Modelling Signalling Networks with Incomplete Protein Activation Information: A P Framework of the KaiABC Oscillator

T.Hinze[1] T.Lenser[2] G.Escuela[2] I.Heiland[1] S.Schuster[1]

{thomas.hinze,heiland.ines,stefan.schu}@uni-jena.de
{thorsten.lenser,gabi.escuela}@uni-jena.de

Friedrich Schiller University Jena

[1]Department Bioinformatics at School of Biology/Pharmacy

[2]Bio Systems Analysis Group

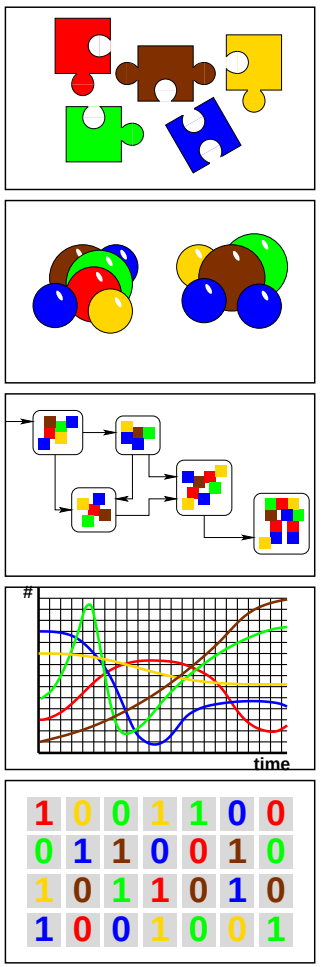
Tenth International Workshop on Membrane Computing (WMC10)

FORSYS Partner Initiative

Outline

CSMs with Incomplete Protein Activation Information: KaiABC Oscillator

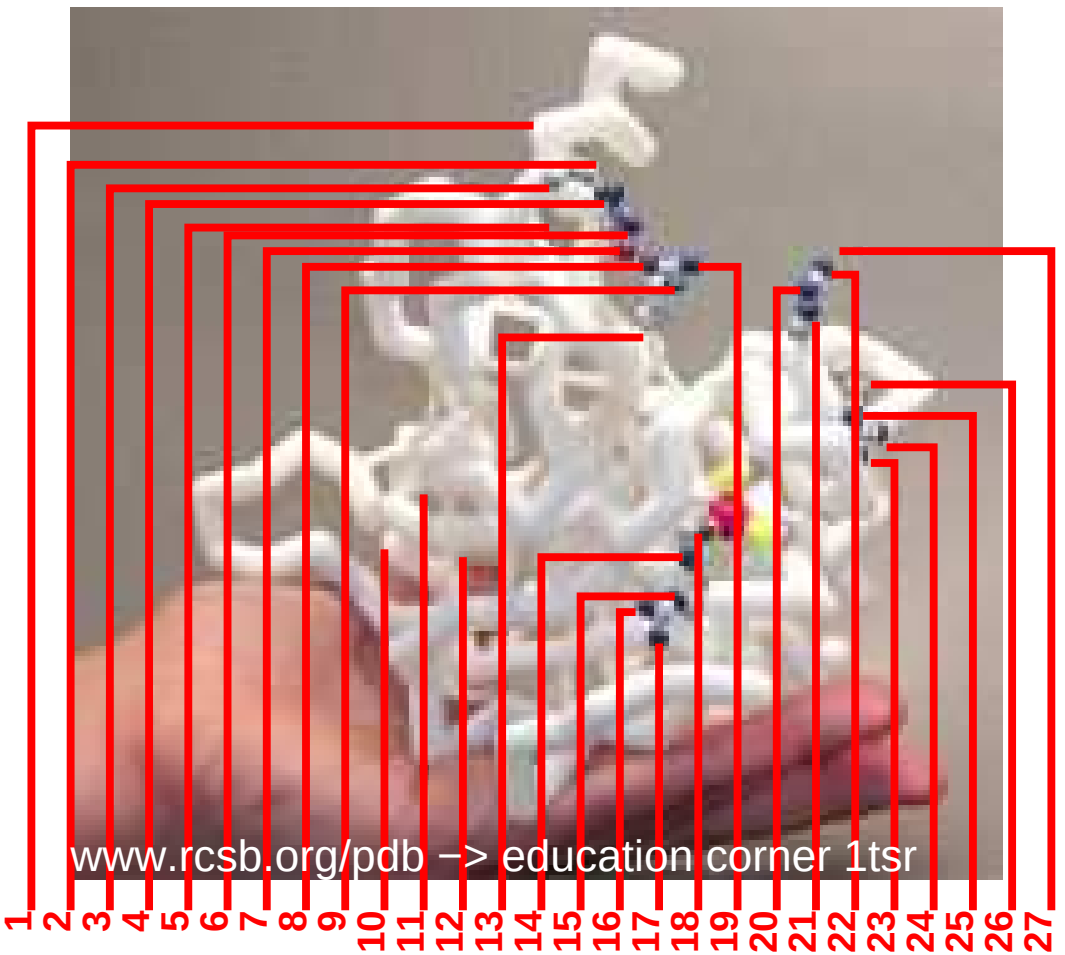
1. Motivation
2. Cell signalling modules (CSM)
3. P system framework Π_{CSM}
4. The KaiABC Oscillator: A circadian clock
5. Case Study Π_{KaiABC}
6. Outlook and acknowledgement



FORSYS Partner Initiative

Combinatorial Explosion of Protein Activation States

- Tumor suppressor protein p53: 27 phosphorylation sites
- Up to $2^{27} = 34,217,728$ distinguishable activation states
- Each state: individual constituent of reaction network



Cell Signalling Module

Characteristics

- Intracellular reaction network acting as functional unit
- Composed of proteins carrying phosphorylation sites
- Interactions between individual activation states

Facts

- Dynamical behaviour essential to understand function
- Often partially unknown
- Reconstruction as challenging task in systems biology
- Reverse engineering by integrative approach

Idea to manage complexity:
Capturing each protein by a specific string-object
instead of separate species per activation state

FORSYS Partner Initiative

Specification of String-Objects

Assuming two alphabets: V (for protein names), V' (for protein properties); w.l.o.g $\#, \neg, * \notin V \cup V'$

Syntax for string-objects by regular set

$$S = V^+ \cdot \left(\{\#\} \cdot ((V')^+ \cup \{\neg\} \cdot (V')^+ \cup \{*\})\right)^*$$

Protein properties

- x: property x present (e.g. specific phosphate attached)
- $\neg x$: property x absent (e.g. specific phosphate removed)
- $*$: placeholder for arbitrary property setting

Examples

- $\mathrm{prot1\#p\#} * \#\neg\mathrm{p}$ (subsumes activation states of prot1)
- $\mathrm{KaiC\#\neg KaiA\#KaiB\#4}$ (prot. complex, 4 ligands attached)

