/C bound to Cdh1 and Emi1	APC(FZR1_CDC20!1,FBX05!2,Ser355~u). FZR1(APC!1,FBX05!3,nTerm~u). FBX05(APC!2,FZR1!3,Ser182~u)
pApc:Cdh1:Emi1 (AU <sub>Cdh1</sub> )	phosphorylated APC/C bound to Cdh1 and Emi1	APC(FZR1_CDC20!1,FBX05!2,Ser355~p). FZR1(APC!1,FBX05!3,nTerm~u). FBX05(APC!2,FZR1!3,Ser182~u)
tFoxM1 (AU <sub>Fox</sub> )	total transcription factor FoxM1	FOXM1(DBD!?,Thr600)
pFoxM1 (AU <sub>Fox</sub> )	phosphorylated (i.e. active) transcription factor FoxM1	FOXM1(DBD,Thr600~p)
tCycB (AU <sub>B</sub> )	cyclin B:(p)Cdk1 complex	CCNB(CDK1_Thr14_Tyr15)
pFoxM1:PCycB (AU <sub>Pb</sub> )	phosphorylated FoxM1 bound to the promoter of cyclin B	FOXM1(DBD!1,Thr600~p). CCNB_promoter(FOXM1!1)
CycB:Cdk1 (AU <sub>B</sub> )	cyclin B in complex with unphosphorylated (i.e. active) Cdk1	CCNB(CDK1_Thr14_Tyr15~u)
Wee1 (AU <sub>Wee</sub> )	unphosphorylated (i.e. active) Wee1 kinase	WEE1(Ser123~u)
pCdc25 (AU <sub>C25</sub> )	phosphorylated (i.e. active) Cdc25 phosphatase	CDC25(pSites~p)
pGw (AU <sub>Gw</sub> )	phosphorylated Greatwall kinase	MASTL(Thr198~p)
pEnsa (AU <sub>Ensa</sub> )	total (incl. pEnsa:B55) phosphorylated forms of Endosulfine alpha and Arpp19	ENSA_ARPP19(PPP2R2B!?,Ser62_Ser67~p)
B55 (AU <sub>Ensa</sub> )	protein phosphatase 2A core dimer in complex with the regulatory subunit B55	PPP2R2B(ENSA_ARPP19)
pEnsa:B55 (AU <sub>Ensa</sub> )	phosphorylated Ensa bound to B55 complex	ENSA_ARPP19(PPP2R2B!1,Ser62_Ser67~p). PPP2R2B(ENSA_ARPP19!1)
tCdc20 (AU <sub>Cdh1</sub> )	total APC/C substrate adaptor subunit Cdc20	CDC20(APC!?)
pFoxM1:PCdc20 (AU <sub>Pc</sub> )	phosphorylated FoxM1 bound to the promoter of Cdc20	FOXM1(DBD!1,Thr600~p). CDC20_promoter(FOXM1!1)
Cdc20 (AU <sub>Cdh1</sub> )	free Cdc20	CDC20(APC)
pApc:Cdc20 (AU <sub>Cdh1</sub> )	phosphorylated APC/C bound to Cdc20	APC(FZR1_CDC20!1,FBX05,Ser355~p).CDC20(APC!1)

<sup>1</sup> As arbitrary units (AU) of dimension "concentration" with characteristic scale indicated in subscript.

\* one of CCNE, CCNA, E2F, FBX05, FOXM1.

**Table B** — Changelog of model versions.

Version	Changes
1.0.0	-
2.0.0	<ul style="list-style-type: none"><li>- Rescaled cell cycle length to 20 h (H1 human embryonic stem cells).</li><li>- Hard coded/eliminated parameters f1-f4 that were once introduced to describe loss of enzymatic access if target is in complex.</li><li>- Renamed species to avoid ambiguity in case-insensitive software.</li><li>- Ran version 1.0.0 for 9950 time units and used this state as initial conditions in version 2.0.0.</li><li>- Deleted observables.</li></ul>
2.1.0	<ul style="list-style-type: none"><li>- Removed the simplifying assumption from the ODE model that (un)binding of transcription factor to promoter does not affect the concentration of free transcription factor.</li></ul>
2.1.1	<ul style="list-style-type: none"><li>- Created separate promoters for Ce, Ca, E2f, Emi and Fox from one E2f activated promoter (formerly called Px).</li></ul>
2.1.2	<ul style="list-style-type: none"><li>- Introduced observables for parameter optimisation purposes (were never used in the end and specified via PEtab instead).</li></ul>
2.1.3	<ul style="list-style-type: none"><li>- Introduced interfaces for small molecule mediated inhibition of Ce and transcriptional/translational inhibition for Cb (were never used in the end).</li></ul>
2.1.4	<ul style="list-style-type: none"><li>- Introduced interface for small molecule mediated inhibition of Wee1 and removed it for Ce.</li></ul>
3.0.0	<ul style="list-style-type: none"><li>- Converted model into BioNetGen language.</li><li>- Introduced a systematic naming convention.</li></ul>
3.0.1	<ul style="list-style-type: none"><li>- Make E2F binding to DNA and its transcriptional activity independent of E2F phosphorylation state.</li><li>- Introduce baseline APC dephosphorylation.</li><li>- Allow dephosphorylation of DNA-bound FOXM1.</li></ul>
3.1.0	<ul style="list-style-type: none"><li>- Added DNA damage checkpoint (SKP2 and TP53 and CDKN1A).</li></ul>
3.2.0	<ul style="list-style-type: none"><li>- Added CDKN1B.</li></ul>
4.0.0	<ul style="list-style-type: none"><li>- Added compartmentalisation.</li></ul>

## Equations A: The restriction point submodel

Please refer to the [/versions/v0.0.1/](#) directory of the cell\_cycle\_model GitHub repository for parameter values and initial conditions. Note that for Equations (1)–(26) it is assumed that transcription factors are present in much higher concentration than promoters, i.e. promoter binding has negligible effect on free transcription factor concentration. This assumption will be removed for full cell cycle model versions  $\geq$  v2.1.0 (see also Table B)

$$\frac{dRb}{dt} = k_{DpRb} \cdot pRb - k_{PhRb} \cdot (CycD + CycE) \cdot Rb + k_{DeE2f} \cdot E2f \cdot Rb \quad (1)$$

$$\frac{dCycE}{dt} = k_{SyCe1} + k_{SyCe2} \cdot E2f \cdot Px - k_{DeCe} \cdot CycE \quad (2)$$

$$\begin{aligned} \frac{dE2f}{dt} &= k_{SyE2f1} + k_{SyE2f2} \cdot E2f \cdot Px - k_{DeE2f} \cdot E2f \\ &\quad + k_{DiE2fRb} \cdot E2f \cdot Rb - k_{AsE2fRb} \cdot E2f \cdot fRb \end{aligned} \quad (3)$$

$$\frac{dtE2f}{dt} = k_{SyE2f1} + k_{SyE2f2} \cdot E2f \cdot Px - k_{DeE2f} \cdot tE2f \quad (4)$$

$$\frac{dE2f \cdot Px}{dt} = k_{AsEPx} \cdot E2f \cdot (tDna - E2f \cdot Px) - k_{DiEPx} \cdot E2f \cdot Px \quad (5)$$

$$pRb = tRb - Rb \quad (6)$$

$$fRb = Rb - (tE2f - E2f) \quad (7)$$

$$E2f \cdot Rb = tE2f - E2f \quad (8)$$

## Equations B: The G1/S transition submodel

$$\begin{aligned} \frac{dtEmi}{dt} &= k_{SyEmi1} + k_{SyEmi2} \cdot E2f:Px - k_{DeEmi1} \cdot tEmi1 \\ &\quad - k_{DeEmi2} \cdot Apc:Cdh1:Emi1 \end{aligned} \quad (9)$$

$$\begin{aligned} \frac{dCycA}{dt} &= k_{SyCa1} + k_{SyCa2} \cdot E2f:Px \\ &\quad - (k_{DeCa1} + k_{DeCa2} \cdot Apc:Cdh1) \cdot CycA \end{aligned} \quad (10)$$

