Hierarchy of models: from qualitative to quantitative analysis of circadian rhythms in cyanobacteria

Madalena Chaves*   Miguel Preto†

Abstract

A hierarchy of models, ranging from high to lower levels of abstraction, is proposed to construct "minimal" but predictive and explanatory models of biological systems. Three hierarchical levels will be considered: Boolean networks, piecewise affine differential (PWA) equations and a class of continuous, ordinary differential equations' models derived from the PWA model. This hierarchy provides different levels of approximation of the biological system and, crucially, allows the use of theoretical tools to more exactly analyze and understand the mechanisms of the system. The Kai ABC oscillator, which is at the core of the cyanobacterial circadian rhythm, is analyzed as a case study, showing how several fundamental properties—order of oscillations, synchronization when mixing oscillating samples, structural robustness, and entrainment by external cues—can be obtained from basic mechanisms.

Author summary

The dynamical behavior of biological systems can be modeled using different mathematical formalisms, each providing a different level of detail. In this work, we use several formalisms to model the Kai ABC oscillator, which is at the core of the cyanobacterial circadian clock. We start with a logical model, that contains the qualitative information available on the system. Building upon this structure, we then develop quantitative "minimal" models—with the smallest number of parameters and state variables—that are able to describe the dynamical properties observed in experiments. The resulting model successfully reproduces many of the defining properties of circadian rhythms, including the entrainment by the "environment", showing that it can synchronize with daily cycles, and it is robust to perturbations in the phase of oscillation.

1 Introduction

A wide range of modeling formalisms is available to describe, with different levels of detail and emphasis, the dynamical behavior of biological systems[7]. In this work, we will illustrate the use of different mathematical formalisms to construct a hierarchy of "minimal" models ranging from high to lower levels of abstraction. The goal is to construct models that not only reproduce the essential properties of a given biological system, but are also more amenable to rigorous theoretical analysis, and whose results at each step can thus provide guidance in the construction of successively more realistic, detailed models. These models are "minimal" in the sense that they aim to identify the basic structure of interactions—including simplified representations of production and degradation, activation and inhibition, conversion or transformation terms— and which are, moreover, characterized by a "minimum" number of parameters and of state variables. They do not intend to provide a realistic description of every reaction step of the system.

Three hierarchical levels will be considered: at a high abstraction level, Boolean networks[33, 34, 37] represent the topology of interactions and describe the logic inherent to the system by assuming that each species is either present or absent—e.g., a protein that is either expressed or not expressed, active

---

*BIOCORE, INRIA, 2004 Route des Lucioles, BP 93, 06902 Sophia Antipolis, France, `madalena.chaves@inria.fr`
†Astronomisches Rechen-Institut, Zentrum für Astronomie, Univ. of Heidelberg, Mönchhof-Strasse 12-14, 69120 Heidelberg, Germany, `miguelp@ari.uni-heidelberg.de`.

or inactive (see Section IIIA). At an intermediate abstraction level, piecewise affine differential (PWA) equations (such as Glass-type models[14, 10, 8, 5]) describe production rates as logical functions, where a species is either not produced, or produced at a constant rate, and degradation rates are treated as being continuous and linear (see Section IIIB). Finally, at a lower abstraction level, we consider continuous, ordinary differential equations' (ODEs) models, which retain the same set of parameters and logic of interactions as the PWA models, but where the constant production rates of the latter are substituted by Hill functions, and for which the PWA model constitutes a limit case (when the Hill coefficient $n \to +\infty$) and is often a reasonable approximation (see Section IV). The Hill function model can be used for a quantitative comparison with experimental data and for parameter estimation (Section V). The advantage of PWA over ODE models is that, by being explicitly integrable in separate regions of its phase space, they are more amenable to rigorous derivation of results, albeit at a cost of severely restricting the topology of its vector fields. Nevertheless, PWA models have revealed themselves very useful in providing insight for the analysis of more general Hill funtion ODE models. Armed with these tools we start to unveil and analyse the properties emerging from these series of models, namely: the existence and stability of steady state or oscillatory solutions, response to environmental perturbations, robustness against intrinsic or extrinsic noise, and so on (Section VI).

We illustrate this methodology through the analysis of a system composed of three proteins—Kai A, Kai B and Kai C—, which is at the core of the circadian clock of cyanobacteria, and for which a large amount of data is available.

2 Circadian rhythms, cyanobacteria and the Kai ABC network

During evolution organisms have developed sophisticated mechanisms enabling them to adapt their metabolism and behavior to changes in the surrounding environment. Among these, the most predictable are associated to daily cycles of light/dark and temperature arising from the rotation of the Earth about its axis. From bacteria to humans, almost all organisms developed the molecular circuitry to keep an internal representation of daily cycles—these are called circadian clocks or oscillators [38, 16, 17]. These biochemical oscillators are characterized by the following properties [9, 22]: (i) a free-running period of about 24 hours; (ii) entrainability by dark/light cycles, with the ability to reset their phase so as to keep synchronized to the exterior; (iii) a free-running period that is almost unaffected by changes in *constant* ambient temperature, a property known as *temperature compensation*, but with the capability of being entrained and reset by cycles or pulses of temperature; and (iv) at least for cyanobacteria, another requirement must be met: the oscillations must be extremely robust to fluctuations against the inherent stochasticity of biochemical reactions and frequent cell divisions— as their phases of oscillation drift very little even after weeks without external cues in constant light environments [2, 23, 1].

The prevailing notion from the mid 1990s—a time when a plethora of new clock genes were being discovered in various organisms—to 2005 was that circadian clocks are driven by a negative transcriptional feedback loop in which the clock genes produce proteins that repress their own transcription [9]. The clock genes and their proteins, in turn, regulate the circadian gene expression of many other genes. In the case of cyanobacteria, the core oscillator was thought to consist of a transcription-translation cycle comprised of just three proteins (Kai A, Kai B and Kai C) together with their genes. These proteins are also involved in a phosphorylation cycle —itself oscillating in a circadian fashion. Kondo and colleagues made a significant breakthrough by demonstrating that the circadian clock required no transcription-translation cycle: robust circadian gene expression persists in the presence of an inhibitor of transcription or translation of the kai ABC genes, or in continuous dark conditions [35]. Moreover, the phosphorylation cycle was shown to be functional in a test tube by combining only those three proteins and adenosine triphosphate (ATP) [25]. This *in vitro* oscillator was found to have all the properties of a *bona fide* circadian clock [31] .

By focusing on the Kai ABC *in vitro* oscillator, the work of Rust et al[32] and the Kondo lab[26] (among others) have unveiled a network of interactions among the proteins Kai A,B,C. The main steps[32, 26] can be briefly summarized as follows. The protein Kai C has two phosphorylation sites that phosphorylate in an ordered sequence: starting with the fully unphosphosrylated form U-Kai C, this is phosphorylated first on the site T432, converting into T-Kai C; then, this is phosphorylated at the

site S431, leading to the doubly phosphorylated form TS-Kai C; finally, TS-Kai C is desphosphorylated at the site T432, leading to accumulation of the form S-Kai C. In the absence of Kai A, Kai C acts essentially as an autophosphatase: incubated alone without Kai A, it dephosphorylates. Free Kai A promotes the phosphorylation of Kai C, but its activity is supressed when it forms a bound complex with Kai B and S-KaiC. Kai B seems to have no direct role in the (de)phosphorylation of Kai C. The dephosphorylation of TS-Kai C into S-Kai C seems to be weakly inhibited by Kai A. This leads to further dephosphorylation of S-Kai C, which then transforms into the fully unphosphorylated U-Kai C while, at the same time, liberating Kai A from the bound complex. The newly freed Kai A can resume its activity of promoting Kai C phosphorylation (U-Kai C → T-Kai C and T-Kai C → TS-Kai C), and a new cycle begins (see Fig. 2). In sum, it is the feedback loop provided by the Kai A ability of switching its activity, between a phase during which it promotes Kai C phosphorylation and a phase when it is inactive in a bound complex (with Kai B and S-Kai C), that serves as modulator of the phosphorylation activity of Kai C and thus drives the phosphorylation cycle oscillator.

However, there is experimental evidence that the clock retains its circadian periodicity even in strains that overexpress Kai A, where the phosphorylation cycle oscillation has been arrested [21]. Whether the transcription-translation cycle oscillation is self-sustained or damped remains to be determined [30]. The discovery of the *in vitro* Kai ABC oscillator, recent results on the circadian rhythmicity in human red blood cells that lack a nucleus [28], among other studies highlighting the role of non-genetic cogs and wheels, lead to a general question: what is the relation between the genetic and post-translational circuitry in driving the circadian clock? Do clock genes play the role of the circadian pacemaker; or, is it driven by post-translational circuits instead? Alternatively, do circadian clocks emerge from the coupling of autonomous genetic and post-translational oscillators complementing each other in a non-hierarchical manner?

In this paper, we take a step in the direction of tackling the relation between the genetic and post-translational components of the circadian circuitry. We focus on constructing a model of the Kai ABC oscillator which is "minimal", but still capable of capturing its defining properties, and that could constitute a module in the subsequent modeling steps.
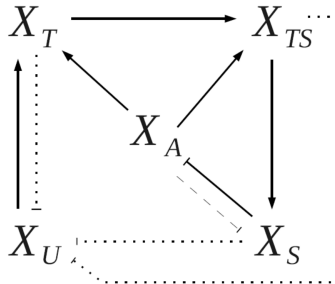

Figure 1: A schematic view of the Kai ABC system. X_A denotes the protein Kai A and X_U denotes the unphosphorylated form of protein Kai C. X_T (resp., X_S) denotes the form T-Kai C (resp., S-Kai C), while X_{TS} denotes the form TS-Kai C. Arrow (resp., bar)-headed arcs represent a positive (resp., negative) effect from the source to the target species. The dashed arc represents a weaker interaction than those represented by solid arcs. The dotted arcs represent the (Boolean) negative effect on X_U due to the conservation of total protein Kai C (that is, $X_U = X_{Tot} - X_T - X_{TS} - X_S$).