$\Longrightarrow$ **Application of reaction rules requires string matching**

Π_{CSM}: System Components

Let $\langle S \rangle$ be the set of all multisets over S.

$$\Pi_{CSM} = (V, V', R_1, \ldots, R_r, \mathrm{f}_1, \ldots, \mathrm{f}_r, A, \mathrm{C}, \Delta\tau)$$

with

$R_i \in \langle S \rangle \times \langle S \rangle$ ..... is a reaction rule composed of two finite multisets

$\mathrm{f}_i : \langle S \rangle \longrightarrow \mathbb{N}$ ..... is a function corresponding to discrete kinetics of reaction R_i

$A \in \langle S \rangle$ ..... is a multiset of axioms representing the initial molecular configuration

$\mathrm{C} \in \mathbb{R}_+$ spatial capacity of the module (vessel or compartment)

$\Delta\tau \in \mathbb{R}_+$ ..... time discretisation interval

Π_{CSM}: Matching

Let S be a string-object syntax. Two string-objects match iff there is at least one common wild card-free representation:

$$
\begin{aligned}
\mathit{Match} &\subseteq S \times S \\
\mathit{Match} &= \bigcup_{m \in \mathbb{N}} \{(p\#p_1\#p_2 \ldots \#p_m,\ s\#s_1\#s_2 \ldots \#s_m) \mid (p = s) \wedge \\
&\qquad \forall j \in \{1, \ldots, m\} : [(p_j = s_j) \vee (p_j = *) \vee (s_j = *) \vee \\
&\qquad\qquad ((p_j = \neg q) \wedge (s_j \neq q)) \vee ((s_j = \neg q) \wedge (p_j \neq q))]\}
\end{aligned}
$$

- *Match* is a symmetric relation
- Requires minimal similarity between string-objects with incomplete information
- Uncertainty interpreted as arbitrary replacement by available properties

FORSYS Partner Initiative

Π_{CSM}: Matching

Let S be a string-object syntax. Two string-objects match iff there is at least one common wild card-free representation:

Example

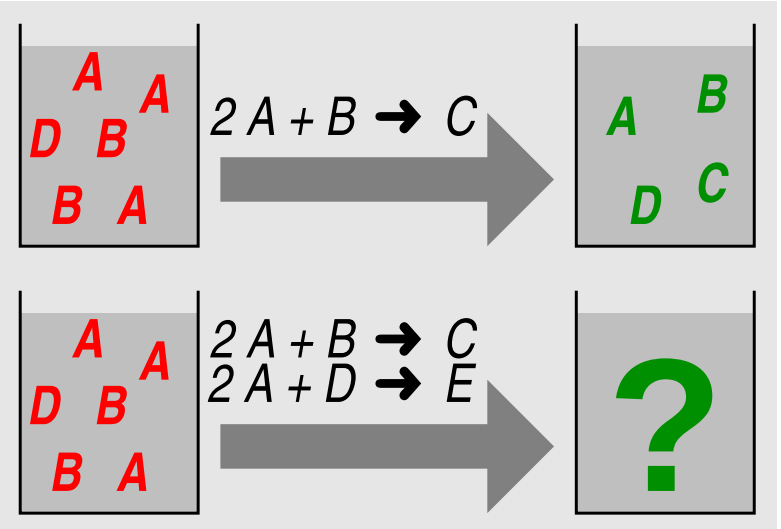
C#D#p	C#D#−p	C#T#p	C#T#−p	C#*#p	C#D#*	C#*#*	E#*#*	
■				■	■	■		C#D#p
	■				■	■		C#D#−p
		■		■		■		C#T#p
			■			■		C#T#−p
■		■		■	■	■		C#*#p
■	■			■	■	■		C#D#*
■	■	■	■	■	■	■		C#*#*
							■	E#*#*

V = {C, E}

V' = {D, T, p}

Π_{CSM}: Dynamical System Behaviour (I)

- Successive progression of configuration $L_t \in \langle S \rangle$ over time $t \in \mathbb{N}$ starting from axioms A
- $\Delta\tau$: span between t and $t+1$
- Conflict handling by prioritisation of reaction rules



$$L_0 = L_{0,0} = A$$

$$L_{t,1} = \begin{cases} L_{t,0} \ominus Reactants_{t,1} \uplus Products_{t,1} & \text{if } Reactants_{t,1} \subseteq L_{t,0} \\ L_{t,0} & \text{otherwise} \end{cases}$$

$$\vdots$$

$$L_{t+1} = L_{t,r} = \begin{cases} L_{t,r-1} \ominus Reactants_{t,r} \uplus Products_{t,r} & \text{if } Reactants_{t,r} \subseteq L_{t,r-1} \\ L_{t,r-1} & \text{otherwise} \end{cases}$$

FORSYS Partner Initiative

Π_{CSM}: Dynamical System Behaviour (II)

Estimation of multisets $Reactants_{t,j}$ and $Products_{t,j}$ at time t concerning reaction $R_j = (A_j, B_j) \in \langle S \rangle \times \langle S \rangle$ denoted

$$A_j(a_1)\ a_1 + \ldots + A_j(a_p)\ a_p \longrightarrow B_j(b_1)\ b_1 + \ldots + B_j(b_q)\ b_q$$

includes

- Matching between string-objects in L_t and those in A_j
- Consideration of stoichiometry captured by multisets A_j, B_j
- Evaluation of kinetic law expressed by scalar function f_j

$$Reactants_{t,j} = \biguplus_{e_1 \in Match(a_1)} \ldots \biguplus_{e_p \in Match(a_p)} f_j\left(\{(e_1,\infty),\ldots,(e_p,\infty)\} \cap L_{t,j-1}\right) \cdot \{(e_1, A_j(a_1)),\ldots,(e_p, A_j(a_p))\}$$

$$Products_{t,j} = \biguplus_{e_1 \in Match(a_1)} \ldots \biguplus_{e_p \in Match(a_p)} f_j\left(\{(e_1,\infty),\ldots,(e_p,\infty)\} \cap L_{t,j-1}\right) \cdot \{(b_1, B_j(b_1)),\ldots,(b_q, B_j(b_q))\}$$

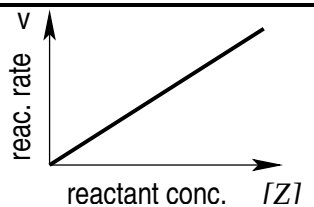
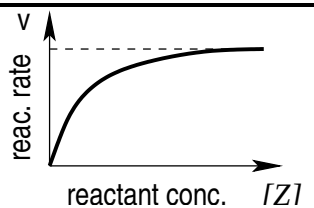
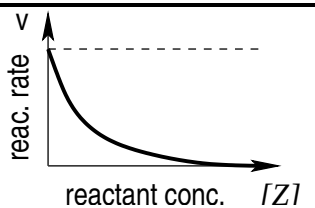
FORSYS Partner Initiative