$$\frac{dCycE}{dt} = k_{SyCe1} + k_{SyCe2} \cdot E2f:Px - k_{DeCe} \cdot CycE \quad (11)$$

$$\begin{aligned} \frac{dApc:Cdh1:Emi1}{dt} &= k_{AsACE} \cdot Apc:Cdh1 \cdot (tEmi - Apc:Cdh1:Emi1) \\ &\quad - (k_{DiACE} + k_{DeEmi1} + k_{DeEmi2} + k_{PhCdhE} \cdot CycE \\ &\quad + k_{PhCdhA} \cdot CycA) \cdot Apc:Cdh1:Emi1 \end{aligned} \quad (12)$$

$$\begin{aligned} \frac{dpCdh1}{dt} &= (k_{PhCdhA} \cdot CycA + k_{PhCdhE} \cdot CycE) \cdot (tCdh1 - pCdh1) \\ &\quad - k_{DpCdh} \cdot pCdh1 \end{aligned} \quad (13)$$

$$\frac{dE2f:Px}{dt} = k_{AsEPx} \cdot E2f \cdot (tDna - E2f:Px) - k_{DiEPx} \cdot E2f:Px \quad (14)$$

$$Apc:Cdh1 = tCdh1 - pCdh1 - Apc:Cdh1:Emi1 \quad (15)$$

## Equations C: The G2/M transition submodel

$$\frac{dpEnsa}{dt} = k_{PhEnsa} \cdot pGw \cdot (tEnsa - pEnsa) - k_{DpEnsa} \cdot pEnsa : B55 \quad (16)$$

$$\frac{dCycB:Cdk1}{dt} = r_{Cdc25} \cdot (tCycB - CycB:Cdk1) - r_{Wee} \cdot CycB:Cdk1 \quad (17)$$

$$\begin{aligned} \frac{dpGw}{dt} &= k_{PhGw} \cdot CycB:Cdk1 \cdot (tGw - pGw) \\ &\quad - (k_{DpGw1} + k_{DpGw2} \cdot B55) \cdot pGw \end{aligned} \quad (18)$$

$$\begin{aligned} \frac{dpEnsa:B55}{dt} &= k_{AspEB55} \cdot B55 \cdot (pEnsa - pEnsa:B55) \\ &\quad - (k_{DipEB55} + k_{DpEnsa}) \cdot pEnsa:B55 \end{aligned} \quad (19)$$

$$B55 = t_{B55} - pEnsa:B55 \quad (20)$$

$$r_{Wee} = k_{Wee1} + \frac{(k_{Wee2} - k_{Wee1}) \cdot k_{DpWee} \cdot B55}{k_{DpWee} \cdot B55 + k_{PhWee} \cdot CycB:Cdk1} \quad (21)$$

$$r_{Cdc25} = k_{Cdc25.1} + \frac{(k_{Cdc25.2} - k_{Cdc25.1}) \cdot k_{PhCdc25} \cdot CycB:Cdk1}{k_{PhCdc25} \cdot CycB:Cdk1 + k_{DpCdc25} \cdot B55} \quad (22)$$

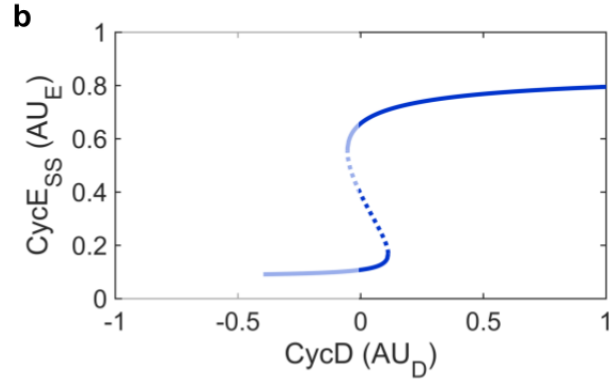
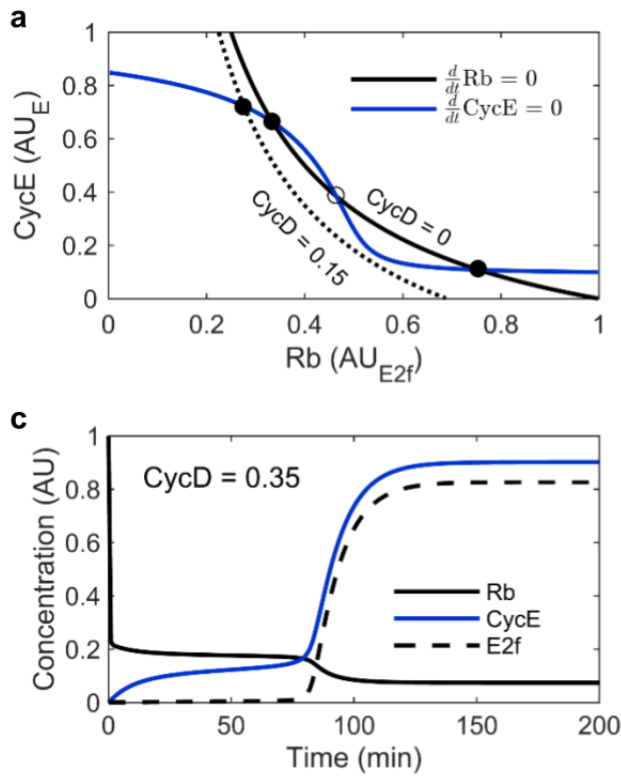
## Equations D: The M/A transition submodel

$$\begin{aligned} \frac{dpApc}{dt} &= k_{PhApc} \cdot (t_{Apc} - pApc) \cdot (CycB : Cdk1) \\ &\quad - k_{DpApc} \cdot B55 \cdot pApc \end{aligned} \quad (23)$$

$$\begin{aligned} \frac{dpApc:Cdc20}{dt} &= k_{AsAC20} \cdot (t_{Cdc20} - pApc:Cdc20) \cdot (pApc - pApc:Cdc20) \\ &\quad - (k_{DiAC20} + k_{DpApc} \cdot B55) \cdot pApc:Cdc20 \end{aligned} \quad (24)$$

$$\frac{dtCycB}{dt} = k_{SyCb} - (k_{DeCb1} + k_{DeCb2} \cdot pApc:Cdc20) \cdot t_{CycB} \quad (25)$$

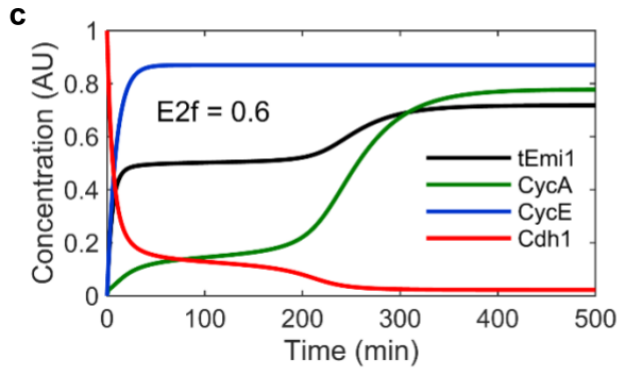
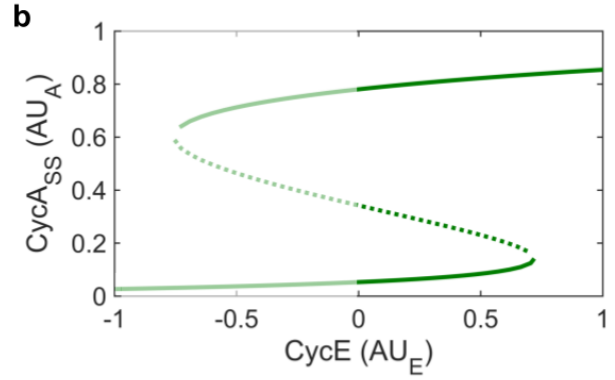
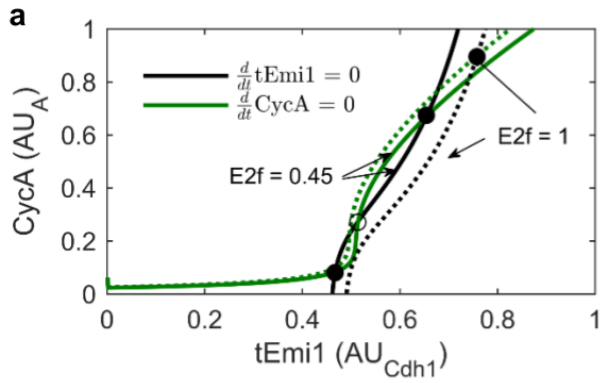
$$\begin{aligned} \frac{dCycB:Cdk1}{dt} &= k_{SyCb} - (k_{DeCb1} + k_{DeCb2} \cdot pApc:Cdc) \cdot CycB:Cdk1 \\ &\quad + r_{Cdc25} \cdot (t_{CycB} - CycB:Cdk1) - r_{Wee} \cdot CycB:Cdk1 \end{aligned} \quad (26)$$