3 Qualitative study of the Kai ABC network

As a first step to constructing a model of the system, we will study the qualitative dynamics entailed by the network of interactions among the various species involved (Fig. 2). First, we will construct a Boolean model consisting of a logical rule for each variable, based only on the network of interactions: for instance, the rule for species X_T will say that it can be produced only if both X_U and X_A are

available, that is, protein Kai C can be phosphorylated on the T-site, only if (non-negligible quantities of) unphosphorylated protein Kai C as well as Kai A are present. Such logical models require little data, but provide a faithful representation of the qualitative dynamical behavior of the system, such as possible trajectories, the number of equilibria (or steady states), or the existence of oscillatory behavior. Second, this qualitative model will be used to construct a "hybrid" (piecewise affine) model which considers continuous concentrations, and discrete activation functions, to obtain more quantitative information (Section 3.2).

3.1 Boolean models: two qualitative levels for Kai A are needed

In Boolean models, each species can take only two values: either $X_i = 1$, meaning that species i is strongly expressed or present in non-negligible quantities in the medium; or $X_i = 0$, meaning that species i is weakly expressed or present in negligible quantities.

More precisely, a Boolean model is composed of a set of variables (or "nodes") $X = (X_1, \ldots, X_n)'$, with state space $\Omega = \{0,1\}^n$, and a set of logical rules, $g_i : \Omega \to \{0,1\}$ $(i = 1, \ldots, n)$ which summarize the effect of the network on each variable in the form: $X_i^+ = g_i(X)$, where X_i^+ denotes the next state of node i given that the current state of the network is X.

To complete the model it is necessary to specify an updating algorithm to determine the dynamical evolution of the Boolean model. In general, an updating algorithm consists of a discrete set of time instants, $t_j : \ j = 0, 1, 2, \ldots$, and a set of variables to be updated at each instant (see, for instance, Wang et al[37] for a recent review on Boolean models). We will adopt the fully asynchronous algorithm (where only one variable is updated at each time instant), for its generality and more realistic interpretation. The updating scheme generates a *state transition diagram* which represents the dynamical evolution of the Boolean network, and consists of a graph where the vertices are the Boolean states and the edges represent possible transitions between states (this graph was first introduced in[13]). The *attractors* of the state transition diagram are of special interest as they correspond to the asymptotic behavior of the system: once the system enters an attractor (which may consist of one or more Boolean states), it cannot leave again.

Following the experiments of Rust et al[32], our model describes the dynamics of the three phosphorylated forms of Kai C (X_T, X_{TS}, and X_S) and the protein Kai A (X_A). In order to simplify the model, it is assumed that protein Kai B is always present in sufficiently large amounts, so that all rates will depend only on the concentrations of Kai A and Kai C. A Boolean model of the interactions in Fig. 2 (including all but the dotted edges) can be written as follows:

$$
\begin{aligned}
X_A^+ &= \texttt{not}X_S \\
X_T^+ &= X_U \texttt{ and } X_A \\
X_{TS}^+ &= X_T \texttt{ and } X_A \\
X_S^+ &= X_{TS} \texttt{ and not}X_A
\end{aligned}
\tag{1}
$$

where the "free" or unphosphorylated protein X_U is obtained by assuming that total Kai C is conserved: $X_U = X_{Tot} - X_T - X_{TS} - X_S$. This constraint may be translated into different logical rules for X_U (for instance, $X_U = \texttt{not}(X_T \texttt{ or } X_{TS} \texttt{ or } X_S)$). However, analysis of model (1) shows that its asymptotic behavior satisfies

$$
X_S = X_{TS} \texttt{ and not}X_A = X_U \texttt{ and not}X_S \texttt{ and } X_S,
$$

independently of the logical rule for X_U. This implies $X_S \equiv 0$ and hence $X_A \equiv 1$ after a few iterations, since $\texttt{not}X_S \texttt{ and } X_S \equiv 0$. However, observations[32] show that the concentration of S-Kai C oscillates, which leads us to conclude that the full inhibition of X_S by X_A is too strong, and so a lower inhibition level should be considered. In other words, there should be at least two (positive) qualitative levels for Kai A: one responsible for activating X_T and X_{TS}, and the other for inhibition of X_S. There are several ways of modifying the Boolean model to account for two levels of X_A. One possibility is to say that the inhibition term of X_A on X_S can be neglected, relative to the activation of X_T and X_{TS}, that is the edge $X_A \dashv X_S$ is "weaker" than the the edges $X_A \to X_T$ or $X_A \to X_{TS}$ so we drop the former

4

to obtain:

$$
\begin{aligned}
X_A^+ &= \; \mathtt{not}\, X_S \\
X_T^+ &= \; X_U \;\; \mathtt{and} \;\; X_A \\
X_{TS}^+ &= \; X_T \;\; \mathtt{and} \;\; X_A \\
X_S^+ &= \; X_{TS},
\end{aligned}
\tag{2}
$$

where only the solid edges of Fig. 2 are considered. Using asynchronous updates, model (2) is found to globally converge to the state 1000 in the case $X_U = 0$, and has one acyclic attractor in the case $X_U = 1$ (shown in Fig. 3.1). Note that different qualitative trajectories are possible (one can identify four cycles in Fig. 3.1), but they all share the same phosphorylation sequence (light shaded states): in the presence of Kai A, first the T-site is phosphorylated and unphosphorylated Kai C is converted into T-Kai C, then the S-site follows and T-Kai C is converted into TS-Kai C, and finally the T-site is dephosphorylated to yield S-Kai C. The remainder of the cycle may happen in different ways corresponding to different depletion orders, suggesting that the Kai interaction network could support different (average) lifetimes for each of the three phosphorylated configurations, depending on experimental conditions. We conclude from the state transition diagram that the inferred logic of interaction is consistent with the properties of the *in vitro* Kai ABC oscillator.
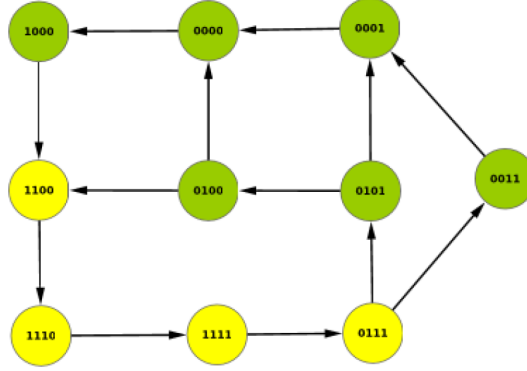

Figure 2: Attractor of asynchronous Boolean model (2), in the case $X_U = 1$. There are a total of $2^4 = 16$ states in the model but here all the transient states have been removed for clarity. The light-shaded nodes represent the states common to all possible cycles and correspond to the observed sequence of phosphorylation of protein Kai C.

3.2 Piecewise affine model of the Kai ABC oscillator

The next step in the modeling hierarchy is to consider continuous variables governed by a system of differential equations with (piecewise) constant production rates—either 0 or 1 given by the Boolean logical rule above—and a linear degradation rate. Following a construction first suggested by Glass and co-authors[14, 10], we introduce a set of continuous variables $x = (x_A, x_T, x_{TS}, x_S)'$ and a set of activity threshold constants $(\theta_A, \theta_T, \theta_{TS}, \theta_S)$ which are related to the Boolean variables by the means of step functions as follows:

$$
X_i := s^+(x_i, \theta_i) = 1 - s^-(x_i, \theta_i) = \begin{cases} 1, & x_i > \theta_i \\ 0, & x_i < \theta_i. \end{cases}
$$

The continuous variables x_i satisfy a system of piecewise affine differential equations of the following type:

$$
\dot{x}_i = k_i f_i(X) - \gamma_i x_i,
$$

where each $f_i(X)$, made of sums of products of step functions, represents a Boolean rule. More general formulations of these piecewise affine systems have been studied elsewhere[8, 3]. Following the Boolean model analysis, we include a term for the inhibition of x_S by Kai A ($s^-(x_A, \theta_{A,2})$), which can be interpreted as follows: $\theta_{A,2} > \theta_A$ is sufficiently large, so that $s^-(x_A, \theta_{A,2}) \equiv 1$ for most x_A; in other words, Kai A inhibits only weakly the production of S-Kai C. In terms of step functions one can therefore write:

$$
\begin{aligned}
\dot{x}_A &= k_A \; s^-(x_S, \theta_S) - \gamma_A x_A \\
\dot{x}_T &= k_T \; s^+(x_U, \theta_U) \; s^+(x_A, \theta_A) - \gamma_T \; x_T \\
\dot{x}_{TS} &= k_{TS} \; s^+(x_T, \theta_T) \; s^+(x_A, \theta_A) - \gamma_{TS} \; x_{TS} \\
\dot{x}_S &= k_S \; s^+(x_{TS}, \theta_{TS}) \; s^-(x_A, \theta_{A,2}) - \gamma_S \; x_S,
\end{aligned}
\tag{3}
$$

where

$$
x_U = x_{Tot} - x_T - x_{TS} - x_S,
\tag{4}
$$

that is, the total amount of protein Kai C (phosphorylated or not) is constant, as observed in test tube experiments. From a mathematical point of view this conservation of Kai C protein implies that some care is needed in the analysis; notably, one must enforce condition $x_U \geq 0$ at all times. Since the region

$$
[0, k_A/\gamma_A] \times [0, k_T/\gamma_T] \times [0, k_{TS}/\gamma_{TS}] \times [0, k_S/\gamma_S]
$$

is an invariant for system (3), a sufficient condition to guarantee $x_U \geq 0$ is given by the following inequality:

$$
\frac{k_T}{\gamma_T} + \frac{k_{TS}}{\gamma_{TS}} + \frac{k_S}{\gamma_S} \leq x_{Tot},
$$

and to consider initial conditions only inside that invariant region. From a numerical point of view, enforcing the conservation of Kai C can be easily done by including a "shutdown" term in each of the production terms for the phosphorylated forms of Kai C which stop production whenever all the available unphosphorylated C has been used: $\dot{x}_i = k_i f_i(x) s^+(x_U, 0) - \gamma_i x_i, \; i \in \{T, TS, S\}$.