Π_{CSM}: Discrete Reaction Kinetics

Scalar function f_j provides number of turns for application of reaction rule R_j. Rate constant: $k_j = \hat{k}_j \cdot C \cdot \Delta\tau$ (Euler).

$$f_j(L_t) = \left\lfloor k_j \prod_{\forall \alpha \in Match(A_j) \cap Match(L_t) : (R_j = (A_j, B_j))} \hat{f}(L_t(\alpha))^{|Match(A_j) \cap \{(\alpha, \infty)\}|} \right\rfloor$$

Selected kinetic laws $\hat{f}([Z])$

Kinetics	Activation	Repression
Mass-Action (no saturation)	 $\hat{f}([Z]) = [Z]$	–
Michaelis-Menten (saturation)	 $\hat{f}([Z]) = \frac{[Z]}{\Theta + [Z]}$	 $\hat{f}([Z]) = \left(1 - \frac{[Z]}{\Theta + [Z]}\right)$

Circadian Clocks

Characteristics

- Self-sustained biochemical oscillators
- Period of approx. 24 hours persisting under constant environmental conditions (e.g. constant darkness)
- Temperature compensation within physiological range
- Capability of entrainment by external stimuli (e.g. light/dark or temperature cycles)
- Reaction system with at least one feedback loop

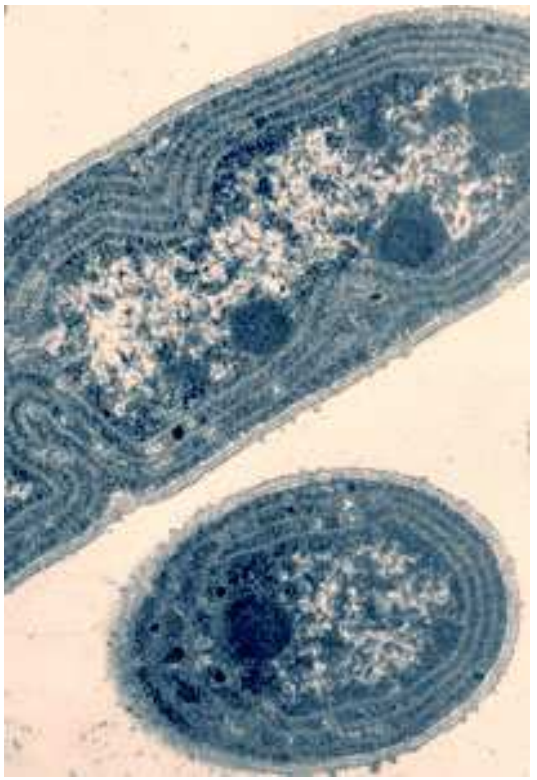
High scientific impact because . . .

- Circadian clock as a potential universal property of life
- Self-sustainability and high precision of bio-oscillators
- Chronobiological control systems for manifold processes
- Several independent evolutionary origins assumed

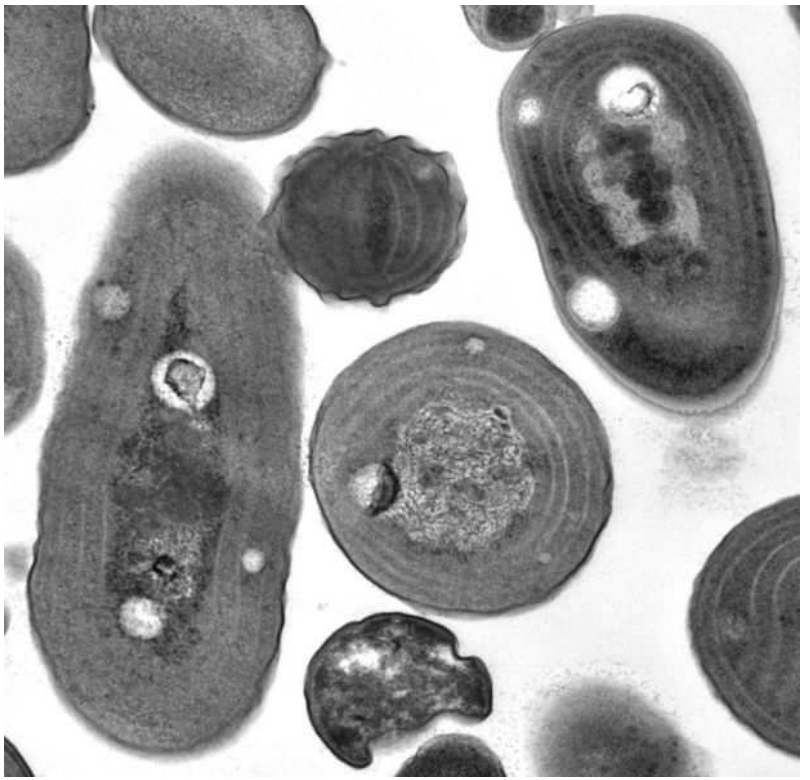
FORSYS Partner Initiative

Cyanobacterium *Synechococcus elongatus*

"Simplest cells known to exhibit circadian phenomena"



www.genome.jgi–psf.org



1µm www.wikipedia.org

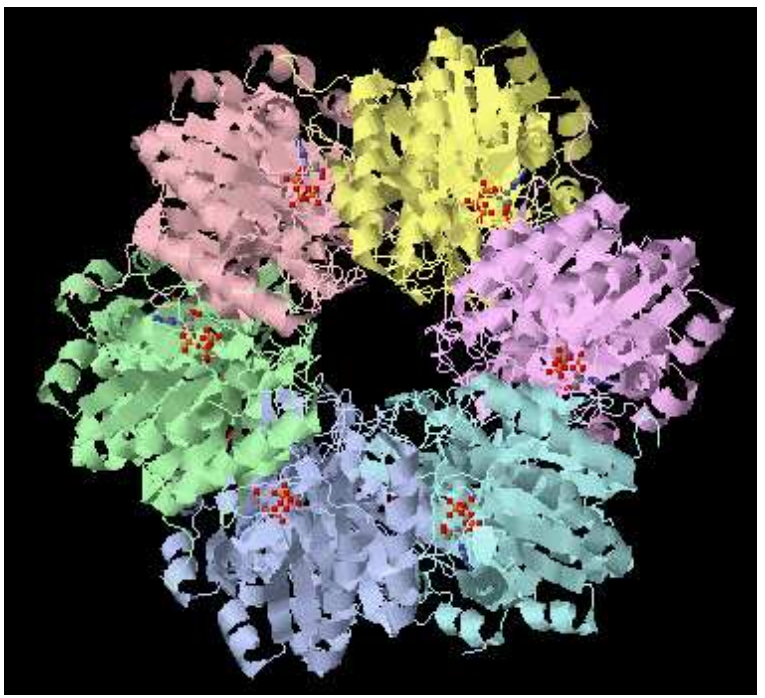
Prokaryotic autotrophic picoplankton in tropical oceans