**Fig A — Restriction point submodel.** **a** Phase plane showing the nullclines for CycE and Rb intersecting at two stable (filled circles) and one unstable (unfilled circle) steady states when CycD is set to 0 AU<sub>D</sub>. There is only one steady state at CycD = 0.15 AU<sub>D</sub>. The system was reduced to two dimensions by making a steady state approximation for E2f and calculating E2f:Rb and pRb via conservation laws. **b** Bifurcation diagrams of the non-reduced system with constant E2F, showing stable (solid line) and unstable (dotted line) steady states of CycE. Line endings within the axes limits indicate disappearance of a steady state. Unphysiological regions are semi-transparent. **c** Time course of unphosphorylated Rb, CycE and E2f at CycD = 0.35 AU<sub>D</sub>.

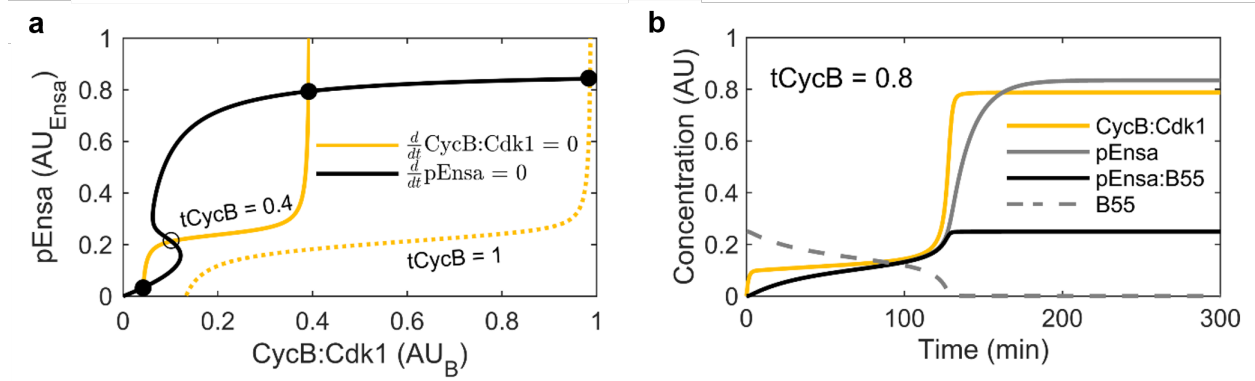
For variable abbreviations please refer to Table A. Parameter values and initial conditions are available the </versions/v0.0.1/> directory of the cell\_cycle\_model GitHub repository.



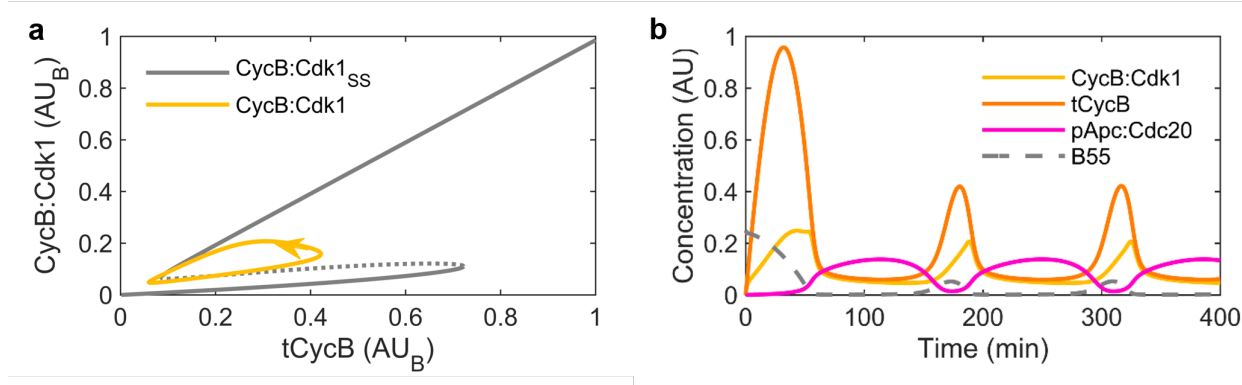


**Fig B — G1/S transition submodel.** **a** Phase plane showing the nullclines for total Emi1 (tEmi1) and CycA intersecting at two stable (filled circles) and one unstable (unfilled circle) steady states when E2f is set to 0.45 AU<sub>E2f</sub>. There is only one steady state at E2f = 1 AU<sub>E2f</sub>. The system was reduced to the two plotted dimensions by making steady state approximations for the remaining variables. **b** Bifurcation diagram of the non-reduced system at E2f = 0.7 AU<sub>E2f</sub> with CycE as bifurcation parameter. Unphysiological regions are semi-transparent. **c** Time course of tEmi, CycA, CycE and Cdh1 at E2f = 0.6 AU<sub>E2f</sub>. For variable abbreviations please refer to Table A. Parameter values and initial conditions are available

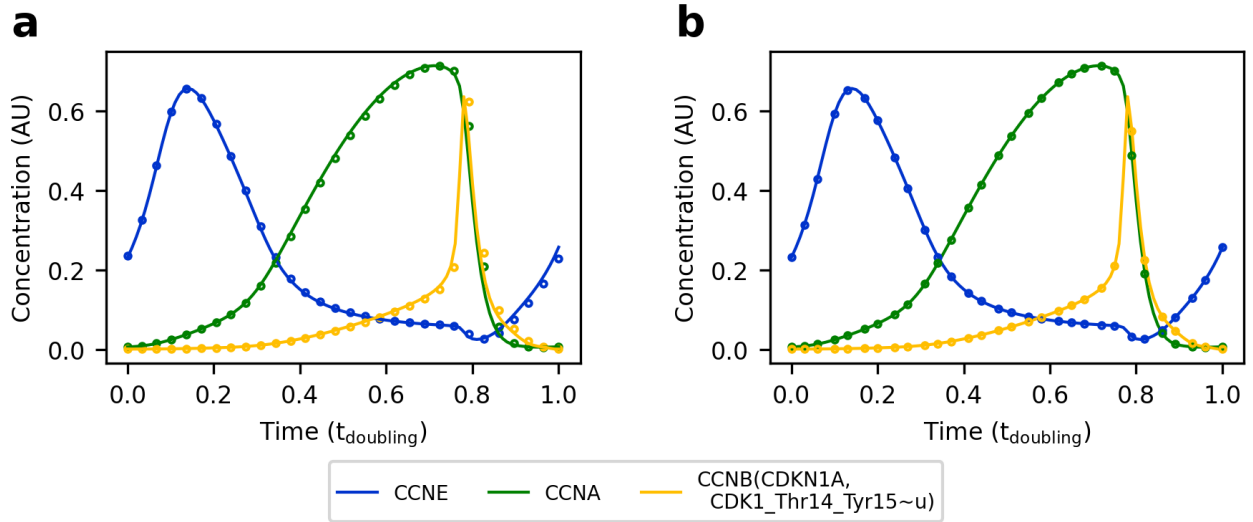
able the [/versions/v0.0.1/](#) directory of the cell\_cycle\_model GitHub repository.