The continuous state space of piecewise affine (PWA) systems can be partitioned into hyperrectangles, or regular *domains*, defined by the thresholds values. In model (3), all variables have only one threshold except for x_A, which has two (see Fig. 8 for an example of the partition of the state space). In each of the domains, the vector field is linear and solutions can be explicitly computed. However, the vector field changes as the solution moves from one domain to another, and hence global trajectories have to be formed by concatenating the solutions on each domain. At the boundaries of these regions—i.e., at the loci where one or more variables take their threshold values—the vector field is discontinuous but the PWA system can be interpreted as a differential inclusion, and global solutions are defined using Filippov-type constructions[18, 3, 12] (see also Section 4.2). If the vector fields on two adjacent domains do not point in opposite directions, then the solutions can be continued, and one says that there is a transition from one domain to the next. We can again compute a state transition diagram associated with (3) by identifying each domain with a discrete state. For instance, "1110" corresponds to $\theta_A < x_A < \theta_{A,2}$, $x_T > \theta_T$, $x_{TS} > \theta_{TS}$, and $x_S < \theta_S$. In this way, there is a natural correspondence between PWA and discrete or Boolean models[6, 4], where each domain corresponds to a vertex in a discrete model, and there is an edge connecting two vertices if there is a transition between the associated domains.

In the case $s^+(x_U, \theta_U) \equiv 0$, corresponding to the absence of unphosphorylated Kai C protein, $\dot{x}_T = -\gamma_T x_T$ for all times, and so eventually all x_i will fall below their thresholds and converge to 0. This in turn implies that, after a suitable time T_A, the equation for x_A becomes $\dot{x}_A = k_A - \gamma_A x_A$, and so $x_A(t) \to k_A/\gamma_A$.

In the case $s^+(x_U, \theta_U) \equiv 1$, corresponding to a permanent pool of unphosphorylated Kai C protein, it can be shown that no equilibria exists with either $x_S > \theta_S$ or $x_S < \theta_S$. Hence, any equilibria can only be of the form $x_i = \theta_i$, for $i \in \{A, T, TS, S\}$. Two possible qualitative solutions are obtained depending on the parameters: either a self-sustained periodic oscillation, or a damped oscillatory solution

6

which converges towards the point $(\theta_{A,j}, \theta_T, \theta_{TS}, \theta_S)$. This analysis is supported by the state transition diagram of model (3) with $s^+(x_U, \theta_U) = 1$, which coincides with that of Fig. 3.1. Further analysis of system (3) is given in Section 4.2.

4 From qualitative to quantitative models

The advantages of piecewise affine (PWA) models include the possibility of a theoretical/analytical study [8, 5, 4], as the solutions can be easily computed in each hyperrectangle. They also contain a minimum number of parameters. However, its vector fields are discontinuous, which is an unphysical feature that may lead to complex behavior around the points of discontinuity. An alternative model can be constructed using continuous sigmoidal functions in place of step functions. This will still involve a small number of parameters in comparison with detailed stoichiometric models and obviate unphysical transitions. Furthermore, this will greatly facilitate parameter estimation.

4.1 A continuous (ODE) model

The new model then becomes

$$
\begin{aligned}
\dot{x}_A &= k_A h^-(x_S, \theta_S, n) - \gamma_A x_A \\
\dot{x}_T &= k_T h^+(x_U, \theta_U, n) h^+(x_A, \theta_A, n) - \gamma_T x_T \\
\dot{x}_{TS} &= k_{TS} h^+(x_T, \theta_T, n) h^+(x_A, \theta_A, n) - \gamma_{TS} x_{TS} \\
\dot{x}_S &= k_S h^+(x_{TS}, \theta_{TS}, n) h^-(x_A, \theta_{A,2}, n) - \gamma_S x_S \\
x_U &= x_{Tot} - x_T - x_{TS} - x_S.
\end{aligned}
\tag{5}
$$

The functions h^+ and h^- represent the increasing and decreasing Hill functions:

$$
h^+(x, \theta, n) = \frac{x^n}{x^n + \theta^n}, \quad h^-(x, \theta, n) = \frac{\theta^n}{x^n + \theta^n},
$$

where n represents the Hill coefficient. Note that the PWA model is recovered in the limit $n \to +\infty$.

Before advancing any further, let us make a few considerations about the way we have chosen to model the Kai ABC oscillator. We did not aim to be realistic at the mechanistic level; instead we only require our model to be able to reproduce the essential properties of the system, in a way that can be easily controlled by manipulating its parameters. We do not try to model the oscillator at the hexamer level. We designed the model to represent the logical structure as inferred from experiments; and, in order to keep it analytically manageable and to minimize the number of its parameters, we did not try to write the detailed balance in the interconversion of Kai C phosphoforms. So, in our framework, the decay terms $-\gamma_i x_i$ are purported to represent all losses of molecular species x_i, either due to conversion into the next form (eg., $x_T \to x_{TS}$ or $x_{TS} \to x_S$) or "natural" degradation which, in view of the conservation of total Kai C, turns out to represent the conversion into the unphosphorylated form of Kai C. For instance, we did not try to balance each dephosphorylation event of, say, TS-Kai C by adding an S-Kai C. On the contrary, in order to keep the model as simple as possible, we purposefully neglect to keep such mechanistic balance. Then we investigate whether it reproduces the detailed properties of the Kai ABC oscillator. It is remarkable that even after all the simplifications, our model is able to reproduce the qualitative and quantitative properties of the Kai ABC oscillator.

4.2 Asymptotic behavior: two modes

We may reasonably expect models (3) and (5) to admit a stable periodic orbit (or "limit cycle" in 4 dimensions), for some sets of parameters, but this is difficult to prove theoretically even for the PWA model (existence of limit cycles has been proved for some particular examples, such as negative loops[15, 11]). However, independently of the existence or not of a periodic orbit, it is interesting to

note that the PWA model (3) *always* admits an asymptotic solution corresponding to a high value of x_A and low x_S. To prove this, consider the region:

$$R_A = \left\{ x: \; \theta_{A,2} < x_A \le k_A/\gamma_A, \; 0 < x_T < \frac{k_T}{\gamma_T}, \; 0 \le x_{TS} \le \frac{k_{TS}}{\gamma_{TS}}, \; x_S < \theta_S \right\}$$

and observe first that it is forward-invariant. Indeed, for all $x \in R_A$, model (3) becomes:

$$
\begin{aligned}
\dot{x}_A &= k_A - \gamma_A x_A \\
\dot{x}_T &= k_T s^+(x_U, \theta_U) - \gamma_T x_T \\
\dot{x}_{TS} &= k_{TS} s^+(x_T, \theta_T) - \gamma_{TS} x_{TS} \\
\dot{x}_S &= -\gamma_S x_S
\end{aligned}
\tag{6}
$$

which implies that $x_A(t) \to k_A/\gamma_A$, $x_S(t) \to 0$ (with monotone convergence), hence the trajectories do not leave the region R_A. In particular, this shows that, for any set of parameters, there exists an asymptotic solution of (3) where the concentrations of $x_A = k_A/\gamma_A$ and $x_S = 0$ are constant.

Lemma 1 *The system (3) with (4) and $\theta_U > \theta_T + \theta_{TS}$, has a stable steady state at*

$$x = \left(\frac{k_A}{\gamma_A}, x_{Tot} - \theta_U - \frac{k_{TS}}{\gamma_{TS}}, \frac{k_{TS}}{\gamma_{TS}}, 0 \right).$$

To prove this, note that, in region R_A, x_T and x_{TS} form a 2-dim system which becomes decoupled from x_A and x_S, as $x_A \to k_A/\gamma_A$ and $x_S \to 0$, and can therefore be separately studied. The phase plane of the decoupled system (x_T, x_{TS}) is shown in Fig. 9, in the case when the lines $x_{Tot} - x_T - x_{TS} = \theta_U$ and $x_{Tot} - x_T - x_{TS} = 0$ both intersect the domain

$$R_T = \left\{ (x_T, x_{TS}) : \theta_T < x_T < \frac{k_T}{\gamma_T}, \; \theta_{TS} < x_{TS} < \frac{k_{TS}}{\gamma_{TS}} \right\}$$

(this will be true for the nominal case resulting from the parameter set estimated below). Analysis of the decoupled system $(\dot{x}_T, \dot{x}_{TS})$ shows that trajectories converge to the domain R_T, while remaining at all times below the line $x_T + x_{TS} = x_{Tot}$ due to the constraint imposed by the conservation law. More precisely, it is easy to check that, within R_T, it holds that $\dot{x}_{TS} = k_{TS} - \gamma_{TS} x_{TS}$, so eventually $x_{TS} \to k_{TS}/\gamma_{TS}$, which is its maximal value. In contrast, the vector field of variable x_T has opposite signs on each side of the line L_T:

$$L_T = \{ (x_T, x_{TS}) : \; x_{TS} = x_{Tot} - \theta_U - x_T \}. \tag{7}$$

Since

$$
\dot{x}_T =
\begin{cases}
k_T - \gamma_T x_T, & \text{if } x_T < x_{Tot} - \theta_U - x_{TS} \\
-\gamma_T x_T, & \text{if } x_T > x_{Tot} - \theta_U - x_{TS}
\end{cases}
$$

the trajectories cannot cross L_T. However, a solution exists along this line, by defining the system as a differential inclusion and following a construction due to Filippov[12, 18] (by requiring that $\dot{x}_T + \dot{x}_{TS} = 0$, among other conditions). This type of solution is also called a "sliding" mode, as the trajectory will evolve on L_T. Therefore, as $x_{TS} \to k_{TS}/\gamma_{TS}$, x_T converges to $x_{Tot} - \theta_U - k_{TS}/\gamma_{TS}$, as stated in Lemma 1.

This analysis shows that there is no oscillatory behavior of the system in R_A, and the system converges to a steady state which is a Filippov type equilibrium point[3].

To conclude this section, it is interesting to note that these results also show that, for a given set of parameters, a steady state solution can coexist with a "limit cycle"; which one corresponds to the asymptotic behaviour will depend on the initial conditions. See Section 6.2 for further analysis.

5 Parameter estimation

A large number of *in vitro* experiments are available on the phosphorylation cycle that allow the estimation of most of the parameters of the models (3) or (5). Here, we will use the data from Rust et al.[32] who have performed non-oscillatory partial reactions experiments, in which one or more species are absent at a time, to separately reproduce the dephosphorylation and phosphorylation phases. From the experimental time series thus obtained, it is possible to infer the best-fit values for the rate constants of our model. Incidentally, one of the advantages of writing minimal models is that one minimizes a recurring issue in this type of fits—namely, that some rate constants are poorly constrained by the available data. The caveat is that no such fit is ever able to prove that a given reaction scheme holds.