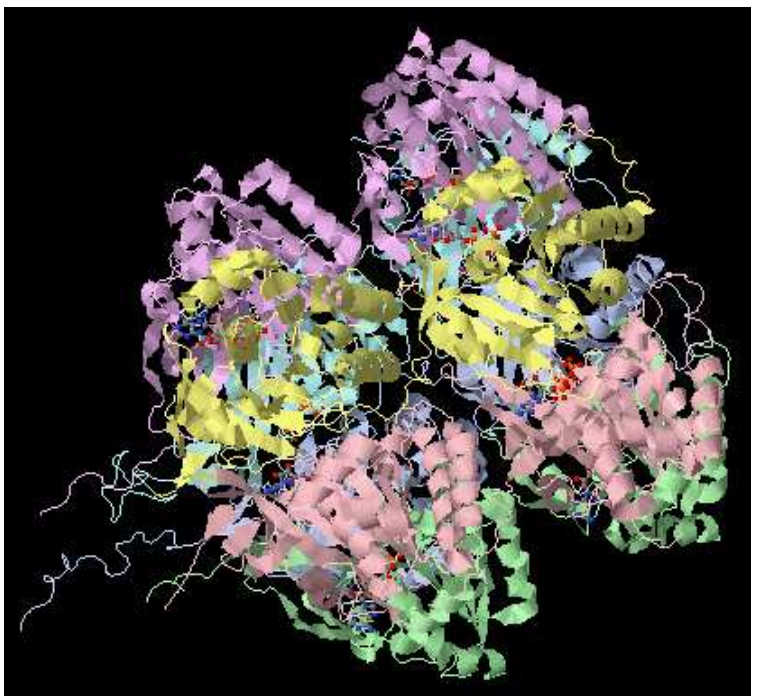
Genome: 2.4 . . . 2.7 Mbp

FORSYS Partner Initiative

Components of Circadian Clock: Key Protein KaiC

- Homohexamer ("double doughnut") with 12 ATP molecules
- Protein kinase (transferase), length: 519 residues



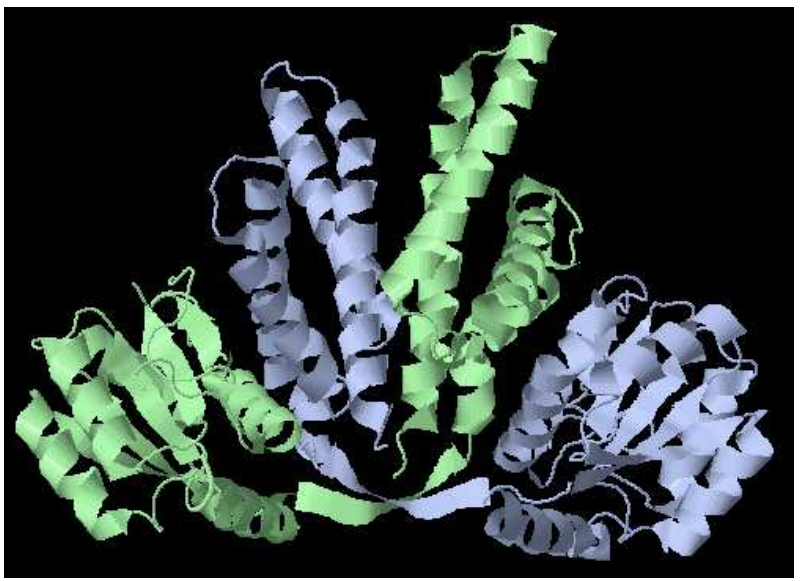


PDB Protein Data Bank, ID: 2gbl, www.rcsb.org/pdb

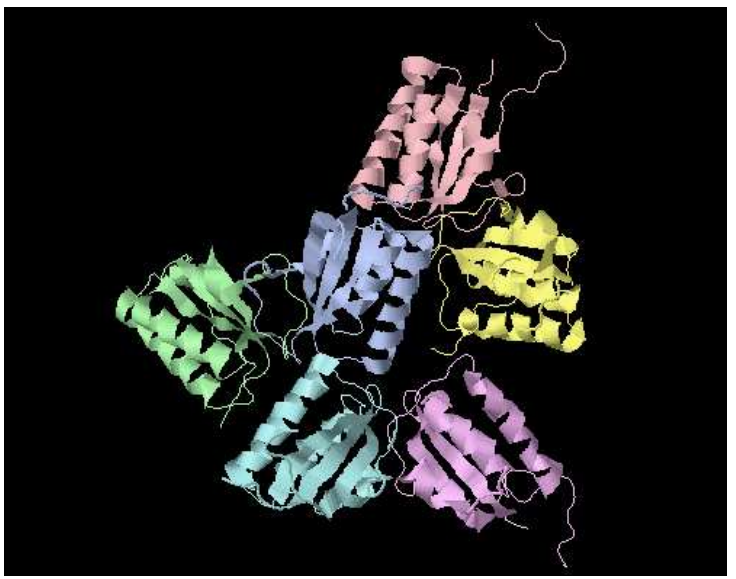
FORSYS Partner Initiative

Key Clock Proteins KaiA and KaiB

- KaiA: protein binding molecular function reported, length: 289 residues
- KaiB: no further molecular function reported, length: 108 residues