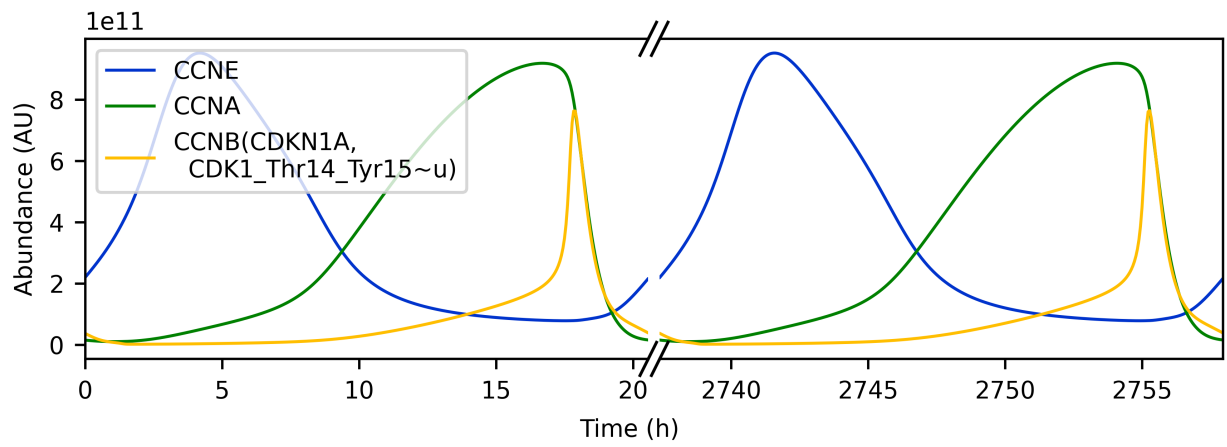
**Fig C** — G2/M transition submodel as developed by Vinod and Novak [1]. **a** Phase plane showing the nullclines for CycB:Cdk1 and pEnsa intersecting at two stable (filled circle) and one unstable (unfilled circle) steady states when total cyclin B (tCycB) is set to 0.4 AU<sub>B</sub>. There is only one steady state at tCycB = 1 AU<sub>B</sub>. The system was reduced to the two plotted dimensions by making steady state approximations for the remaining variables. Plot recreated from [1]. **b** Time course of CycB:Cdk1, pEnsa, pEnsa:B55, and B55 at tCycB = 0.8 AU<sub>B</sub>. For variable abbreviations please refer to Table A. Parameter values and initial conditions are available from the [/versions/v0.0.1/](#) directory of the cell\_cycle\_model GitHub repository.



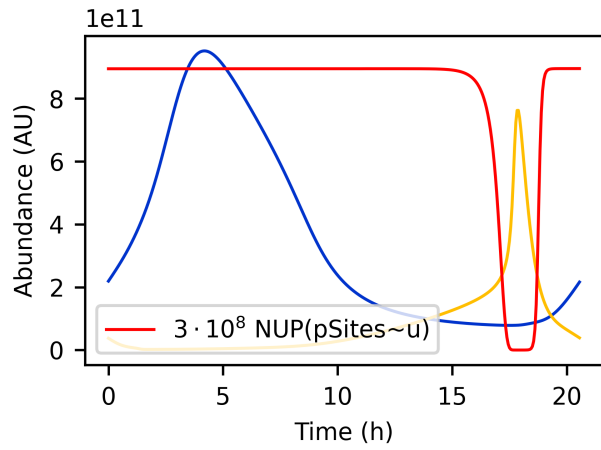
**Fig D — G2/M transition submodel with negative feedback of the M/A transition.** **a** CycB:Cdk1 steady states (grey) of the G2/M transition submodel shown in the bifurcation diagram of Fig 1F are overlaid with a stable limit cycle trajectory (yellow) of the system in the CycB:Cdk1 – tCycB subspace at a cyclin B synthesis rate of  $k_{\text{SyCb}} = 0.05 \text{ min}^{-1}$ . The periodic over- and undershooting of the upper and lower bifurcation point, respectively, results from combining the toggle switch of the G2/M submodel with the negative feedback of the M/A submodel. **b** Time course of CycB:Cdk1, tCycB, pApc:Cdc20 and B55 at  $k_{\text{SyCb}} = 0.05 \text{ min}^{-1}$ . For variable abbreviations please refer to Table A. Parameter values and initial conditions are available the </versions/v0.0.1/> directory of the cell.cycle.model GitHub repository.



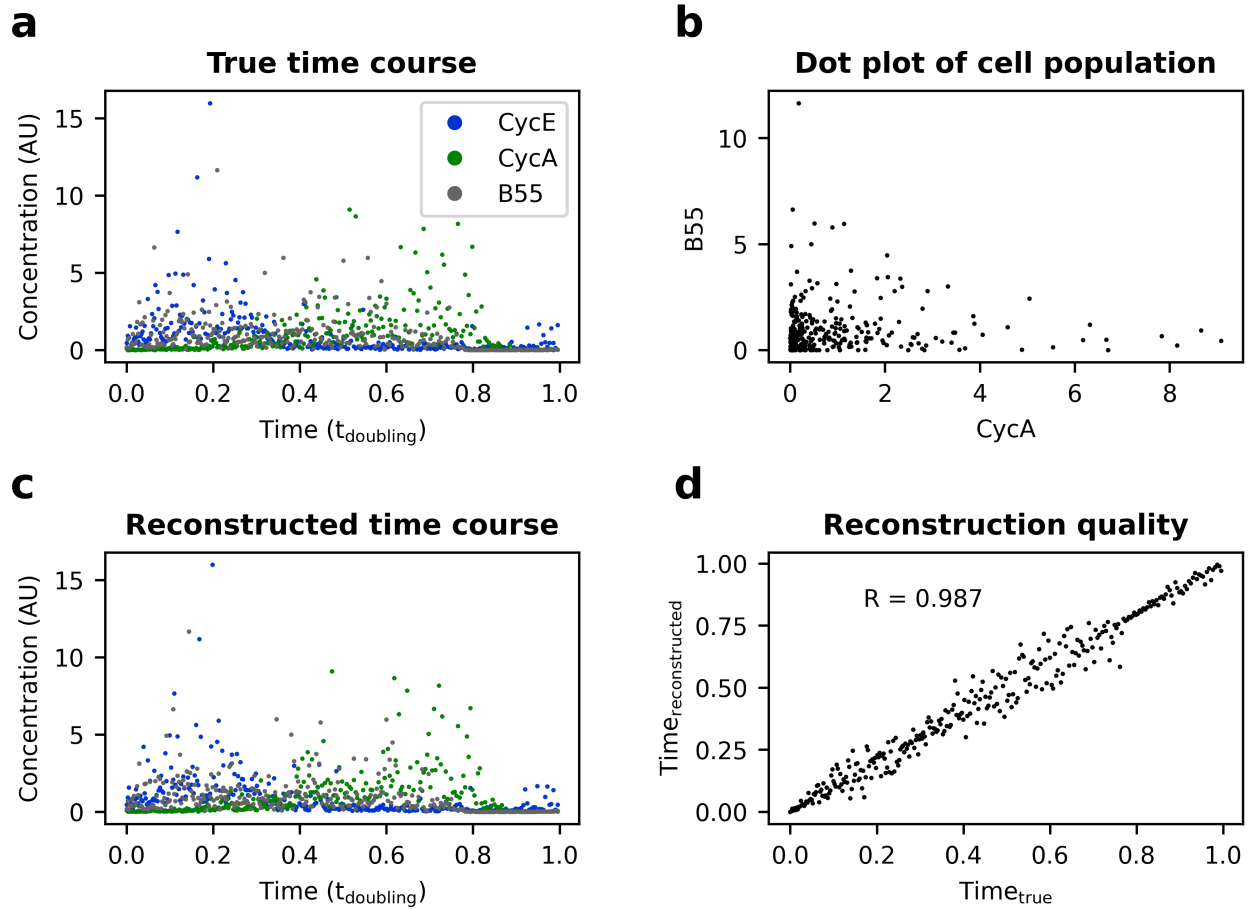
**Fig E — Comparison between model versions.** All simulations were performed with Copasi LSODA and default settings. Lines represent simulations of model version 2.1.4. **(a)** Circles represent second cycle of model version 1.0.0. First cycle is not shown, as the initial conditions of this model do not lie on the limit cycle trajectory. **(b)** Circles represent model version 3.0.0, simulated after export to SBML.