The estimation procedure is an optimization problem, with the goal of minimizing the difference between the experimental data points and the corresponding predictions of the mathematical model. In other words, we wish to find a set of parameters p^* that minimizes the following cost function:

$$J(p) = \sum_{i \in \{T, TS, S, U\}} \sum_{j=1}^{L} |x_i(t_j) - x_i^{obs}(t_j)|^2,$$

with $J(p^*) = \min_p J(p)$, where $x_i^{obs}(t_j)$ are the experimental points (normalized to 100) at the L time instants t_j, and $x_i(t_j)$ are the corresponding model predictions at time t_j. The resulting nonlinear least squares problems were solved in Matlab, using the built in function `lsqnonlin`.

To reduce the number of parameters to be estimated, and simplify the procedure, we have assumed that the exponents n in all Hill functions are the same, and we have tried different values of n in different optimization runs. The results obtained for $n = 2, 4$ and 15 are shown in Table 1, together with the corresponding final cost value $J(p^*)$. Other values for n were tried ($n = 6$, not shown), but we find that the case $n = 4$ yields the best fits, as it corresponds to the lower final cost. In the analysis of the model (Section 6), we will therefore use the parameters in the column $n = 4$ of Table 1.

5.1 Dephosphorylation phase

Kai C acts essentially as an autophosphatase when it is incubated alone, in the absence of Kai A. In fact, in the absence of Kai A, it was observed by Rust et al.[32] that very little phosphorylation takes place and that the system converges to a fully unphosphorylated steady state, with $x_T = x_{TS} = x_S = 0$ and $x_U = x_{Tot}$. In our model, the dephosphorylation phase can be used to fit a subset of the rate constants if we force $x_A = 0$ at all times, which implies that both phosphorylated forms x_T and x_{TS} can only decay (into x_U and x_S, respectively). Finally, since $h^-(x_A, \theta_{A,2}, n) \equiv 1$, x_S can transiently increase (due to decay of x_{TS}), until it too eventually decays. The initial conditions of the various phosphoforms of Kai C were taken from Fig. 2B of Rust *et al.* [32]. The partial dephosphorylation reactions thus become

$$\begin{aligned}
\dot{x}_T &= -\gamma_T x_T \\
\dot{x}_{TS} &= -\gamma_{TS} x_{TS} \\
\dot{x}_S &= k_S h^+(x_{TS}, \theta_{TS}, n) - \gamma_S x_S.
\end{aligned} \tag{8}$$

This model was compared to the dephosphorylation data (Fig. 2B) of Rust et al[32], by a nonlinear least squares method as described above. The result of the fitting is shown in Fig 3, and the corresponding parameters are listed in Table 1.

5.2 Phosphorylation phase

When Kai C is incubated in the presence of a large amount of Kai A and in the absence of Kai B, the phosphorylated forms of Kai C accumulate until they reach a steady state. In fact, in the absence of Kai B, the activity of Kai A is never inhibited because it cannot be sequestred to form the complex Kai A•Kai B•S-Kai C. This is mimicked in our model by setting x_A to be a constant large enough for $h^+(x_A, \theta_A, n) \approx 1$, but not so large as to fully inhibit the formation of S-Kai C. Indeed, we choose Kai A
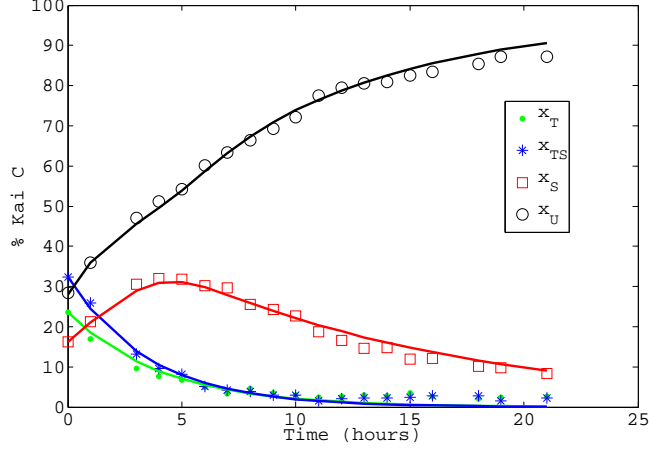
Figure 3: Fitting model (8) to experimental partial dephosphorylation reactions from Rust *et al.* [32]

in the range to satisfy the condition $h^-(x_A, \theta_{A,2}, n) = \varepsilon$, where ε is very small but different from zero. In other words, we assume that there is only a residual level of production of x_S, represented by $k_S^0 = k_S\varepsilon$. The consistency of this assumption needs to be confirmed *a posteriori* in the full simulation of the Kai ABC oscillator. This leads to the following set of equations to describe the partial phosphorylation reactions:

$$
\begin{aligned}
\dot{x}_T &= k_T h^+(x_U, \theta_U, n) - \gamma_T x_T \\
\dot{x}_{TS} &= k_{TS} h^+(x_T, \theta_T, n) - \gamma_{TS} x_{TS} \\
\dot{x}_S &= k_S^0 h^+(x_{TS}, \theta_{TS}, n) - \gamma_S x_S \\
x_U &= 100 - x_T - x_{TS} - x_S.
\end{aligned}
\tag{9}
$$

The solution of this system, starting from a fully unphosphorylated initial condition, was compared to the phosphorylation time series (Fig. 2A) of Rust et al.[32], by using the parameters already estimated from the dephosphorylation data (γ_T, γ_{TS}, γ_S, and θ_{TS}). The parameters to be estimated are thus k_T, k_S^0, θ_U, k_{TS} and θ_T. The results are listed in Table 1 and the final fit is shown in Fig. 4.
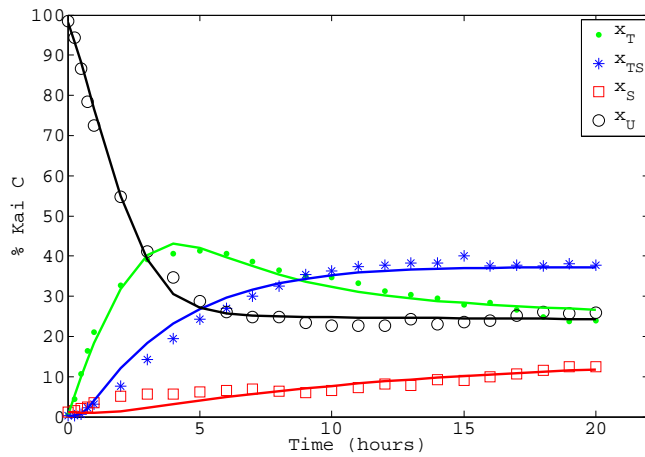

Figure 4: Fitting model (9) to experimental partial phosphorylation reactions from Rust *et al.* [32]

10

5.3 Kai A parameters

The only four parameters not estimated from data are those involved in the x_A equation, since time series for the evolution of Kai A concentration are not available in the literature. We have therefore set k_A and θ_A to values of the same order of those already estimated for k_i and θ_i, and then fine-tuned γ_A and θ_S in order to obtain a circadian period. Note that experiments indicate that the formation and dissolution of Kai A•Kai B•S-Kai C complex is slow, with a timescale similar to those of the phosphorylation and dephosphorylation terms [29]. Nevertheless, we have performed Monte Carlo simulations to assess the sensitivity of the model and its behavior to the dynamics of x_A (see Section 6). Note that the choice is not unique: by varying k_A and γ_A in a coordinated way we are able to keep the oscillation period and amplitude approximately constant. These changes affect, however, the basins of attraction of the coexisting steady state and the circadian "limit cycle".

5.4 Full model dynamics

The final results (Fig. 6) exhibit a periodic solution, with a period around 24.1 hours, to be compared to those in Fig. 1A of Rust et al.[32] (reproduced in Fig. 5). In spite of the level of abstraction and simplicity of the model, it is able to reproduce fairly well the experimental circadian oscillation. Namely, (i) circadian periodicity; (ii) the same pattern of ordered sequence of phosphorylated forms of Kai C: T-Kai C is the first to peak, followed closely by TS-Kai C and a little later by S-Kai C; (iii) the amplitude of the waveforms for each of the forms of Kai C is similar, with peaks and troughs comparable to experiment (see Fig. 5); (iv) in particular, the low height of both T-Kai C and TS-Kai C is near zero, while low height of S-Kai C is about 1/3 or 1/4 of its peak height.

In order to verify the consistency of the assumption made during the fit regarding the concentration of Kai A, i.e. that it is mostly below $\theta_{A,2}$ (as discussed in Section 3.2), we have checked that the value for ε obtained from the fit to the phosphorylation data indeed corresponds to an amount of Kai A that is larger than those reached during the oscillations of the full model. In fact, we have

$$\varepsilon = \frac{k_S^0}{k_S} \approx 0.18 = \frac{\theta^n}{\theta^n + \bar{X}^n},$$

and this yields

$$\bar{X} = \begin{cases} 14.6, & \text{if } \theta = \theta_A = 10 \\ 19.3, & \text{if } \theta = \theta_{A,2} = 13 \end{cases}$$

From Fig. 6 it can be observed that x_A remains clearly below $\bar{X}$ during the oscillations.

6 Analysis of the Kai ABC system

In this section, several properties of the quantitative model will be tested, compared to experimental observations and deduced from theoretical analysis of the PWA model.