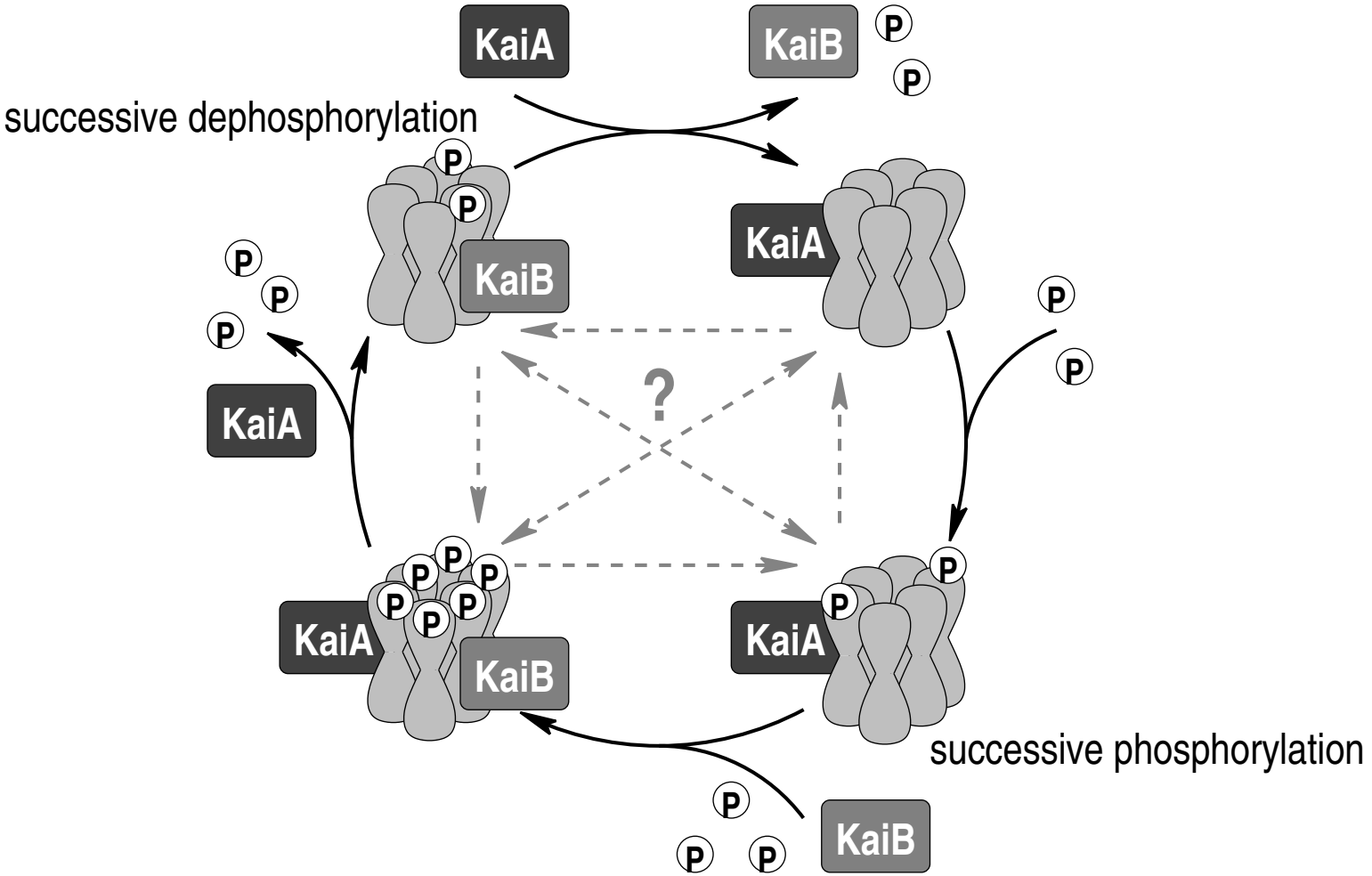
KaiA protein structure from PDB Protein Data Bank, ID: 1r8j
www.rcsb.org/pdb



KaiB protein structure from PDB Protein Data Bank, ID: 2qke

FORSYS Partner Initiative

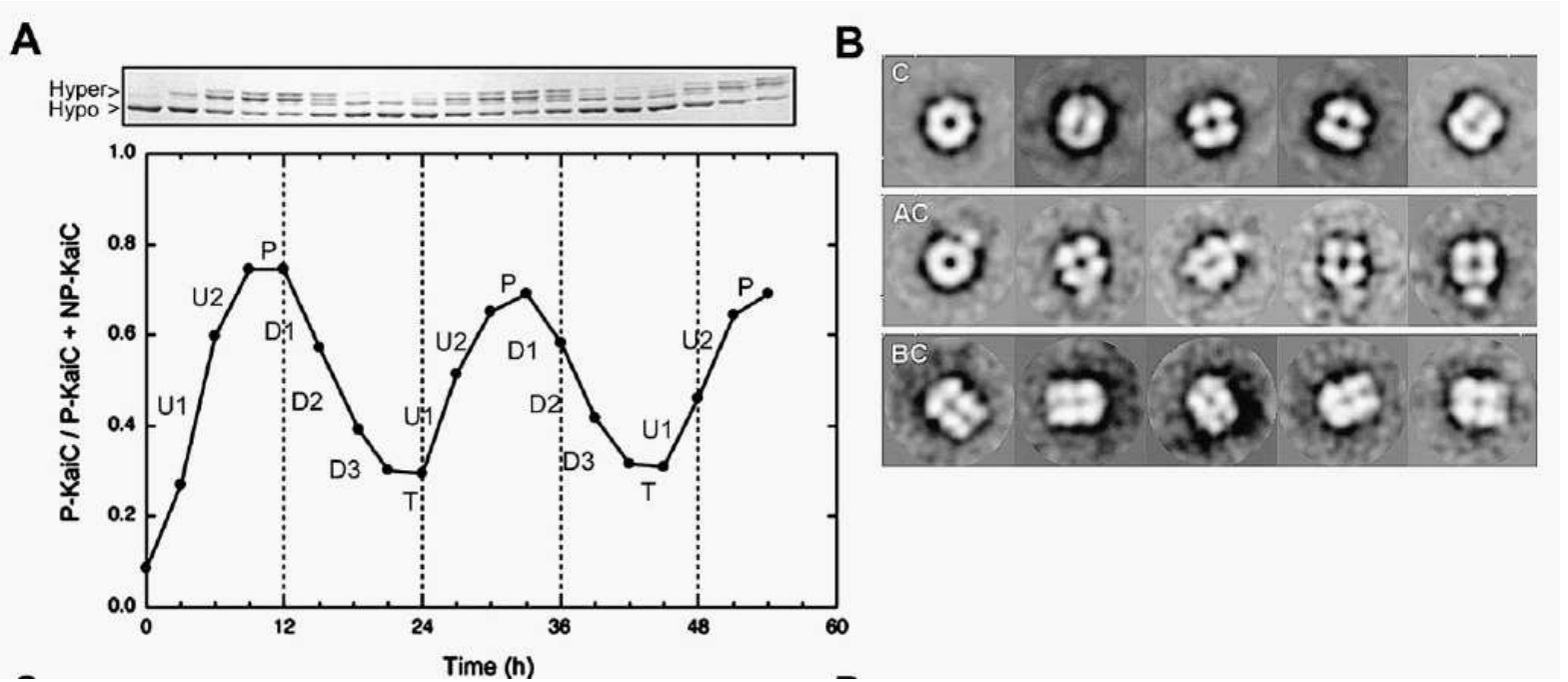
KaiABC Oscillator: Reaction Cycle



Incomplete information about interphase feedback loops

FORSYS Partner Initiative

KaiC Oscillating Behaviour in Seven Phases



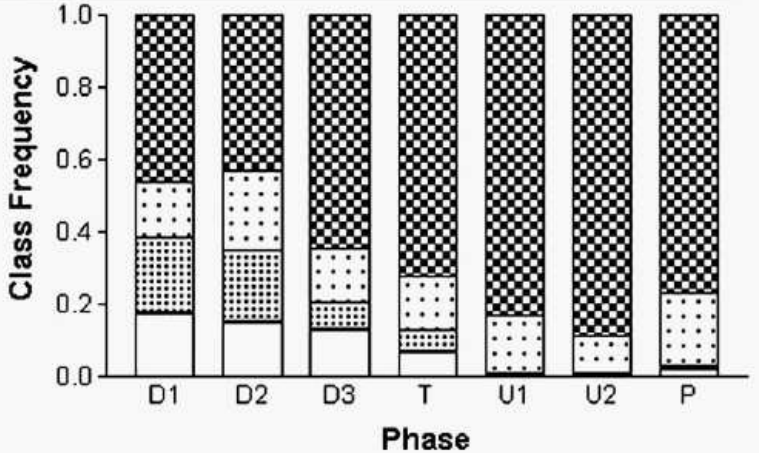
(A) PAGE gel and oscillation phases: U1, U2 (upward), P (peak), D1, D2, D3 (descent), T (trough)
(B) Representative electron microscopic images of KaiC (C) and complexes KaiA●KaiC (AC), KaiB●KaiC (BC)

