A computational study of the robustness of cellular oscillators

JORIS PAIJMANS

This thesis was reviewed by:

prof.dr. F.J. Bruggeman	Vrije Universiteit Amsterdam
prof.dr. M.J. Rust	University of Chicago, Chicago, USA
prof.dr. K.J. Hellingwerf	Universiteit van Amsterdam
prof.dr. W.T.S. Huck	Radbout Universiteit Nijmegen
dr. G.J. Stephens	Vrije Universiteit Amsterdam
dr. T.F.A. de Greef	Technische Universiteit Eindhoven

The research described in this thesis was performed at the FOM Institute AMOLF, Science Park 104, 1098 XG Amsterdam, The Netherlands. This work is part of the research programme of the Foundation for fundamental Research on Matter (FOM), which is part of the Netherlands Organization for Scientific Research.

Cover Illustration: Circadian rhythms in the clock protein concentration of different cells. From left to right, the rhythms synchronize because in each cell the clock couples to the cell cycle, and all cells grow and divide in phase.
Cover Design: Marrigje Paijmans, 2016
Printed by: GVO, Ede, The Netherlands

ISBN 9789492323095
Nederlandse titel: Een numerieke studie naar de robuustheid van cellulaire oscillatoren

A digital copy of this thesis is available online at `www.ub.vu.nl` and `www.amolf.nl`.
Printed copies can be obtained by request via `library@amolf.nl`.

VRIJE UNIVERSITEIT

A COMPUTATIONAL STUDY OF THE ROBUSTNESS OF CELLULAR OSCILLATORS

ACADEMISCH PROEFSCHRIFT

ter verkrijging van de graad Doctor aan
de Vrije Universiteit Amsterdam,
op gezag van de rector magnificus
prof.dr. V. Subramaniam,
in het openbaar te verdedigen
ten overstaan van de promotiecommissie
van de Faculteit der Exacte Wetenschappen
op vrijdag 20 januari 2017 om 11.45 uur
in de aula van de universiteit,
De Boelelaan 1105

door

Josephus Johannes Paijmans

geboren te Leiden

promotor:	prof.dr. P.R. ten Wolde
copromotor:	prof.dr. D.K. Lubensky

To get things done, you must love the doing,
not the secondary consequences.
The work, not the people.
Your own action,
not any possible object of your charity.

— Haword Roark in *The Fountainhead* by Ayn Rand

CONTENTS

1

INTRODUCTION

Throughout history, people have tried to accurately measure time to predict the rhythmic changes caused by the sun. Dating back to earliest days of civilization, by using the position of the moon relative to the stars, people could plan their agriculture and religious ceremonies throughout the year. Around 3500 BC, ancient Egyptians would coordinate their daily activities by using an obelisk as a sundial to track time. Later, methods that do not depend on the sun where developed, such as the clepsydra, candle clock and hourglass, as shown in Fig. 1.1. However, for longer time intervals lasting days or even months, these methods are too inaccurate to measure time reliably. It was not before 1656, that the first precise mechanical clock was invented by the famous Dutch physicist Christiaan Huygens. His design for a pendulum clock reduced the error in the period to less than 10 seconds a day or one ten thousands of a second.

Given humanities long strive to accurately measure time, it is remarkable that precise clocks have been on earth for billions of years already. Cyanobacteria, single cellular organisms that use sunlight for their growth, are likely to have mastered accurate time measurement at least 2 billion years ago. They contain a 24 hour rhythm, resulting from chemical reactions between proteins, which is called a circadian (Latin: *circa* for "around" and *dies* "day") clock. Nowadays, circadian clocks are found in organisms ranging from bacteria and fungi, to plants, insects and animals and allow them to anticipate the changes between day and night. Remarkably, these clocks can often maintain stable rhythms for months or even years in the absence of any daily cue from the environment, such as light-dark or temperature cycles [1]. This robustness is surprising, given the fact that experiments in recent years have vividly demonstrated that protein synthesis, which is required to sustain the clock, is highly stochastic [2]. Clearly, circadian clocks tend to be designed in such a way that they have become resilient to the intrinsic stochasticity of the underlying biochemical reactions.