6.1 Robustness of the model with respect to Kai A dynamics

Since the parameters involved in the Kai A equation have not been estimated from the data, we analyzed the robustness of the model with respect to these parameters, by simultaneously randomly varying the four parameters k_A, γ_A, θ_A, and θ_S within a range of $\pm 15\%$ of their nominal values. The results show a high robustness of the properties of the model with respect to these parameters. In fact, almost all cases (99%) still exhibited a stable periodic orbit, although the period can vary. Fig. 7 shows the dependence of the period on each of the four parameters, from which it is clear that only γ_A has an effect on the period which is strong enough to stand out and show a clear correlation: the period increases/decreases with decreasing/increasing γ_A. The opposite trend is true for k_A, but its effect on the period is significantly weaker so the scatter is larger; an even weaker trend is discernible in the panels for the threshold values θ_A and θ_S. The solid lines result from singling out the effect on the

Table 1: Parameters for model (5) obtained from the experimental data in Rust et al[32]. The exponent n was chosen *a priori*, and equal for all Hill functions. Each column $n = i$ shows the estimated parameters for different choices of n. The Data column specifies the data used to estimate the given parameter.

| Parameter	$n = 2$	$n = 4$	$n = 15$	Data
k_S	10.05	6.61	5.76	Fig. 2B[32]
γ_T	0.24	0.24	0.24	Fig. 2B[32]
γ_{TS}	0.28	0.28	0.28	Fig. 2B[32]
γ_S	0.087	0.081	0.077	Fig. 2B[32]
θ_{TS}	15.37	10.16	9.15	Fig. 2B[32]
$J(p^*)$	175.6	167.4	173.3	
k_T	26.92	20.51	18.80	Fig. 2A[32]
k_{TS}	11.54	10.74	10.48	Fig. 2A[32]
k_S^0	1.57	1.22	1.15	Fig. 2A[32]
θ_U	42.08	29.95	26.21	Fig. 2A[32]
θ_T	12.02	11.42	9.82	Fig. 2A[32]
$J(p^*)$	308.9	303.9	436.7	
k_A		10		Order(k_i)
θ_A		10.0		Order(θ_i)
$\theta_{A,2}$		$1.3\theta_A$		$> \theta_i$
θ_S		5.0		(for 24h period)
γ_A		0.45		(for 24h period)
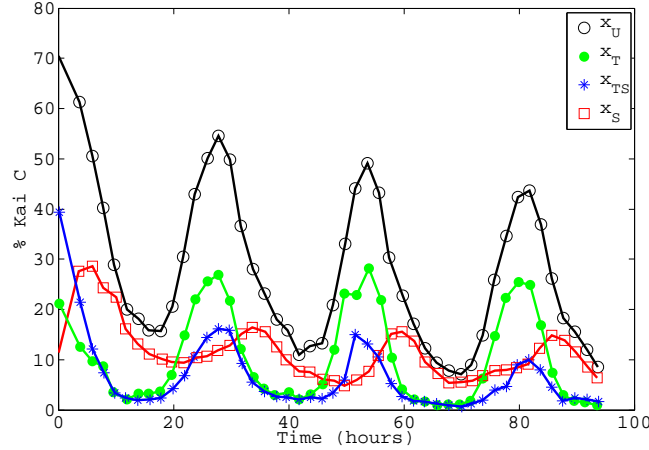
Figure 5: Self-sustained circadian oscillations of the *in vitro* Kai ABC oscillator. Taken from Fig. 1A of Rust et al.
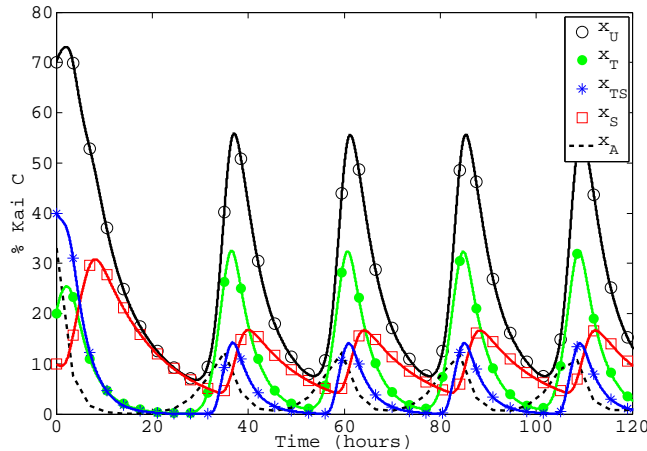
Figure 6: Solutions of model (5) with the estimated parameters, as given in Table 1. The period of the solution is approximately 24.1 h. A comparison with Fig. 5 shows that our fiducial oscillator, obtained from the fits to the non-oscillatory partial reactions, do indeed reproduce the experimental curves fairly well.

period of varying each parameter at a time, while keeping the others fixed (see caption in Fig. 7). This can be understood heuristically as follows. Assuming the other γ's fixed by the fit, the timescale for the decay of A acts in the model as the rate at which Kai A forms the complex with Kai B and S-Kai C. Therefore, increasing γ_A is equivalent, in fact, to forming this complex at a faster rate, and thus to setting up the conditions for an earlier beginning of the dephosphorylation phase.

6.2 Asymptotic behavior and stability

Following the analysis of Section 4.2, we expect to recover similar qualitative behavior for the continuous system (5). Namely, the system is expected to admit a stable steady state x^* with $x_S^* \approx 0$ and x_A^* close to k_A/γ_A. For some values of the rate constants, a steady state and a periodic orbit can coexist, each with its basin of attraction. The rectangular regions which partition the state space of system (3) will, in general, be deformed into regions defined by the nullclines of system (5), smoothing over the unphysical jump of the vector field when the trajectory crosses the domains' boundaries.

This is illustrated in Fig. 8,[1] where the filled (resp., open) circles represent initial conditions (projected onto the $x_S - x_A$ plane), leading to the steady state x^* (resp., periodic orbit). The closed circles delineate this basin of attraction of x^* (represented by $*$ inside a square). As expected, the basin of attraction has a large overlap with the (projection of the) foward-invariant rectangular region R_A of (3), defined by the upper left dashed rectangle. In contrast, the periodic orbit has a much larger basin of attraction. Interestingly, the (projected) limit cycle is fairly close to the boundary of the two basins of attraction, suggesting that it may be possible to drive the system away from the periodic orbit and into the steady state—especially if the appropriate perturbation is added at the most critical phase in the cycle (the region where $x_S \approx \theta_S$ and $x_A \approx \theta_{A,2}$). In fact, this happens when a large amount of Kai A is added to the mixture at that critical juncture (data not shown). This phenomenon is also somewhat reminiscent, although not equivalent, to the observation that a single pulse of light at a given phase may arrest the clock progression, at least transiently—as reported for circadian oscillations in different species[38].

The question then arises: how robust is the circadian oscillation, as described by this "limit cycle" solution, against fluctuations—inherent to the stochasticity of biochemical reactions—in the number of Kai A and Kai C proteins? A very rough estimate can be made as follows. Following Zwicker et al. [39], we take the number of Kai C monomers present in the test tube to be $N_C \approx 10^4$. We can then compute the effect of the expected number fluctuations on each form of Kai C and Kai A along the limit cycle.

---

[1]To better illustrate this behavior, the parameters used in Fig. 8 are those in Table 1, for $n = 4$, but with $k_A \rightsquigarrow 2k_A$. For lower k_A, the basin of attraction of the steady state decreases.
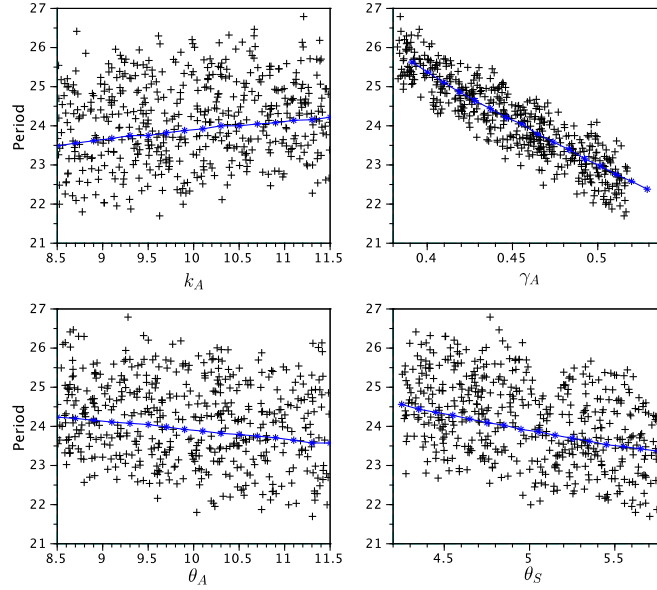
Figure 7: Perturbations of 15% for the (non-fitted) parameters: the period of the perturbed system in terms of k_A, γ_A, θ_A, and θ_S. The lines in each plot represent the changes in the period due to perturbations in each single parameter, keeping all others fixed.

The only assumptions are: (i) the size of the fluctuations, at each instant of time, follows a Poisson distribution [36], *i.e.* $x(t) = \langle x \rangle \pm \sqrt{\langle x(t) \rangle}$, where the average $\langle x \rangle$ is equal to the value taken on the limit cycle; and (ii) successive fluctuations, estimated only at a finite number of points along one period, are uncorrelated with each other. These are not unreasonable assumptions provided we seek only a rough order of magnitude estimate. The result is in Fig. 8 and shows the two dashed lines that form a "tube" around the "limit cycle". It is therefore very unlikely that fluctuations are important enough to arrest the circadian clock by throwing its trajectory into the steady state's basin of attraction.

The behavior of the variables x_T and x_{TS}, which in the PWA model decouple within the steady state' basin of attraction, is described in Fig. 9. The sliding mode solution of system (3) found in Section 4.2 is obviously no longer present in system (5), since solutions are smooth, but there exists a stable steady state near $(x_{Tot} - \theta_U - \frac{k_{TS}}{\gamma_{TS}}, \frac{k_{TS}}{\gamma_{TS}})$. A sample trajectory starting above the line L_T is depicted in Fig. 9.

6.3 Phase synchronization of a mixture

Ito *et al.* [19] have shown that when several oscillating samples, evolving with different phases, are mixed together, the Kai C phosphorylating rhythm persists in a robust manner, maintaining approximately the (common) original period. By labelling the Kai C proteins from different samples with different fluorescent tags, they have shown that the persistence of rhythmicity results from the rapid (within one period) synchronization of the different subsystems. They have hypothesized that the synchronization mechanism is based on the monomer exchange between Kai C hexamers. Subsequently, this has been shown to be theoretically feasible [24]. Here we show that monomer exchange may not be strictly necessary (even if present) for prompt synchronization. The capacity of our model to reproduce the synchronization of the sub-samples after the mixing, suggests that synchronization can be effectively achieved through Kai A sequestration alone. Note, however, that the two mechanisms are not mutually exclusive and may act simultaneously [22].