T. Mori, D.R. Williams, M.O. Byrne, X. Qin, M. Egli, H.S. Mchaourab, P.L. Stewart, C.H. Johnson. Elucidating the Ticking of an In Vitro Circadian Clockwork. PLoS Biology **5(4)**:841–853, 2007, doi: 10.1371/journal.pbio.0050093

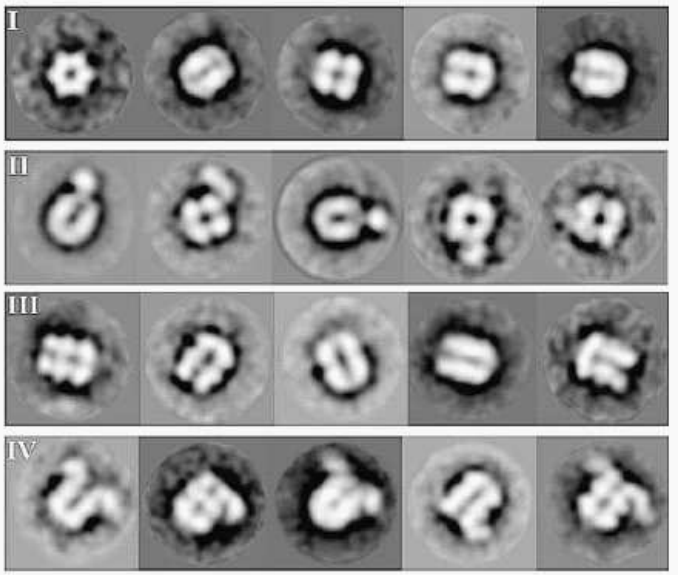
Relative Frequencies of KaiC and Complexes in Phases D1, D2, D3, T, U1, U2, P

C

	D1	D2	D3	T	U1	U2	P
I	0.43	0.40	0.52	0.69	0.83	0.90	0.76
II	0.16	0.23	0.20	0.15	0.16	0.10	0.20
III	0.22	0.20	0.12	0.07	0.00	0.00	0.01
IV	0.19	0.18	0.16	0.09	0.01	0.01	0.04



D



(C) Assignment of frequency classes
I KaiC hexamers alone, **II** KaiA●KaiC, **III** KaiB●KaiC, **IV** KaiA●KaiB●KaiC

(D) Representative electron microscopic images of classes

T. Mori, D.R. Williams, M.O. Byrne, X. Qin, M. Egli, H.S. Mchaourab, P.L. Stewart, C.H. Johnson. Elucidating the Ticking of an In Vitro Circadian Clockwork. PLoS Biology **5(4)**:841–853, 2007, doi: 10.1371/journal.pbio.0050093