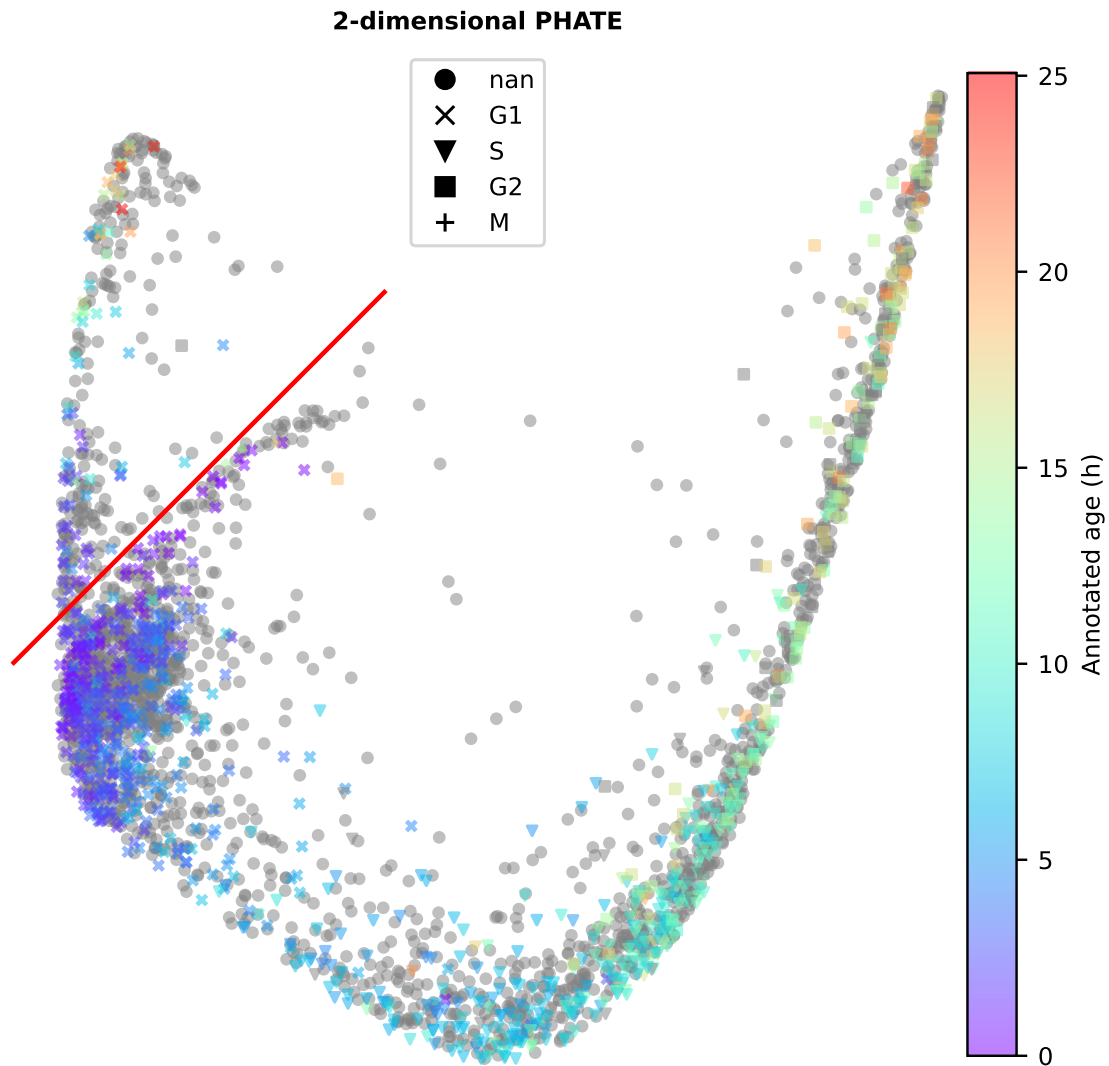
**Fig F** — Cell cycle model with nuclear and cytoplasmic compartments. Time course of three representative species over 2758 h. Model version 4.0.0.



**Fig G — Cell cycle model with nuclear pore phosphorylation.** Time course of three representative species. Unphosphorylated nuclear pore complex NUP(pSites~u) are a proxy for an intact nuclear membrane. Model based on version 4.0.0 plus nuclear pore phosphorylation.

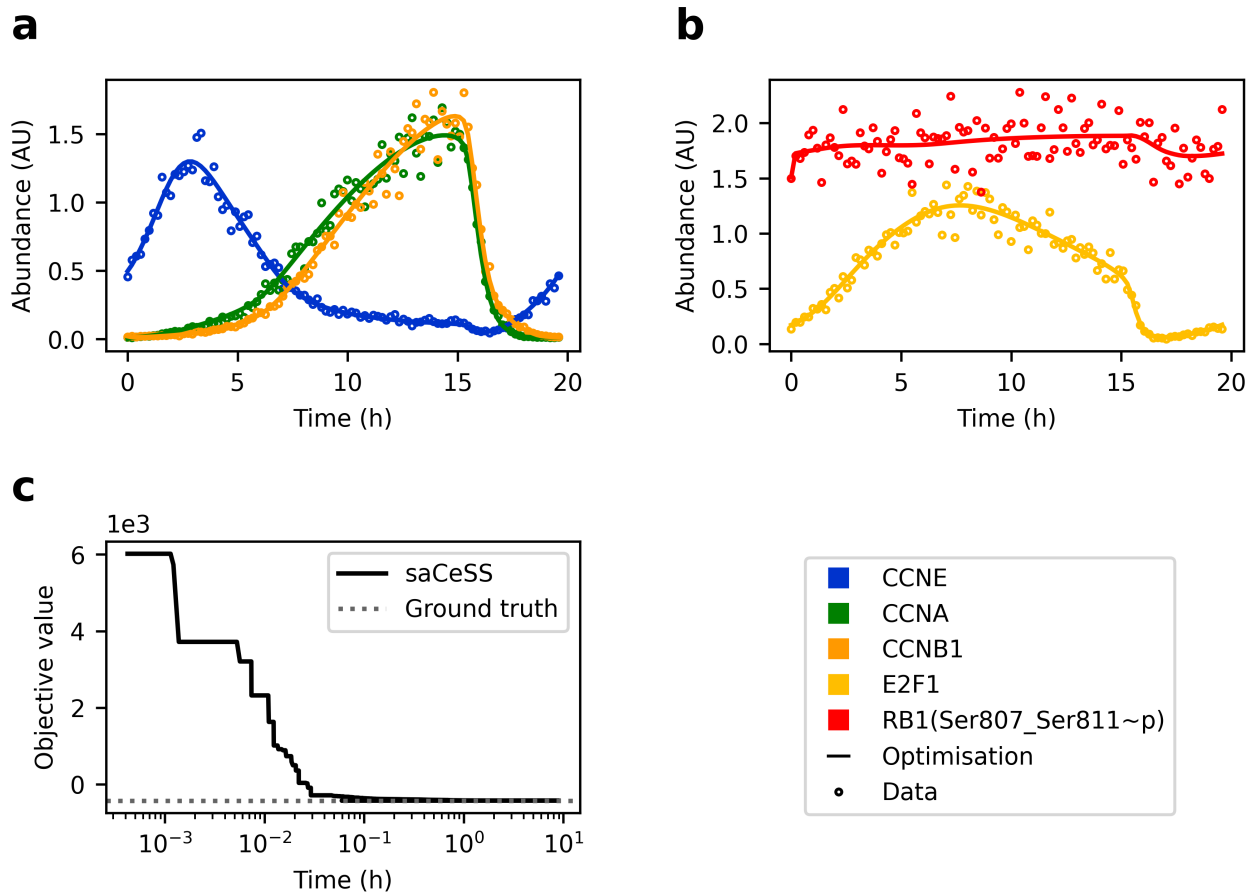


**Fig H** — Cell cycle trajectory reconstruction from noise-free simulated data with reCAT. (a) 300 cells were sampled across one cell cycle model simulation (version 2.1.4), such that cell density decreases exponentially over the cell cycle with one cell cycle length half-life. Of all 9 variables that were supplied to reCAT, only cyclin E, cyclin A and B55 are shown. (b). Scatterplot of the sampled cells from A. (c) Reconstructed cell cycle trajectory. (d) Correlation between true and reconstructed cell cycle time for each cell.