In fact, the sudden injection of new S-Kai C will presumably act to inhibit the activity of the free Kai A from the sample which, before mixing, was closer to the phosphorylation phase. Therefore, one expects that sequestration of Kai A after mixing with the formation of the Kai A•Kai B•S-Kai C complex would, for most phase combinations, favor dephosphorylation rather than phosphorylation.
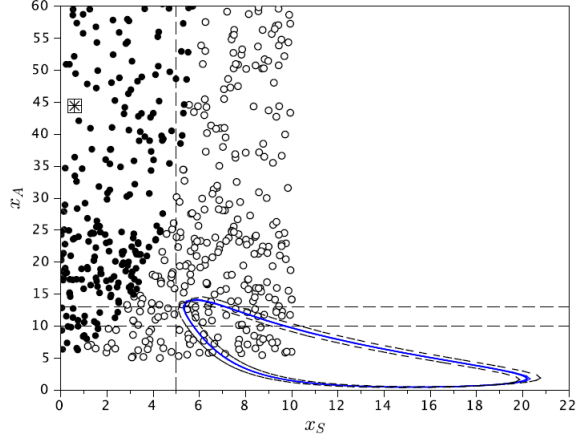
14

Figure 8: Projection of the dynamics of system (5) into the plane (x_S, x_A). The filled (resp., open) circles represent partially the basin of attraction of the steady state x^* (resp., periodic orbit). The symbol * inside a square denotes x^*. Vertical and horizontal dashed lines represent the threshold values of the PWA model. Dashed lines around the limit cycle represent the bounds on the effect of number fluctuations (of Kai A and Kai C) on the clock (see text for further details). The parameters are as in Table 1 for $n = 4$, but with $k_A \rightsquigarrow 2k_A$ to better illustrate the two asymptotic behaviors.
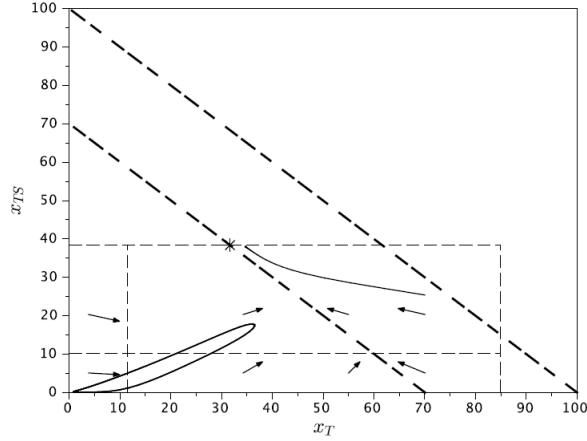

Figure 9: Projection of the dynamics of system (5) into the plane (x_T, x_{TS}). The projection of the periodic orbit and a trajectory converging to the new steady state near x^* are shown. The small arrows represent the direction of the vector field in each region. The large dashed rectangle represents a projection of the forward-invariant region R_A of the PWA model (3). The dashed line connecting the points (100,0) and (0,100) represents the conservation of Kai C, while the dashed line connecting (70,0) and (0,70) represents the line L_T (7) where a sliding mode exists. The parameters are as in Fig. 8.

15

This is indeed seen in the results by Ito *et al.* [19]. Upon mixing, the sample that was originally in the phosphorylation phase reverses its direction course to dephosphorylation, while the ones which were originally desphosphorylating maintain their course on track. Remarkably, our model reproduces these experimental facts, as will be shown below.

We now use model (3) to analyze the synchronization dynamics from a theoretical point of view and predict some results regarding mixtures of systems. To reproduce the conditions of the experiment, consider N similar systems of the form (3) and assume that the Kai C proteins have been tagged with different markers so they are distinguishable, while the Kai A proteins remain undistinguishable. The mixed system can then be described using variables x_j^i where $j \in \{T, TS, S\}$ and $i \in \{1, \ldots, N\}$ for the tagged Kai C proteins, each governed by a set of equations similar, but not equal, to (3), and a common variable for Kai A, which can bind to all x_S^i. The equations are then given by

$$\dot{y}_A = \frac{k_A}{N^2} \sum_{i=1}^{N} s^-(y_S^i, \frac{\theta_S}{N}) - \gamma_A y_A \tag{10}$$

$$\begin{cases} \dot{y}_T^i = \frac{k_T}{N} s^+(y_U^i, \frac{\theta_U}{N}) s^+(y_A, \frac{\theta_A}{N}) - \gamma_T y_T^i \\ \dot{y}_{TS}^i = \frac{k_{TS}}{N} s^+(y_T^i, \frac{\theta_T}{N}) s^+(y_A, \frac{\theta_A}{N}) - \gamma_{TS} y_{TS}^i \\ \dot{y}_S^i = \frac{k_S}{N} s^+(y_{TS}^i, \frac{\theta_{TS}}{N}) s^-(y_A, \frac{\theta_{A,2}}{N}) - \gamma_S y_S^i \end{cases} \tag{11}$$

$$y_U^i = y_{C,Tot} - y_T^i - y_{TS}^i - y_S^i. \tag{12}$$

The concentrations of each species just after the mixing change as follows (the derivation is found in the Appendix): :

$$y_j^i = x_j^i/N, \quad y_A = \frac{1}{N} \sum_i^N \frac{x_A^i}{N}, \quad y_{C,Tot} = x_{Tot}/N.$$

Now, observe that the mixed system can be viewed as the coupling of N identical systems which receive the same signal. Intuitively, since the signal is equal for individual systems (x_A), starting from the same initial conditions, the N systems x^i will follow exactly the same dynamics and their solutions will be exactly the same. Generally, when the systems are mixed they are in different phases and do not have the same initial conditions. We next show that the asymptotic solutions of the "single" system (3) can be extended to provide a solution of a mixture of N similar systems. Let $\mathbb{R}_{\geq 0} = [0, \infty)$ and let $\phi : \mathbb{R}_{\geq 0} \to \mathbb{R}_{\geq 0}^4$ denote a solution of system (3):

$$\phi(t) := (\phi_A(t), \phi_C(t)) := (\phi_A(t), \phi_T(t), \phi_{TS}(t), \phi_S(t))'.$$

Lemma 2 *Assume that system (3) has an asymptotically stable solution (periodic or not), ϕ. Then, the mixture of N similar systems (10)-(12) also admits an asymptotically stable solution, $\psi : \mathbb{R}_{\geq 0} \to \mathbb{R}_{\geq 0}^{3N+1}$, given by*

$$\psi(t) = \left(\frac{\phi_A(t)}{N}, \frac{\phi_C(t)}{N}, \ldots, \frac{\phi_C(t)}{N} \right)'. \tag{13}$$

This result follows by first multiplying the equations (10)-(12) by a factor N, which translates into an equation for $[Ny_j^i]$ similar to the ones for the single system (3):

$$[\dot{Ny_A}] = k_A \frac{1}{N} \sum_{i=1}^{N} s^-([Ny_S^i], \theta_S) - \gamma_A [Ny_A] \tag{14}$$

$$[\dot{Ny_j^i}] = k_j s^+([Ny_{j-1}^i], \theta_{j-1}) s^\pm([Ny_A^i], \theta_A) - \gamma_j [Ny_j^i]$$

for $j \in \{T, TS, S\}$. Next, substitution of $\psi(t)$ into these equations (ie. $[Ny_A] = \phi_A$, $[Ny_j^i] = \phi_j$) yields

$$\begin{aligned} \dot{\phi}_A &= k_A \frac{1}{N} \sum_{i=1}^{N} s^-(\phi_S^i, \theta_S) - \gamma_A \phi_A \\ &= k_A \frac{1}{N} N s^-(\phi_S, \theta_S) - \gamma_A \phi_A \\ &= k_A s^-(\phi_S, \theta_S) - \gamma_A \phi_A \end{aligned}$$

and

$$\dot{\phi}_j = k_j s^+(\phi_{j-1}, \theta_{j-1}) s^{\pm}(\phi_A, \theta_A) - \gamma_j \phi_j,$$

which are true equalities, as $(\phi_A(t), \phi_C(t))$ is the solution of system (3). Therefore, we conclude that (13) is an asymptotically stable solution of the mixed system (10)-(12) induced by that of the single system. The Lemma does not imply that the mixed system always synchronizes, but it guarantees that a mixture of N systems always has at least one synchronized solution and that, if two systems are in phase, they keep their phase when mixed together, as seen in Fig 10.
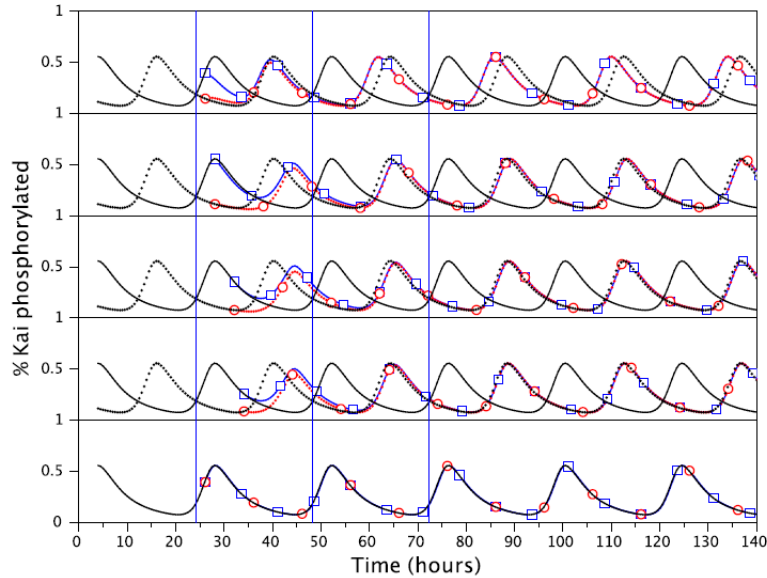

Figure 10: Phase synchronization after mixing two systems, at different stages in the cycle. The trajectories of the two separated systems are shown in solid or dotted lines. After mixing, the trajectories corresponding to each are marked by open squares or circles, respectively. The vertical lines are separated by 24h. In the first four rows, the two samples were running with opposite phases and they were mixed with several relative phases. In the bottom row, the samples were already in phase when they were mixed. Synchronization (of the curves marked with open squares or circles) obtains in all cases in less than one period of oscillation.

In order to make a quantitative comparison with experiments, we integrate numerically several cases of mixing two samples using equations (10)-(12) with Hill functions (using $n = 4$) in place of the step functions. The first four rows in Fig. 10 show the results of mixing two samples, running with opposite phases, at different relative phases. Our model reproduces the synchronization results from Ito et al. [19](see their Fig. 2), in particular, the total phosphorylation fraction from the two samples promptly synchronizes in less than a period of oscillation. The bottom panel represents samples that were running in phase, and then were mixed close to the peak of their oscillation. This shows that our model describing the Kai ABC oscillator at the monomer level is capable of reproducing the results of mixing experiments.