FORSYS Partner Initiative

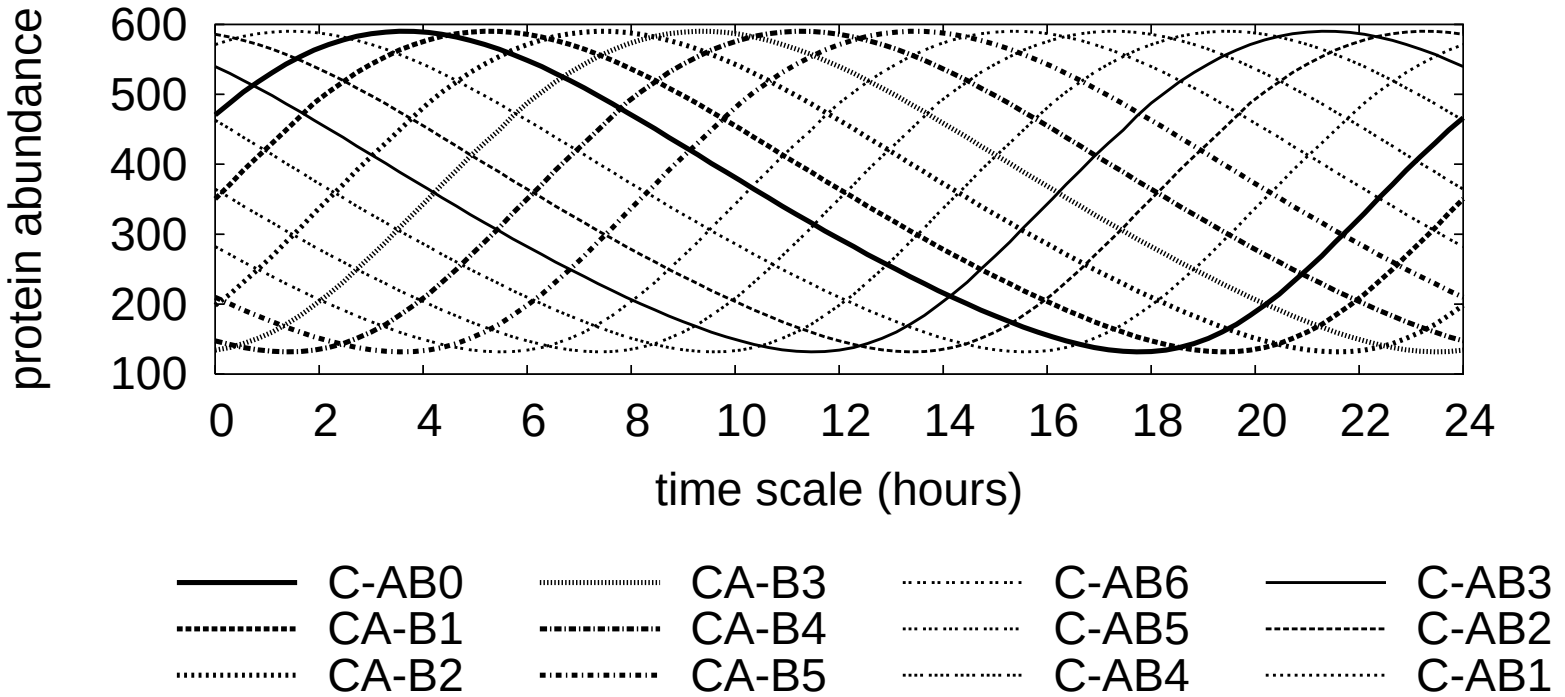
The Model Π_{KaiABC} at a Glance

$$
\begin{aligned}
\Pi_{KaiABC} &= (V, V', R_1, \ldots, R_{17}, \mathrm{f}_1, \ldots, \mathrm{f}_{17}, A, \mathrm{C}, \Delta\tau) \\
V &= \{\mathrm{A}, \mathrm{B}, \mathrm{C}\} \ldots\ldots\ldots\ldots\ldots\ldots \text{identifiers of proteins KaiA, KaiB, KaiC} \\
V' &= \{\mathrm{A}, \mathrm{B}\} \cup \ldots\ldots\ldots\ldots\ldots \text{KaiA, KaiB within a complex associated to KaiC} \\
&\quad \{0, 1, 2, 3, 4, 5, 6\} \ldots\ldots \text{number of attached phosphates} \\
R_1 &= \mathrm{C}\#\neg\mathrm{A}\#\mathrm{B}\#0 + \mathrm{A} \longrightarrow \mathrm{C}\#\mathrm{A}\#\neg\mathrm{B}\#1 + \mathrm{B} \\
R_2 &= \mathrm{C}\#\mathrm{A}\#*\#1 + \mathrm{A} \longrightarrow \mathrm{C}\#\mathrm{A}\#*\#2 + \mathrm{A} \\
R_3 &= \mathrm{C}\#\mathrm{A}\#*\#2 + \mathrm{A} \longrightarrow \mathrm{C}\#\mathrm{A}\#*\#3 + \mathrm{A} \\
R_4 &= \mathrm{C}\#\mathrm{A}\#*\#3 + \mathrm{A} \longrightarrow \mathrm{C}\#\mathrm{A}\#*\#4 + \mathrm{A} \\
R_5 &= \mathrm{C}\#\mathrm{A}\#*\#4 + \mathrm{A} \longrightarrow \mathrm{C}\#\mathrm{A}\#*\#5 + \mathrm{A} \\
R_6 &= \mathrm{C}\#\mathrm{A}\#\neg\mathrm{B}\#5 + \mathrm{B} \longrightarrow \mathrm{C}\#\neg\mathrm{A}\#\mathrm{B}\#6 + \mathrm{A} \\
R_7 &= \mathrm{C}\#*\#\mathrm{B}\#6 + \mathrm{B} \longrightarrow \mathrm{C}\#*\#\mathrm{B}\#5 + \mathrm{B} \\
R_8 &= \mathrm{C}\#*\#\mathrm{B}\#5 + \mathrm{B} \longrightarrow \mathrm{C}\#*\#\mathrm{B}\#4 + \mathrm{B} \\
R_9 &= \mathrm{C}\#*\#\mathrm{B}\#4 + \mathrm{B} \longrightarrow \mathrm{C}\#*\#\mathrm{B}\#3 + \mathrm{B} \\
R_{10} &= \mathrm{C}\#*\#\mathrm{B}\#3 + \mathrm{B} \longrightarrow \mathrm{C}\#*\#\mathrm{B}\#2 + \mathrm{B} \\
R_{11} &= \mathrm{C}\#*\#\mathrm{B}\#2 + \mathrm{B} \longrightarrow \mathrm{C}\#*\#\mathrm{B}\#1 + \mathrm{B} \\
R_{12} &= \mathrm{C}\#*\#\mathrm{B}\#1 + \mathrm{B} \longrightarrow \mathrm{C}\#*\#\mathrm{B}\#0 + \mathrm{B} \\
R_{13} &= \mathrm{C}\#\neg\mathrm{A}\#\mathrm{B}\#* + \mathrm{A} \longrightarrow \mathrm{C}\#\mathrm{A}\#\neg\mathrm{B}\#* + \mathrm{B} \\
R_{14} &= \mathrm{C}\#\mathrm{A}\#\neg\mathrm{B}\#* + \mathrm{B} \longrightarrow \mathrm{C}\#\neg\mathrm{A}\#\mathrm{B}\#* + \mathrm{A} \\
R_{15} &= \mathrm{A} \longrightarrow \emptyset \\
R_{16} &= \mathrm{B} \longrightarrow \emptyset \\
R_{17} &= \mathrm{C}\#*\#*\#* \longrightarrow \emptyset \\
&\ldots
\end{aligned}
$$