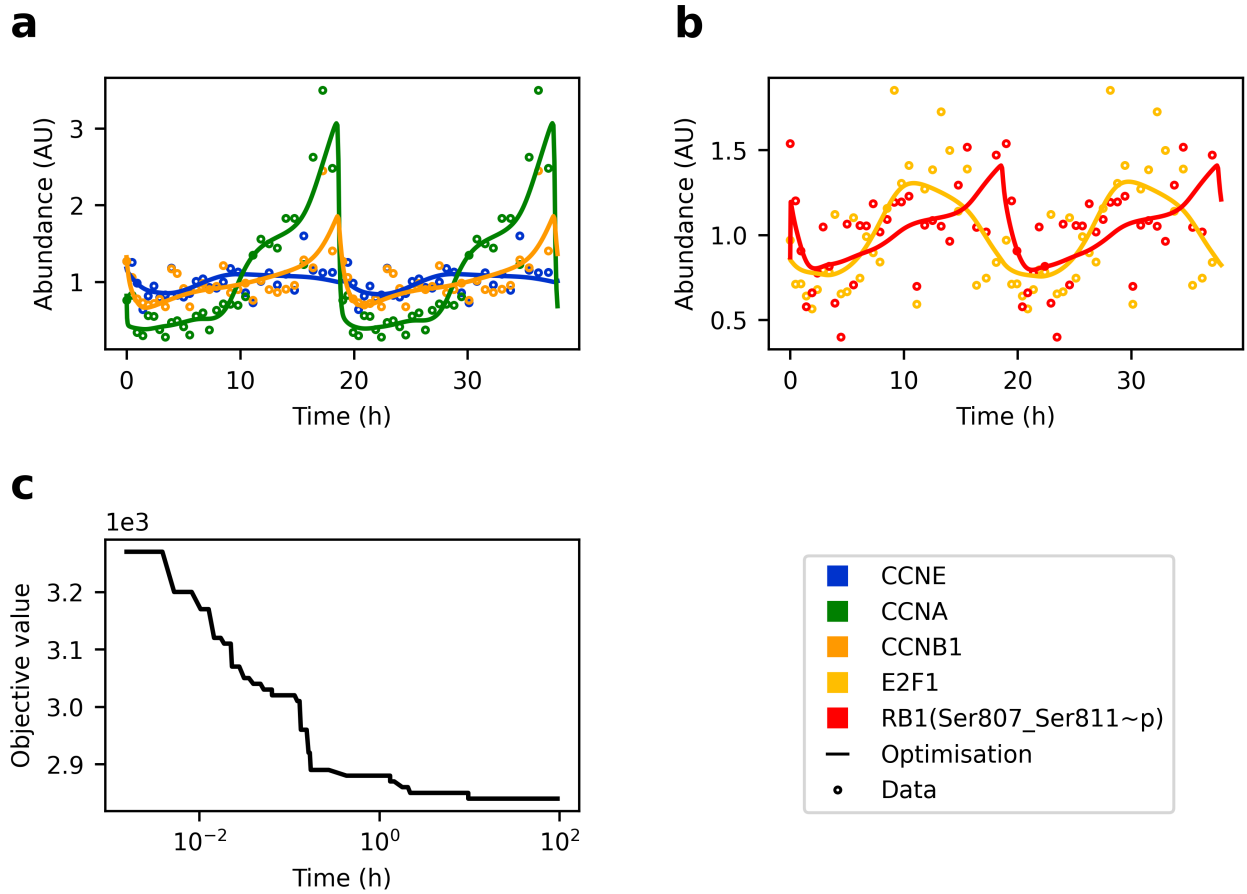


**Fig I — Gating proliferating cells using PHATE.** PHATE [2] was performed to embed publicly available 4i measurements of an asynchronously dividing RPE-1 population [3] in two dimensions (PHATE1 and PHATE2). PHATE separates the cell population into a left, central and right arm. The left arm corresponds to G0 cells [3]. Cells above the line indicated by the manually adjusted red line segment were discarded before the proliferative trajectory was reconstructed with reCAT. For better visualisation only one third of all cells are shown.

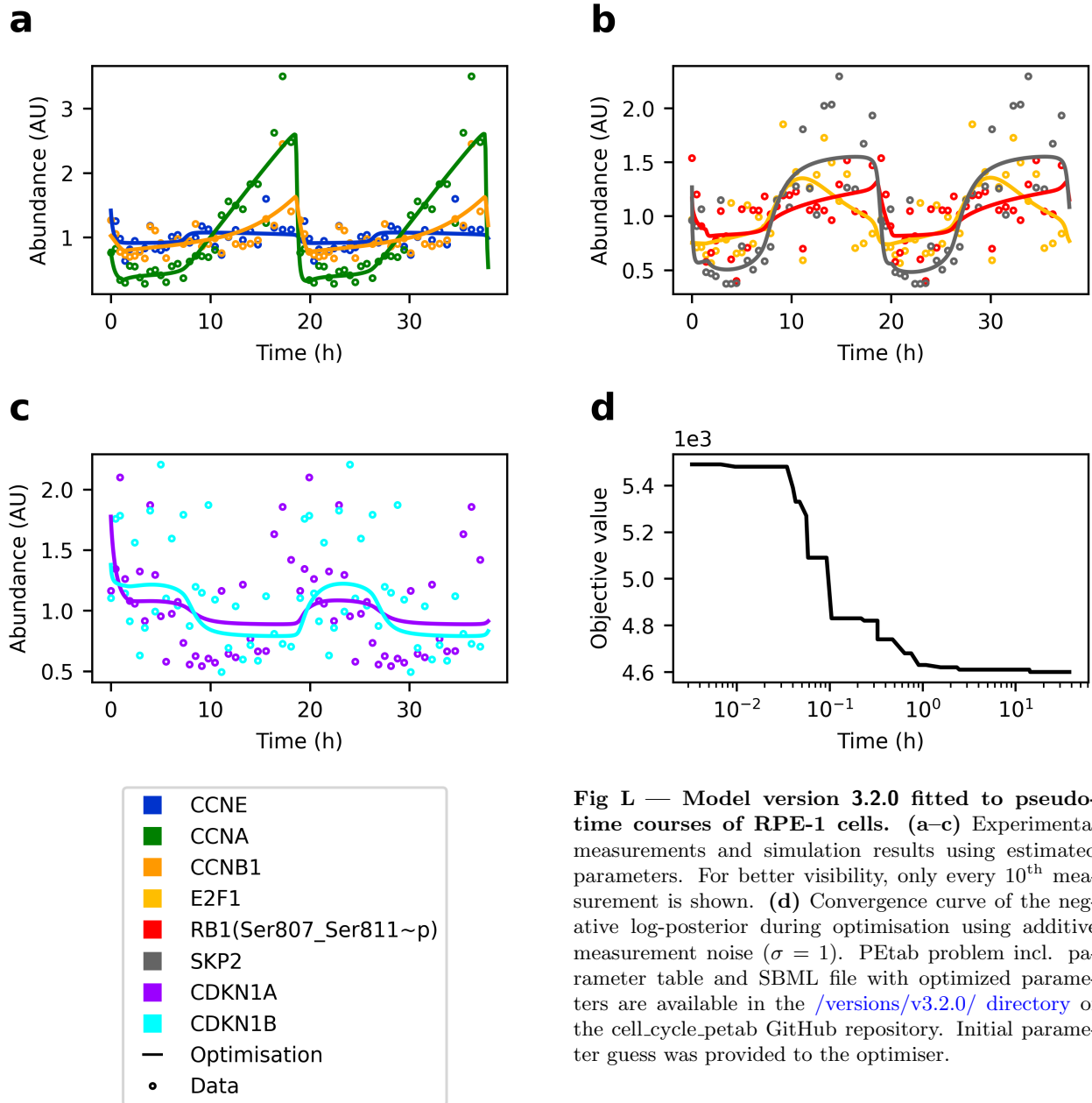




**Fig J — Testing self-adaptive Cooperative enhanced Scatter Search (saCeSS).** 300 datapoints of five observables with exponentially increasing spacing were generated from simulation results of model version 3.0.0. The datapoints were corrupted with  $\mathcal{N}(1, 0.1^2)$  multiplicative noise (circles). saCeSS optimised the parameters from within a  $[0.1 \cdot \theta_{true}, 10 \cdot \theta_{true}]$ . **a, b** The lines show simulated time courses for the observables, using the parameter set found by saCeSS (legend in lower right panel). **c** Convergence curve of the negative log-likelihood during optimisation using additive measurement noise ( $\sigma = 0.1$ ). PETab problem incl. parameter table and SBML file with optimized parameters are available in the [/versions/v3.0.0/PETab\\_PL\\_v3.0.0sim](#) directory of the cell\_cycle\_petab GitHub repository. Initial parameter guess was ignored.