Next to the biological clock, all living cells maintain another rhythm that is essential to their existence: The continuous repetition of growth and division called the cell cycle. Be it bacterial cells, yeast cells, or cells in growing tissues; they all have to duplicate all their components during the cell cycle. Interestingly, experiments in recent years

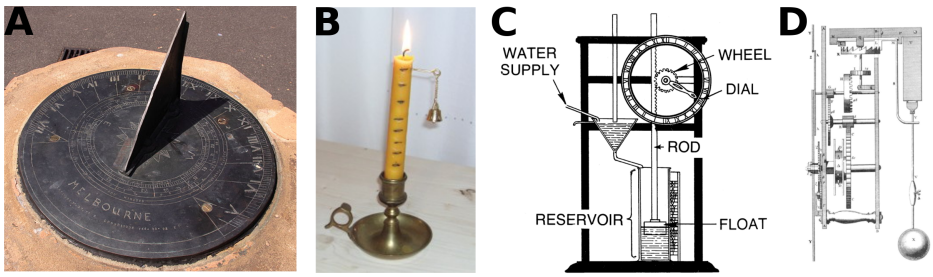

Figure 1.1: Different methods for keeping track of time. (A) A shadow cast by a sundial, tracking the sun's position, can be used to tell the time of the day. However, it only works when the sun shines. A candle clock (B) and a clepsydra or water clock (C) do not have this problem. The regular burning of candle wax or the constant filling of a vat with water can be used to measure time. However, these methods are too inaccurate to be able to tell time over the course of months. This problem was finally solved by Christiaan Huygens' design of the escapement mechanism of the pendulum clock (D). In his design, the driving of the pendulum by the falling weight (to compensate friction), did not affect the period of the swing.

have shown that the clock of one of the best characterized model systems in biology, the cyanobacterium *S. elongatus*, is not only robust to biochemical noise, but is also insensitive to variations in the growth rate: the circadian clock maintains a stable period of 24 hours, even when the cell cycle (i.e. the time between successive cell divisions) is reduced from 24 hours to 10 hours [3]. Since the cell cycle inevitably leads to oscillations in the protein synthesis rate [4], which one expects to affect the clock, the questions is how it can be that the clock's rhythm is unaffected by the cell cycle.

A fascinating property of circadian clocks is that they continue to tick with a nearly 24 hr period even when cues in the environment, like temperature, change, a phenomenon known as input compensation [5]. Yet, while these clocks are robust to a permanent change in the environment, they are still sensitive to transient changes in the environment [6, 7]. The latter phenomenon is known as entrainment, and it is necessary to keep the clock in phase with the day-night rhythm despite the presence of inevitable biochemical noise. How clocks can exhibit both input compensation and entrainability, is still poorly understood. Changes in the environment such as those in temperature typically lead to changes in the reaction rates, which tend to affect the period of the clock. Input compensation means that the clock is designed such that it has become insensitive to permanent changes in the biochemical reaction rates. Entrainability means, however, that the clock is still sensitive to periodic changes in these rates. How can the clock pick up transient changes yet ignore permanent changes in reaction rates?

In this thesis, I use mathematical modeling to study the robustness of cellular oscillators in the chaotic environment of the growing and dividing cell and in vitro. I will focus on the Kai system of the cyanobacterium *S. elongatus*, although I will also study the implications of our findings for the design of synthetic oscillators. First, in the remainder of this chapter, I provide a short overview of the history of the field followed by my view on the role of modeling in systems biology. I end with a summary of the chapters in this

thesis.