Discrete Michaelis-Menten kinetics

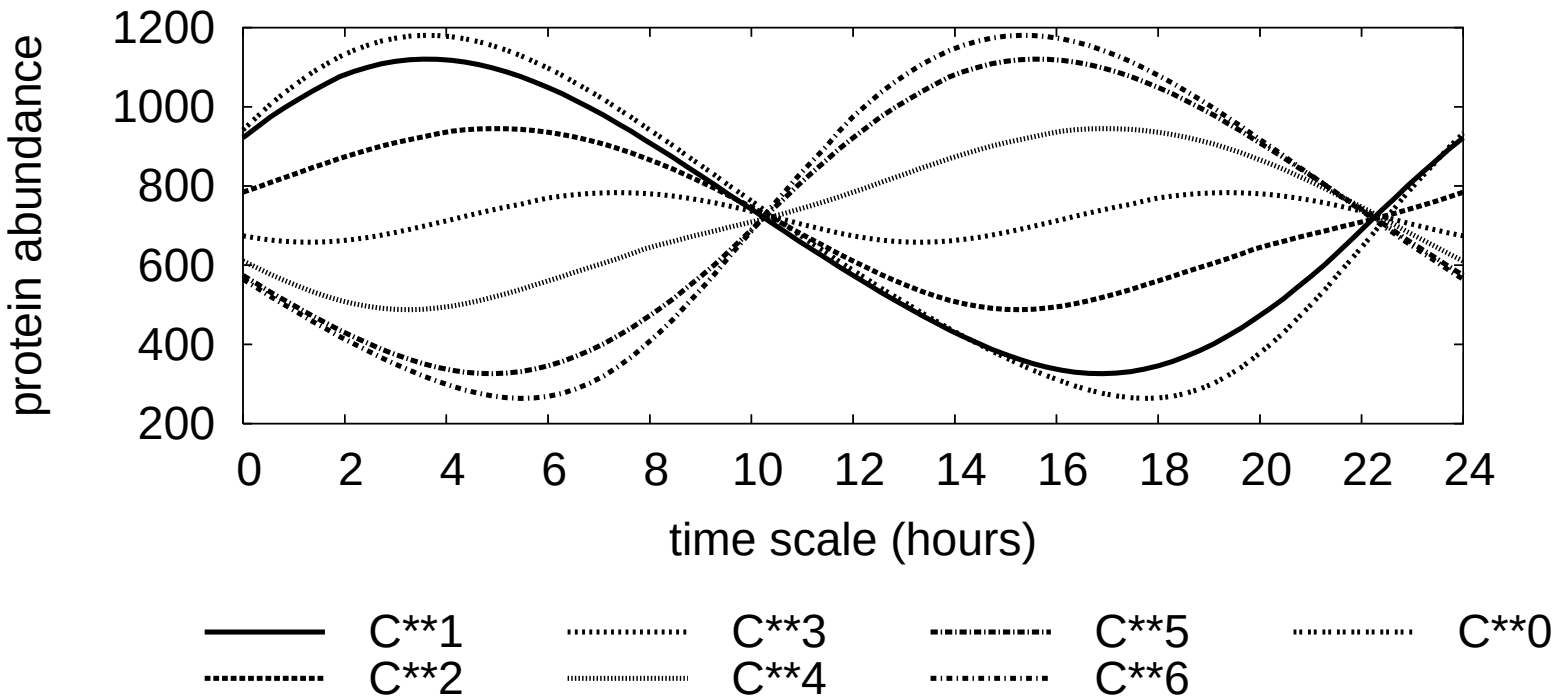
FORSYS Partner Initiative

Simulation Results: Individual KaiABC Subproducts



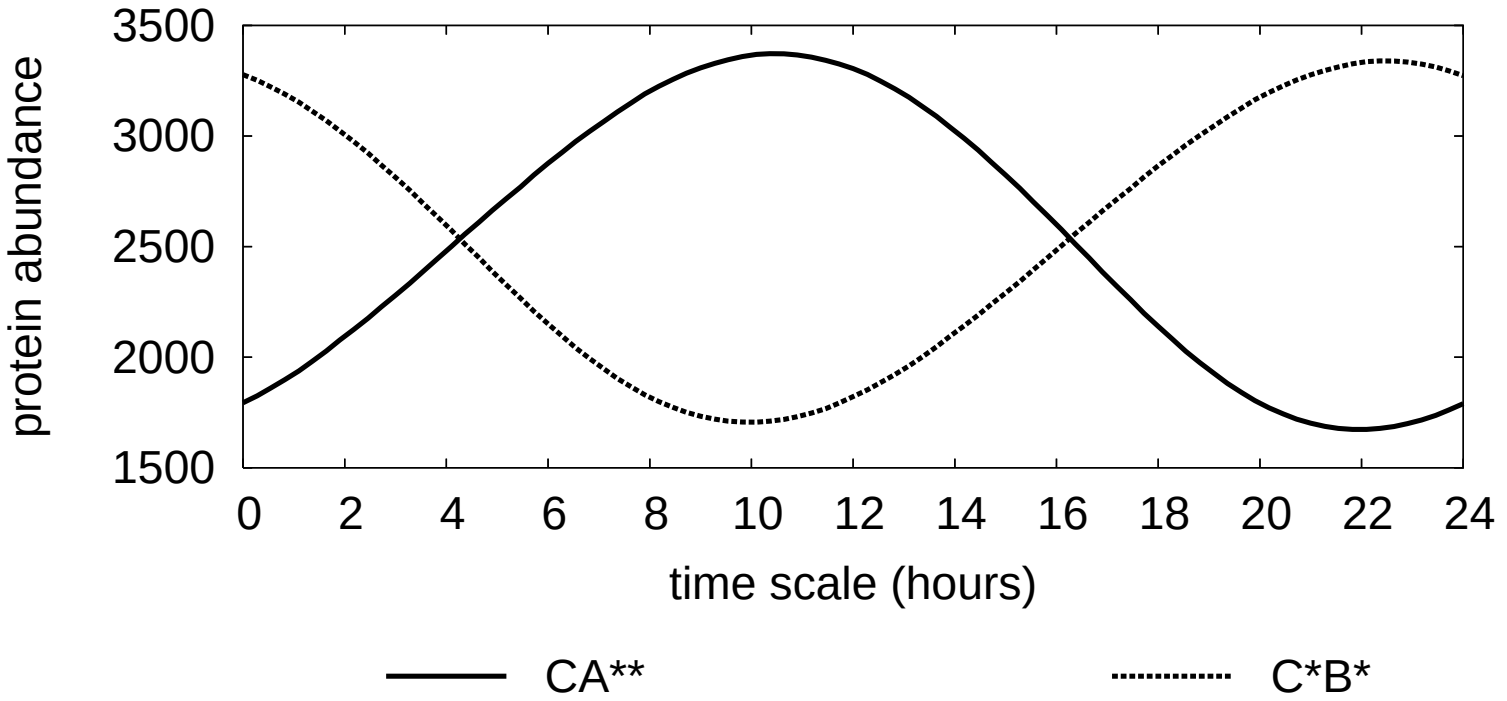
Temporal courses of 12 specific KaiABC subproducts representing the process status of the reaction cycle. Kinetic parameters and initial amounts adjusted in a way to obtain a period of $\approx$ 24 hours and symmetry among individual oscillations.

Focussing on the Level of Phosphorylation



Temporal courses of KaiABC subproducts subsumed by their level of phosphorylation ranging from 0 to 6. Kinetic parameters and initial amounts adjusted in a way to obtain a period of $\approx$ 24 hours and symmetry among individual oscillations.

Focussing on KaiABC Complex Formation



Temporal courses of KaiABC subproducts separated into two groups by association of KaiA resp. KaiB to KaiC. Kinetic parameters and initial amounts adjusted in a way to obtain a period of $\approx$ 24 hours and symmetry among individual oscillations.

Outlook

Conclusions

- String-objects denoted by regular expressions can manage descriptional complexity of protein binding states
- Coping with incomplete information by means of integrative approach
- Practicability for application in systems biology

Further work

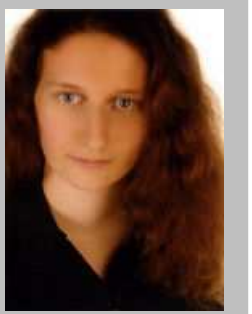
- Extending the system by symmetry-breaking assumptions
  - Premature association/dissociation
  - Spontaneous dephosphorylation
  - Monomer shuffle, . . .
- Integration of temperature within the model
- Perturbance analysis of the model

FORSYS Partner Initiative

Special Thanks go to . . .

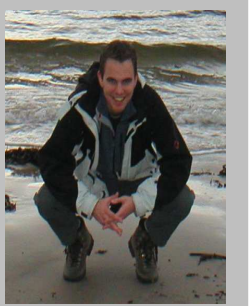
... my coworkers

Gabi Escuela
Bio Systems Analysis Group, FSU Jena



Ines Heiland
Department Bioinformatics, FSU Jena

Thorsten Lenser
Bio Systems Analysis Group, FSU Jena



Stefan Schuster
Department Bioinformatics, FSU Jena

... the funding organization

German Federal Ministry of Education and Research, project 0315260A within Research Initiative in Systems Biology

Bundesministerium für Bildung und Forschung

FORSYS Partner Initiative

... you for your attention. Questions?