**Fig K — Model version 3.0.1 fitted to pseudo-time courses of RPE-1 cells.** (A, B) Experimental measurements and simulation results using estimated parameters. For better visibility, only every 10<sup>th</sup> measurement is shown. (C) Convergence curve of the negative log-posterior during optimisation using additive measurement noise ( $\sigma = 1$ ). PETab problem incl. parameter table and SBML file with optimized parameters are available in the [/versions/v3.0.1/](#) directory of the cell\_cycle\_petab GitHub repository. Initial parameter guess was provided.



**Fig L** — Model version 3.2.0 fitted to pseudo-time courses of RPE-1 cells. (a–c) Experimental measurements and simulation results using estimated parameters. For better visibility, only every 10<sup>th</sup> measurement is shown. (d) Convergence curve of the negative log-posterior during optimisation using additive measurement noise ( $\sigma = 1$ ). PETab problem incl. parameter table and SBML file with optimized parameters are available in the [/versions/v3.2.0/](#) directory of the cell\_cycle\_petab GitHub repository. Initial parameter guess was provided to the optimiser.

## Text A: Merging submodels

When merging the submodels to a full cell cycle model, we had to introduce (a) new reactions (e.g.: CycA mediated Rb phosphorylation; Apc:Cdh1 mediated CycB degradation) and (b) new species (e.g.: pApc:Cdh). To connect the G1/S submodel with the G2/M submodel, we also introduced the transcription factor FOXM1. FOXM1 is synthesized by E2f and activated by Cdks [4]. It binds to the promoter of cyclin B to promote its synthesis [5]. Furthermore, certain molecules that had accumulated during earlier cell cycle stages (tE2F, tEmi1, CycA, tFoxM1 tCdc20) still had to be degraded. For the SCF substrates tE2f [6] and tEmi1 [7] (but for simplicity not for CycE and CycA) this was accomplished by introducing a new, phosphorylated species that was targeted for ubiquitination. These species are phosphorylated by CycA and CycB, thus introducing additional negative feedback loops on E2F (e.g.: E2F activates CycA, which inhibits E2F; and E2F activates CycB via CycA mediated FOXM1 phosphorylation, and CycB inhibits E2F) and on Emi1 (e.g.: Emi1 inhibits Apc:Cdh1 inhibits CycA/B inhibits Emi1). For the (p)Apc:Cdh1 substrates tFoxM1 [8] and tCdc20 [9] a (p)Apc:Cdh1 mediated degradation reaction was introduced. This again results in new feedback loops. For instance, FoxM1 activates CycB. CycB and (p)Apc:Cdh1 mutually inhibit each other. (p)Apc:Cdh1 then inhibits FoxM1. The (p)Apc:Cdh1 mediated degradation of tFoxM1 ensures that tCycB is kept low in G1 phase. Similarly, for CycA/B a pApc:Cdc20 mediated degradation reaction was introduced, resulting in yet another negative feedback loop (CycA inhibits (p)Apc:Cdh1, inhibits pApc:Cdc20, inhibits CycA). These changes also required modifications in the parameters to preserve characteristic features of each submodel:

- RP: The restriction point is a point after which cell cycle progression is mitogen independent.
- G1/S: There is a delay between CycE and CycA accumulation.
- G2/M: There is a sharp increase in CycB:Cdk1.
- M/A: Cyclin B is rapidly and almost completely degraded.

To reduce the complexity of this task, the submodels were fused in a stepwise manner. The resulting model of the whole cell cycle (`cell_cycle.v1.0.0`) comprised 37 species (Table A). The corresponding equations are available in file [/versions/v1.0.0/cell\\_cycle.v1.0.0.ode](#) of the `cell_cycle_model` GitHub repository.

## Text B: Naming conventions in the BioNetGen model

The BNGL syntax also lends itself well to add semantic meaning to the model variables. This enables to conceive the following naming convention:

- Proteins
  - Protein molecule types were named by their Human Genome Organization Gene Nomenclature Committee (HGNC) short name. No distinction is made between protein isoforms that are transcribed from the same gene. Example: `RB1()`.
  - Protein family types were described by a shortened protein HGNC name where unambiguously possible. Example: the set of `CCNE1` and `CCNE2` was collectively referred to as `CCNE()`. Otherwise, they were described as the concatenation of HGNC names separated by underscores. Example: `ENSA_ARPP19()`.
- Promoters
  - Promoters were described by the protein they activate, suffixed by `_promoter`. Example: `CDC20_promoter()`.
- Binding sites
  - If the binding site is bound by a protein or protein family, the binding site is named by the protein (family) name. Example: `PPP2R2B(ENSA_ARPP19)`.
  - If the binding site is bound by a promoter, the binding site is named `DBD` (short for DNA binding domain). Example: `FOX1(DBD)`.
- Phosphorylation sites
  - The unphosphorylated state is indicated with a `u`, the phosphorylated state with a `p`.
  - Single site phosphorylation were named by the three letter amino acid code, suffixed with the amino acid position. Example: `MASTL(Thr198~u~p)`.
  - Known multisite phosphorylations were named by the concatenation of the single site phosphorylation names, separated with an underscore. Example `RB1(Ser807_Ser811~u~p)`. Note that no comma separates `Ser807` and `Ser811` in this example. Therefore, `Ser807_Ser811` is treated as a single site.
  - Unknown phosphorylation sites could have any name. Example `FZR1(APC,FBX05,nTerm~u~p)`.
- Exceptions
  - `E2F()` means only the activating E2F transcription factors `E2F1`, `E2F2` and `E2F3`.
  - The anaphase-promoting complex was named `APC()`.
  - The nuclear pore complex was named `NUP()`.

The exceptions indicate that some generalisability was traded off against concise syntax. Of note, as much as any biochemical model is incomplete, the use of the above syntax shall not imply that no other phosphorylation sites, binding sites or molecule types than those explicitly described by the model are involved in the real phenomenon of cell cycle control. Rather, it shall just help to describe more precisely which biological phenomena are considered in the model and which are not.

## **Text C: Adding CDKN1B to the cell cycle model**

Like CDKN1A, the CDKN1B cycling kinase inhibitor, also known as p27, confers ultrasensitivity. Together, the RP, the G1/S transition, the DNA damage checkpoint, and the mutual inhibition between CDKN1B and cyclin:Cdk complexes represent four partially redundant mechanisms to ensure bistability in G1. Redundancy confers robustness in varying or perturbed conditions. As robustness under perturbation seems to be an important feature of the cell cycle, CDKN1B was added to the model, along with FOXO transcription factors which activate its expression [10]. Unlike CDKN1A, CDKN1B does not bind cyclin B:Cdk complexes [11]. There is also clear evidence that Thr187 phosphorylation not only prevents cyclin:Cdk binding, but also targets CDKN1B for SKP2 mediated degradation.