1.1. A BRIEF HISTORY OF THE RESEARCH OF CIRCADIAN CLOCKS

The research of circadian rhythms started by the 18'th century astronomer Jean-Jacques d'Ortous, who discovered that the leaves of heliotrope plans show a rhythmic motion of their leaves in response to light. More importantly, even when the plants were placed in continuous darkness, their leaves continued to move with a 24 hour period [8]. This was the first indication that some organisms have an internal rhythm, or clock, which is not simply the result of the daily rhythm in light and temperature. In the early 1970's the research in circadian rhythms took a major leap when Ronald Konopka and Seymour Benzer revealed the genetic basis of the circadian activity of the fruit fly *Drosophila melanogaster*. By studying the chromosomes of mutants, which have a different rhythm in their activity compared to wild-type flies, they could track down the difference in behavior to a single genetic locus, termed the *period* gene [9]. Flies with mutations in this gene had a different period in their activity than normal flies.

For long, scientists thought that circadian clocks can only exist in eukaryotic cells because a cell nucleus, intercellular communication and cell division times longer than 24 hours were necessary for a biological clock to function. However, in the mid 1980's, it was shown that cyanobacteria have clear circadian rhythms in their nitrogen fixation, amino acid uptake and cell division. In 1993, Kondo and coworkers were able to implement a luciferase reporter gene on the chromosome of the cyanobacterium *Synechococcus elongatus* PCC7942, *"to demonstrate unequivocally that cyanobacteria exhibit circadian behaviors that are fundamentally the same as circadian rhythms in eukaryotes"* [5]. Shortly after, the genes that form the internal clock of *S. elongatus* were identified by Ishiura *et. al.* and denoted KaiA, KaiB and KaiC, after the Japanese word for cycle [10]. They found a negative control loop where KaiC suppresses its own expression which generates the circadian oscillations in cyanobacteria. KaiC can be phosphorylated, which is regulated positively by KaiA and negatively by KaiB [11]. Finally, by a series of groundbreaking experiments, Nakajima and coworkers showed that the circadian oscillation of KaiC phosphorylation could be reconstituted in a test tube with only KaiA, KaiB, KaiC and ATP [12]. This is the only known system that shows circadian oscillations without a transcription-translation cycle. Isolated from all the other processes inside the cell, the Kai circadian clock of *S. elongatus* has become the ultimate system for studying the mechanisms involved in generating robust post-translational oscillations, both for experiments and theoretical modeling.

1.2. THE ROLE OF MODELING IN BIOLOGY

The natural world can be studied through experiments, theoretical analysis and, recently, also through computational modeling. In physics, theoretical analysis has been very successful in making quantitative predictions and explaining physical phenomena. In biology, on the other hand, most discoveries are made through experimental exploration, as the systems studied are often regarded as too complex to explain through the-

oretical analysis [13]. However, over the past decades experiments in biology have improved to such a high level that we now have very quantitative data, down to the level of single cells and even proteins [2]. This quantitative data has made it possible to verify the predications from theoretical models. Therefore, improvements in experimental techniques in biology has made the use of modeling in biology to answer questions more common. Because models of biological systems are often very complex, computers are used to simulate the models and make predictions.

In this thesis, I will use the quantitative data gathered on the cell cycle and the circadian clock of *S. elongatus* to understand how a biological clock can be so robust in its chaotic environment. However, the behavior of circadian clocks, and of biological oscillators in general, is notoriously hard to grasp intuitively or analytically, because of the many non-linear effects, feedback loops and delays that are present in their reaction networks. This makes computational modeling an essential tool for a systematic study on the robustness of biochemical oscillators, and their complex interaction with the cell cycle. Previously, modeling has been very successful by illuminating several mechanisms in the Kai circadian clockwork of *S. elongatus*, such as the role of synchronization of the phosphorylation of KaiC [14–18], the necessity of having both a phosphorylation cycle and a transcription-translation cycle [15, 19–21], the function of having two phosphorylation sites per monomer [18, 22–25], and the presence of a long temporal delay in the cycle [15, 19, 26, 27].