6.4 Response to ATP/ADP pulses

The most fundamental property of circadian clocks is their capacity to be entrained by environmental cues, such as cycles of light and dark (also of temperature) and, concomitantly, of resetting their phase in response to light/dark pulses [38, 20]. The *in vitro* oscillator cannot be entrained by light, since it lacks the required cellular machinery. Nevertheless, Rust *et al.* [31] have shown that changes of illumination in cyanobacteria affect the ratio of adenosine triphosphate (ATP) to adenosine diphosphate

(ADP). Furthermore, by controlling the relative amount of these nucleotides in the test tube, they were able to induce phase shifts in the circadian Kai ABC oscillation similar to those observed *in vivo*. The conclusion follows that pulses of ADP are natural proxies in the test tube for dark pulses *in vivo*. Mechanistically speaking, ADP acts as a competitive inhibitor of Kai C's kinase activity. As a result, since ADP cannot provide the necessary phosphoryl group in the kinase reaction, by varying the the ATP/ADP ratio during the phosphorylation phase, it is possible to modulate the circadian cycle. When the relative concentration of ADP is suddenly increased, Kai C's kinase activity and its phosphorylation rate both decrease, and the oscillation is phase shifted. Following Rust *et al.*, we model this effect by multiplying the production terms of x_T and x_{TS} in our model by a term of the form $f([ADP]) = [ATP]/([ATP] + K_r[ADP])$, where K_r measures their relative affinity for ATP and ADP. The dephosphorylation terms remain unchanged upon variations of ATP/ADP, since this ratio has no discernible effect on the dephosphorylation rate. In what follows, we describe the results obtained using our model to reproduce the experimental results.

Fig. 11 shows the effect of decreasing the ATP/ADP ratio on the non-oscillatory partial reaction for phosphorylation (cf. Section V). As the percentage of ADP increases in the mixture, the asymptotic level of Kai C phosphorylation decreases accordingly. The quantitative agreement with the experimental results (see Fig. 3 of Phong *et al* [29]) is remarkably good.
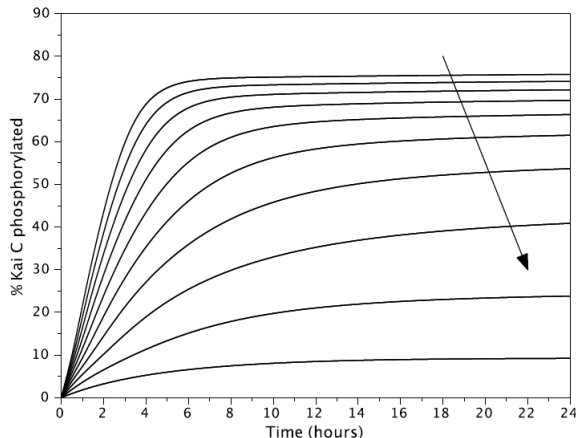

Figure 11: The result of non-oscillatory phosphorylation partial reactions for different *constant* levels of ADP. As the ADP fraction in the tube increases (in the sense of the arrow), the total phosphorylation fraction decreases. Parameters are as in Table 1 for $n = 4$, and $K_r = 1$.

Next, we compute the result of applying an ADP pulse to our fiducial clock. Following Rust *et al.* [31], we apply a 5-hour pulse by adding ADP to bring the ATP/(ATP+ADP) ratio down to the value of 0.5. We repeat this procedure at several points in the cycle, and register the resulting phase shift, $\Delta\phi$. The phase shifts were defined as follows: if the perturbed oscillation leads the control oscillation by less than half-cycle, then $\Delta\phi > 0$ and we say the clock has been advanced; if it lags by less than a half-cycle, then $\Delta\phi < 0$ and it has been delayed. In this way, we construct a phase response curve for our model. Note that this curve is not unique to a model, since it generally depends on the duration and amplitude of the perturbing signal. The results are shown in Fig. 12. It turns out that the phase response curve is fairly sensitive to the shape of the circadian clock oscillation, and we found that by adopting $n = 1$ for the Hill functions of Kai A, $h^+(A)$ and $h^-(A)$, which are not constrained by the fit to the data, does improve considerably the theoretical phase response curve, as can be readily observed from a glimpse at Fig. 12. The first thing to note is that the response $\Delta\phi$ is non-uniform along the cycle: the clock is maximally perturbed by pulses applied a few hours just before peak phosphorylation, and is minimally responsive during the dephosphorylation phase. Increasing the relative affinity K_r for ADP, the system becomes more sensitive to ADP pulses. The same is true if we increase the pulse amplitude (data not shown). The big black dots are data points from Rust *et al.* [31]. Our results show—at least, qualitatively— a remarkably good agreement with the experimental data, even though

18

we underestimate the response of the clock during the dephosphorylation just like the Rust model does. A full exploration of the phase resetting properties of the Kai ABC oscillations, including the response to other perturbations, is beyond the scope of this paper and is the subject of our current work.

The Kai ABC oscillations persist also if, instead of applying a transient pulse, the *constant* value of the ATP/ADP ratio is varied [29]. The period and amplitude of the oscillations are affected, as shown in Fig. 13. It is again surprising to see that our toy model is effectively capable to capture qualitatively most of the trends: (i) the period decreases with the ATP/ADP ratio while inflecting at very low values just before oscillations become damped; (ii) the amplitude of the oscillations also decrease with the same ratio. Quantitatively, our results are somewhat offset, and we don't capture correctly the (significantly weaker) trend of the troughs. The Rust model also fails to capture the same trend [29].

We have also verified that our fiducial 24 hour period oscillator can be entrained by different periods, with a range of entrainability that can vary between ~ 18 hours and ~ 50 hours depending on the amplitude of the forcing, *e.g.* $\max(\text{ADP/ATP}) \in [0.5, 1.0]$ (data not shown). The timescale for reaching steady state after entrainment starts is of the order of, at least, a couple of days.
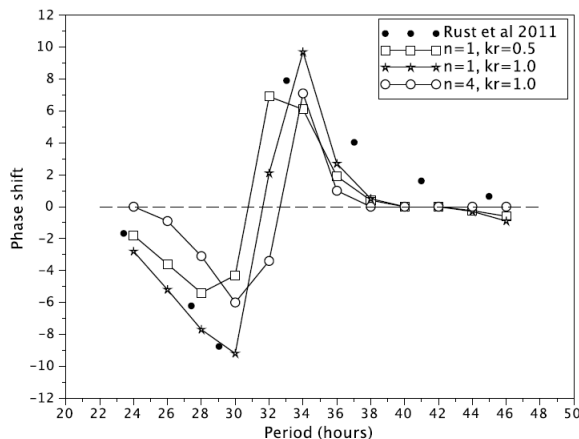

Figure 12: Phase response curves for 5h-pulses of ADP that made the ratio ATP/ADP decrease from 1 to 0.5. The dark points are experimental data from Rust et al [31]. Our model can reproduce the form of the experimental phase response curve, even though it underestimates the phase shift during the phosphorylation phase (see text for the definition of $\Delta\phi$). Peak phosphorylation is at 36 hours.

7   Conclusions

In this paper we argue for combining mathematical frameworks with differents levels of abstraction to analyze a biological system and deduce its essential underlying mechanisms. Each level in this hierarchy of models provides a group of different properties that contributes the necessary elements to built an informed model, capable of describing the main characteristics of the system and has real predictive power. At the more abstract level, Boolean models analyze the feasibility of the wiring network among the main components of the system, and provide qualitative information, such as the ordered sequence of phosphorylation/dephosphorylation. Piecewise affine systems go one level further, by modeling continuous variables but partitioning the variable space into domains where the vector field is approximated by an affine function. From a theoretical point of view, this framework has the advantage of being more tractable, allowing a rigorous study of its generic dynamical behavior, independently of the choice of parameters. In this case, it was possible to infer, for a given set of parameters, the coexistence of one steady state and a periodic orbit; and, in particular, it suggested and helped in quantifying the robustness of a critical phase of the circadian rhythm. The analysis and the synchronization solution of mixing N similar systems was also derived using the PWA framework. The third framework used in this paper consisted on smoothing the discontinuities of the PWA vector field, by substituting step functions by Hill functions with a relatively high exponent ($n = 4$). This permitted
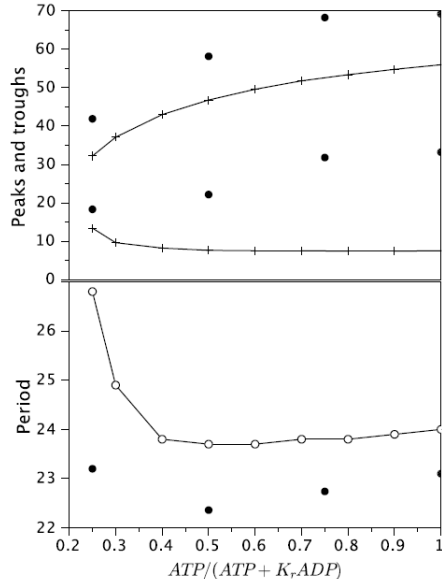
Figure 13: Comparison of peak and trough height (left), and period (right) of sustained oscillations for different *constant* levels of the ATP/ADP ratio. The experimental trends (dark points represent experimental measurements taken from Phong et al [29]) are qualitatively captured by our model, even though the values appear with a relative off-set. Below a certain fraction of ATP, the oscillation is no longer self-sustained and dampens out.

a more realistic quantitative model whose parameters were estimated by fitting to experimental data. Several properties of circadian rhythms could then be verified and compared to experiments: period, amplitudes and the ordered sequence of the phosphoforms involved, phase response curves to ADP pulses and self-sustained oscillations under different amounts of ADP (ADP pulses being a natural proxy for dark pulses, in the cyanobacterial clock), as well as the behavior of a mixture of N similar systems.

A "minimal" model which includes the basic mechanisms and a small, identifiable, number of parameters was thus validated for the Kai ABC system and used to make verifiable predictions, namely on the robustness of the circadian rhythm. These promising results suggest several possible follow-ups, ranging from an extension of the modeling formalism to include further kinetic details (eg., the detailed balancing of the interconversion of phosphoforms, as in the Rust model; or extending the description to the hexamer level [30]), to the exploration of the coupling between the protein phosphorylation cycle with the genetic regulatory network involved in the production of the Kai proteins (work currently in preparation) and with the cell division cycle. Other properties of circadian rhythms that could be further studied include temperature compensation and entrainment, as well as the circadian gating of the cell division cycle.