## Text D: Introducing compartmentalisation

To implement nucleocytoplasmic shuttling, the following points were considered:

- The nucleocytoplasmic transport reactions described in the literature may only represent a subset of the nucleocytoplasmic transport reactions that are happening in the cell.
- We aim at identifying an extensive biochemical reaction network for which a parameterisation exists, such that the network is capable of explaining the regulation of the full cell cycle.
- We aim to demonstrate that the model can generate sustained oscillations with manually chosen parameters.

In agreement with these considerations, translocation reactions for all molecular species, except for those that are exclusively located in either compartment (i.e. promoters), were added to the model. The kinetic constants for the transport reactions were set to very large and equal values (to be later constrained during parameter estimation). This ensured that the oscillatory behaviour of single-compartment versions of the model was retained. More specifically, the following adaptations were made to version 3.2.0:

- Four compartments were defined using BNGL: a  $1 \times 10^{-7}$  dm thick plasma membrane (`P1m`,  $2 \times 10^{-14}$  dm<sup>3</sup>), cytoplasm (`Cyt`,  $1 \times 10^{-12}$  dm<sup>3</sup>), a  $1 \times 10^{-7}$  dm thick nuclear membrane (`Num`,  $9 \times 10^{-15}$  dm<sup>3</sup>) and a nucleus (`Nuc`,  $1 \times 10^{-12}$  dm<sup>3</sup>).
- Nuclear pores `NUP(pSites~u~p)` were inserted into the nuclear membrane.
- Instead of describing the complete Ran-GTP pathway, simple import and export reactions through unphosphorylated nuclear pores were added for all species that do not contain promoters. The kinetic constants for import and export were set to same very large value to allow quasi-free diffusion of molecules.
- The initial concentration of promoters was set to zero in the cytoplasm and a positive value in the nucleus.
- Reaction rules were made independent of localisation. I.e. reactions are unaware of the compartment they occur in, use the same kinetic constants independent of the compartment, and fire at any location where all reactants are present. Only protein synthesis reactions contain a notion of localisation, namely that proteins are synthesized into the cytoplasm.
- Nuclear envelope breakdown was imitated by introducing a rule for `CCNB(CDKN1A,CDK1_Thr14_Tyr15~u)` mediated nuclear pore phosphorylation, allowing for quasi-free diffusion of all pure protein species.

The above changes lead to model version 4.0.0. Simulation results are shown in Fig G. However, due to stiffness arising from nuclear envelope breakdown, we switched it off throughout the present study.

## Text E: Considerations on the effect of cell cycle arrest and trajectory reconstruction

Stallaert *et al.* [3] found that a substantial fraction of RPE-1 cells exit the proliferative cell cycle trajectory into a non-proliferative G0 arm. Removing such cells results in a purely proliferative trajectory. However, it is unclear whether the distribution of the cells along the trajectory follows the pseudo-time expressed by Equation 1 in the main text. This is because G0 cells may not enter the proliferative cell cycle at the same point of the trajectory where they left. Should they do parts of the proliferative trajectory twice, the calculated pseudo-time would become artificially stretched in these regions. Conversely, should they skip parts of the proliferative trajectory, the calculated pseudo-time would become artificially compressed. It is clear that repeating mitosis and skipping parts of S-phase via a G0 trajectory cannot be the norm, as this would lead to regular chromosome aberrations. Nevertheless, former G0 cells may move along a trajectory that is mostly parallel to the proliferative trajectory, but shifted in a few states for some time, before completely merging with the proliferative trajectory. Such parallel trajectories would bias the reconstructed trajectory in the direction of the shift. Yet, overall the trajectory reconstruction would only be strongly affected if alternative trajectories skip or redo large stretches of the proliferative trajectory and a large amount of cells at any given time would move on these trajectories. As both can be considered highly speculative at present, it seems reasonable to assume that removal of G0 cells from an asynchronous cell population leads to a quasi asynchronous cell population that mostly moves along the proliferative cell cycle trajectory. That is, the cell population can be used for reconstructing a proliferative trajectory from snapshot data and for calculating pseudo-time from ranks via Equation 1.



## Text F: Handling real-world data in parameter estimation

Using real experimental data instead of an artificial dataset required us to make five minor modifications to the optimisation strategy. First, simulations with the optimised parameters did not result in oscillations. By duplicating the experimental data to obtain two consecutive cycles, oscillatory behaviour could be enforced. Second, simulations with these optimised parameters unexpectedly showed small dampened oscillations of observables for a short period of time, just after the metaphase anaphase transition. Attempts of manually adjusting the parameters led to the following hypothesis. Residual CCNA and CCNB1 fluorescence signal after anaphase in the Stallaert dataset enforce too high levels of these two species in the model. As a consequence, there is substantial E2F1 phosphorylation, and thus degradation in the model, especially when the activity of the counteracting phosphatase PPP2R2B is suppressed by action of CCNB1. To overcome these high degradation levels, E2F1 synthesis rates must be high. In other words, E2F abundance is highly responsive to changes in kinase/phosphatase ratio. Once kinase/phosphatase ratio drops, E2F1 shoots up, activates slow CCNA synthesis, which results in E2F1 phosphorylation and thus degradation with a delay. Attempts to manually reduce E2F1 degradation and synthesis rates failed to produce sustained oscillations. However, the active concentration of cyclins may be lower than the antibody signal suggests, for instance due to incomplete background subtraction in the dataset. To account for this possibility, the offset parameters  $o_k$  in the observation model defined in Equation (2) of the main text were allowed to deviate from zero. Third, using real world data, the global optimum may lie out of the specified parameter bounds. Therefore, active windows (i.e. windows where the best possible solution found by the optimiser lies at a boundary) were shifted in the corresponding direction by a factor of 10. These shifts were performed iteratively between the multiple job submissions to that were required due to limitations of maximal job execution times on the cluster. Fourth, saCeSS tended to perform better with DHC than parPE/ipopt as local solver option. Last, to further incorporate knowledge that cyclins are present in higher concentrations than E2F1 and RB1, Laplacian priors were introduced for the scaling parameters of the cyclins.

## References

- [1] P. K. Vinod and Bela Novak. “Model scenarios for switch-like mitotic transitions”. *FEBS Letters* 589.6 (2015), pp. 667–671.
- [2] Kevin R. Moon et al. “Visualizing structure and transitions in high-dimensional biological data”. *Nature Biotechnology* 37.12 (2019), pp. 1482–1492.
- [3] Wayne Stallaert et al. “The structure of the human cell cycle”. *Cell Systems* 0.0 (2021).
- [4] J. Millour et al. “ATM and p53 regulate FOXM1 expression via E2F in breast cancer epirubicin treatment and resistance”. *Mol Cancer Ther* 10.6 (2011), pp. 1046–58.
- [5] H. Murakami et al. “Regulation of yeast forkhead transcription factors and FoxM1 by cyclin-dependent and polo-like kinases”. *Cell Cycle* 9.16 (2010), pp. 3233–42.
- [6] A. Marti et al. “Interaction between ubiquitin–protein ligase SCF and E2F-1 underlies the regulation of E2F-1 degradation”. *Nature Cell Biology* 1.1 (1999), p. 14.
- [7] Daniele Guardavaccaro et al. “Control of Meiotic and Mitotic Progression by the F Box Protein  $\beta$ -Trcp1 In Vivo”. *Developmental Cell* 4.6 (2003), pp. 799–812.
- [8] J. Laoukili et al. “FoxM1 is degraded at mitotic exit in a Cdh1-dependent manner”. *Cell Cycle* 7.17 (2008), pp. 2720–6.
- [9] C. M. Pflieger and M. W. Kirschner. “The KEN box: an APC recognition signal distinct from the D box targeted by Cdh1”. *Genes Dev* 14.6 (2000), pp. 655–65.
- [10] Fangfang Yang et al. “The Akt/FoxO/p27Kip1 axis contributes to the anti-proliferation of pentoxifylline in hypertrophic scars”. *Journal of Cellular and Molecular Medicine* 23.9 (2019), pp. 6164–6172.
- [11] Alessia Montagnoli et al. “Ubiquitination of p27 is regulated by Cdk-dependent phosphorylation and trimeric complex formation”. *Genes & Development* 13.9 (1999), pp. 1181–1189.