I will use different models of the circadian clock of *S. elongatus* to investigate the clock's robustness. This might sound strange –why not use the 'best' or most detailed model of the system? The reason is that one wants to use the simplest possible model to answer a particular question. Using a too complex model might obscure a researchers view from gaining insight into the system's inner workings. The motive for modeling, in general, is not to make the most detailed possible description of the system —since the best model is the biological system itself— but to illuminate and understand a particular mechanisms at work. Furthermore, by using a simplified model of the clock, we emphasize that a certain mechanism does not critically depend on the details of the particular system studied like *S. elongatus*. We expect to find the same mechanism in biological oscillators of different organisms as well. Therefore, in the first chapter we used the relatively simple model of the phosphorylation cycle by Van Zon *et. al.* [15]. In chapter 4 we develop a new model to reveal the thermodynamic drive behind the Kai circadian clock. To this end, we had to design a model that includes all hydrolysis, phosphorylation and dephosphorylation reactions of the Kai system, and make sure that all these reactions fulfill detailed balance. This required us to make a model that describes the clockwork in a much more detail than the models we employed in previous chapters.

1.3. SCOPE OF THIS THESIS

First, in chapter 2, we address the question of how the circadian clock can maintain its 24 hour period inside a growing and dividing cell. Cells that grow and divide, be it bacterial cells, yeast cells, or cells in growing tissues, all have to duplicate all their components during the cell cycle. Interestingly, experiments in recent years have shown that the clock of *S. elongatus* is insensitive to variations in the growth rate: the circadian clock main-

tains a stable period of 24 hours, even when the cell cycle (i.e. the time between successive cell divisions) is reduced from 24 hours to 10 hours [3]. Previous studies focused on the effect of intrinsic number fluctuations in the concentration of clock proteins and the noise in gene expression and their effect on the clock's period. However, in these studies, the direct effects of the cell cycle on the clock were not taken into account. Here we include the effect of chromosome replication which causes a doubling of the transcription rate of clock related proteins, with a period of that of the cell division cycle. We find that circadian oscillators are particularly sensitive, as they can phase-lock to the cell cycle, so that the clock period tracks the cell division time, or exhibits erratic behavior. Circadian clocks employ two general mechanisms that insulate them from the cell cycle: i) A phosphorylation-based protein modification oscillator, together with its accompanying push–pull read-out circuit that responds primarily to the ratios of different phosphoform concentrations, which makes the clock less susceptible to perturbations in protein synthesis rate; ii) the presence of multiple, asynchronously replicating copies of the same chromosome diminishes the effect of replicating any single copy of a gene.

We extend our analysis of the effects of the cell cycle in chapter 3, to two famous synthetic oscillators: the repressilator by Elowitz and Leibler [28] and the dual-feedback oscillator by Stricker *et. al* [29]. Given that the networks of these oscillators have a simpler design than biological oscillators, as the reaction networks contain fewer genes and do not involve post-translation modification, we want to know how sensitive synthetic oscillators are to the effects of the cell cycle. We find that the period of both oscillators is strongly perturbed by the cell cycle, and that the effect critically depends on the position of the genes of the oscillator on the chromosome. Even in the limit of high levels of noise in the replication times of the genes, which reduces the coupling between the clock and the cell cycle, both oscillators show clear signatures of locking. We argue that the study of synthetic oscillators can be used to test which mechanisms are critically important for a robust biological oscillator inside growing and diving cells.

In chapter 4 we present a new, thermodynamically consistent, statistical-mechanical model of the Kai circadian clock. We were inspired by the remarkable recent observation that ATP is regenerated during the dephosphorylation of KaiC, which suggests KaiC goes through its phosphorylation cycle without consuming ATP —something that clearly violates the second law of thermodynamics. We set out to answer the question of how the clock is thermodynamically driven. In our new model, building on ideas from previous mathematical models, KaiC consists of the CI and the CII domain. Hydrolysis of ATP in the CI domain provides the thermodynamic driving force for the conformational switch in KaiC, which determines whether the phosphorylation level of KaiC increases or decreases. Phosphorylation of the CII domain acts as a timer for when the hexamer switches confirmation. Using a dedicated kinetic Monte Carlo algorithm, which makes it possible to efficiently simulate the system which contains over a billion reactions, we show that our model can describe the majority of experimental results. Remarkably, the model shows very strong input compensation: When we lower the ATP level in the buffer from 100 to 50%, which slows down the rates at which KaiC phosphorylates, the period in the phosphorylation fraction only slightly changes. Because phosphorylation is the slowest reactions in the biochemical network, and therefore determines its period , we want to know how the period can be constant under changing ATP levels.