Acknowledgments

References

[1] M. Amdaoud, M. Vallade, C. Weiss-Schaber, and I. Mihalcescu. Cyanobacterial clock, a stable phase oscillator with negligible intercellular coupling. *PNAS*, 104:7051–7056, 2007.

[2] N. Barkai and S. Leibler. Circadian clocks limited by noise. *Nature*, 403:267–268, 2000.

[3] R. Casey, H. de Jong, and J.L. Gouzé. Piecewise-linear models of genetic regulatory networks: equilibria and their stability. *J. Math. Biol.*, 52:27–56, 2006.

[4] M. Chaves, E. Farcot, and J.-L. Gouzé. Probabilistic approach for predicting periodic orbits in piecewise affine differential models. *Bull. Math. Biology*, DOI: 10.1007/s11538-012-9773-6, 2013.

[5] M. Chaves, E.D. Sontag, and R. Albert. Methods of robustness analysis for Boolean models of gene control networks. *IEE Proc. Systems Biology*, 153(4):154–167, 2006.

[6] M. Chaves, L. Tournier, and J. L. Gouzé. Comparing Boolean and piecewise affine differential models for genetic networks. *Acta Biotheoretica*, 58(2):217–232, 2010.

[7] H. de Jong. Modeling and simulation of genetic regulatory systems : A literature review. *J. Computational Biology*, 9:67–103, 2002.

[8] H. de Jong, J.L. Gouzé, C. Hernandez, M. Page, T. Sari, and J. Geiselmann. Qualitative simulation of genetic regulatory networks using piecewise linear models. *Bull. Math. Biol.*, 66(2):301–340, 2004.

[9] J.C. Dunlap. Molecular basis for circadian clocks. *Cell*, 96:271–290, 1999.

[10] R. Edwards and L. Glass. Combinatorial explosion in model gene networks. *Chaos*, 10:691–704, 2000.

[11] E. Farcot and J.L. Gouzé. Periodic solutions of piecewise affine gene network models with non uniform decay rates: the case of a negative feedback loop. *Acta Biotheor.*, 57:429–455, 2009.

[12] A.F. Filippov. *Differential equations with discontinuous righthand-sides.* Kluwer Academics Publishers, 1988.

[13] L. Glass. Combinatorial and topological methods in nonlinear chemical kinetics. *J. Chem. Phys.* 63:13251335,1975.

[14] L. Glass and S.A. Kauffman. The logical analysis of continuous, nonlinear biochemical control networks. *J. Theor. Biol.*, 39:103–129, 1973.

[15] L. Glass and J.S. Pasternak. Stable oscillations in mathematical models of biological control systems. *J. Math. Biol.*, 6:207–223, 1978.

[16] A. Goldbeter. *Biochemical oscillations and cellular rhythms.* Cambridge University Press, 1996.

[17] S.S. Golden and S.R. Canales. Cyanobacterial circadian clocks—timing is everything. *Nature Reviews Microbiology*, 1:191–199, 2003.

[18] J.L. Gouzé and T. Sari. A class of piecewise linear differential equations arising in biological models. *Dyn. Syst.*, 17(4):299–316, 2002.

[19] H. Ito, H. Kageyama, M. Mutsuda, M. Nakajima, T.Oyama, and T. Kondo. Autonomous synchronization of the circadian kaic phosphorylation rhythm. *Nature Structural & Molecular Biology*, 14:1084–1088, 2007.

[20] C.H. Johnson, J.A. Elliot, and R. Foster. Entrainment of circadian programs. *Chronobiol. Int.*, 20:741–774, 2003.

[21] Y. Kitayama, T. Nishiwaki, K. Terauchi, and T. Kondo. Dual kaic-based oscillations constitute the circadian system of cyanobacteria. *Genes & Dev.*, 22:1513–1521, 2008.

[22] J.S. Markson and E.K. O'Shea. The molecular clockwork of a protein-based circadian oscillator. *FEBS Letters*, 583:3938–3947, 2009.

[23] I. Mihalcescu, W Hsing, and S. Leibler. Resilient circadian oscillator revealed in individual cyanobacteria. *Nature*, 430:81–85, 2004.

[24] T. Nagai, T.P. Terada, and M. Sasai. Synchronization of circadian oscillation of phosphorylation level of kai c in vitro. *Biophysical Journal*, 98:2469–2477, 2010.

[25] M. Nakajima, K. Imai, H. Ito, T. Nishiwaki, Y. Murayama, H. Iwasaki, T. Oyama, and T. Kondo. Resilient circadian oscillator revealed in individual cyanobacteria. *Science*, 308:414–415, 2005.

[26] T. Nishiwaki, Y. Satomi, Y. Kitayama, K. Terauchi, R. Kiyohara, T. Takao, and T. Kondo. A sequential program of dual phosphorylation of kaic as a basis for circadian rhythm in cyanobacteria. *EMBO Journal*, 26:4029–4037, 2007.

[27] To better illustrate this behavior, the parameters used in Fig. 8 are those in Table 1, for $n = 4$, but with $k_A \rightsquigarrow 2k_A$. For lower k_A, the basin of attraction of the steady state decreases.

[28] J.S. O'Neill and A.B. Reddy. Circadian clocks in human red blood cells. *Nature*, 469:498–503, 2011.

[29] C. Phong, J.S. Markson, C.M. Wilhoite, and M.J. Rust. Robust and tunable circadian rhythms from differentially sensitive catalytic domains. *PNAS*, 110(3):1124–1129, 2013.

[30] X. Qin, M. Byrne, Y. Xu, T. Mori, and C.H. Johnson. Coupling of a core post-translational pacemaker to a slave transcription/translation feedback loop in a circadian system. *Plos Biology*, 8:e1000394, 2010.

[31] M.J. Rust, S.S. Golden, and E.K. OShea. Light-driven changes in energy metabolism directly entrain the cyanobacterial circadian oscillator. *Science*, 331:220, 2011.

[32] M.J. Rust, J.S. Markson, W.S. Lane, D.S. Fisher, and E.K. OShea. Ordered phosphorylation governs oscillation of a three-protein circadian clock. *Science*, 318:809, 2007.

[33] R. Thomas. Boolean formalization of genetic control circuits. *J. Theor. Biol.*, 42:563–585, 1973.

[34] R. Thomas and R. D'Ari. *Biological Feedback*. CRC-Press, Boca Raton, Florida, 1990.

[35] J. Tomita, M. Nakajima, T. Kondo, and H. Iwasaki. No transcription-translation feedback in circadian rhythm of kai c phosphorylation. *Science*, 307:251–254, 2005.

[36] N.G. Van Kampen. *Stochastic Processes in Physics and Chemistry*. North-Holland, Amsterdam, London, New York, Tokyo, 1992.

[37] R. S. Wang, A. Saadatpour, and R. Albert. Boolean modeling in systems biology: an overview of methodology and applications. *Physical Biology*, 9:055001, 2012.

[38] A.T. Winfree. *The Geometry of Biological Time*. Springer-Verlag, Berlin, Heidelberg, New York, 1980.

[39] D. Zwicker, D.K. Lubensky, and P.R. ten Wolde. Robust circadian clocks from coupled protein-modification and transcription-translation cycles. *PNAS*, 107:22540–22545, 2010.

8   Appendix: Derivation of the system of equations for the mixing of samples

To deduce a set of equations for describing the synchronization dynamics after mixing, it helps to consider the particular case of $N = 2$ samples. The first requirement in the derivation is that the dynamics is not affected by the fact that the Kai C proteins have been marked with different tags. So, if we write a single equation for the total concentration of each Kai C phosphoform, $\xi_i = x_i + y_i$, where x_i and y_i are Kai C phosphoforms from each sample, it reads

$$\dot{\xi}_i = k_i \left(s^+(x_{i-1}, \theta_{i-1}) + s^+(y_{i-1}, \theta_{i-1}) \right) - \gamma_i \xi_i.$$

Next, note that there is one case for which a solution is known "a priori": that where both samples have equal numbers of proteins, equal volumes, and are mixed while in phase with each other. In such case, the samples start already synchronized and must remain so after mixing. Therefore, total phosphorylation in the mixed system should be, at all times, equal to twice that from each of the samples. The only way that this can hold using the equations above is if the following conditions are verified:

$$s^+(x_{i-1}, \theta_{i-1}) + s^+(y_{i-1}, \theta_{i-1}) = 2s^+(x_{i-1} + y_{i-1}, \theta_{i-1})$$
$$\theta_i \rightsquigarrow \theta_i/2,$$

and the same applies to the equation for x_A. If we now integrate numerically this particular case, we still do not reproduce the synchronized state after mixing. The reason has to do with the very specific form of the decay term $-\gamma_A x_A$; in the model, it represents the rate at which Kai A forms a bound complex with Kai B and S-Kai C. As a result, the rate at which Kai A decays is effectively independent of the concentration of S-Kai C. This is the price we have to pay for insisting that our model equations have linear decaying terms. Fortunately, there is a way out of this quandary without altering this term. We begin by noting that in a microscopic description of the reactions, each Kai A protein sequestred by proteins from one of the samples cannot be used, at the same time, to form a complex with proteins from the other sample. Also, since both samples are completely equivalent, except for different markers, each Kai A is equally likely to interact with proteins from either sample. The net effect is equivalent to having, at each instant, only *half of the Kai A proteins available* for interaction with Kai C proteins originally from each sample. In mathematical terms, this can be written as two different equations for x_A, each describing the interaction with one of the Kai C sub-systems, represented by x_i and y_i. Thus we divide once more by two the production term of the equation for x_A and its initial condition at the moment of mixing. The equation for x_A in the mixed system thus becomes

$$\dot{x}_A = \frac{1}{2} \frac{[s^-(x_S, \theta_S/2) + s^-(y_S, \theta_S/2)]}{2} - \gamma_A x_A,$$

with initial condition, at the moment of mixing, given by

$$x_A(t_{\text{mix}}) = \frac{1}{2} \frac{x_{A,1}(t_{\text{mix}}) + x_{A,2}(t_{\text{mix}})}{2}.$$

Everything generalizes easily for $N > 2$. The resulting system of equations for mixing N samples is given explicitly in the main text. The crucial difference with respect to the two systems of equations for the (distinguishable) proteins, x_i and y_i lies in its physical interpretation. In successive complex formation events, individual (indistinguishable) Kai A proteins are allowed to interact with proteins originally from both samples, but this exchange must obey a detailed balance at all times.