This motivated us to test how well our new model fulfills two essential criteria for a functional circadian clock: The clock should be entrainable, such that it is always in phase with the time of the day, and should exhibit input compensation, such that the period remains 24 hours in different environments. Experiments show that the daily fluctuation of the ATP fraction inside the cyanobacterium is the main cue for entrainment, while the period of the clock almost does not depend on the bulk ATP fraction [7, 26, 30]. In chapter 5 we compare our new model with the models by Van Zon *et. al.* [15] and Rust *et. al* [22], and find that the new model exhibits the best trade-off between input compensation and entrainability under a change in the ATP fraction of the bulk. Performing stochastic simulations at the level of individual hexamers allows us to identify a new mechanism for input compensation: At lower ATP fraction, the individual hexamers make a shorter cycle in the phosphorylation state space, which compensates for the lower pace at which they traverse the cycle.

2

DISCRETE GENE REPLICATION EVENTS DRIVE COUPLING BETWEEN THE CELL CYCLE AND CIRCADIAN CLOCKS

Many organisms possess both a cell cycle to control DNA replication and a circadian clock to anticipate changes between day and night. In some cases, these two rhythmic systems are known to be coupled by specific, cross-regulatory interactions. Here, we use mathematical modeling to show that, additionally, the cell cycle generically influences circadian clocks in a non-specific fashion: The regular, discrete jumps in gene-copy number arising from DNA replication during the cell cycle cause a periodic driving of the circadian clock, which can dramatically alter its behavior and impair its function. A clock built on negative transcriptional feedback either phase locks to the cell cycle, so that the clock period tracks the cell division time, or exhibits erratic behavior. We argue that the cyanobacterium Synechococcus elongatus *has evolved two features that protect its clock from such disturbances, both of which are needed to fully insulate it from the cell cycle and give it its observed robustness: a phosphorylation-based protein modification oscillator, together with its accompanying push-pull read-out circuit that responds primarily to the ratios of the different phosphoforms, makes the clock less susceptible to perturbations in protein synthesis; and the presence of multiple, asynchronously replicating copies of the same chromosome diminishes the effect of replicating any single copy of a gene.*

The content of this chapter has been published as Joris Paijmans, Mark Bosman, Pieter Rein ten Wolde, and David K. Lubensky. *Discrete gene replication events drive coupling between the cell cycle and circadian clocks.* Proc Natl Acad Sci **113**, 22:4063–4068, 2016 [21].

2.1. INTRODUCTION

Circadian clocks—autonomous oscillators with a roughly 24 hour period that can be entrained to daily cycles of light and dark —are thought to confer important advantages on living cells by allowing them to anticipate diurnal environmental changes. Recent decades have seen considerable progress in elucidating both the architecture and the function of these biological timekeepers. Circadian clocks, however, are not the only oscillatory systems present in living cells. Most notably, cell growth and division are governed by a cell cycle, which can in many contexts be viewed as an autonomous oscillator. Much recent attention has been directed towards the connections between these two rhythmic systems, which are relevant for processes ranging from plants' response to shade [31] to cancer susceptibility [32, 33]. In particular, it is now clear that circadian clocks can exert specific regulatory influences on the cell cycle, and a number of experimental and modeling studies have sought to tease out the implications of this regulation [3, 34–40]. Here, we argue that, in addition to direct, specific regulation of one oscillator by the other, there must also be more generic connections between the circadian clock and the cell cycle [32, 39–41]. In particular, we focus on the consequences of the discrete gene replication events that accompany DNA replication. We show that, as a result of the regular jumps in gene copy number caused by these events, the cell cycle must, very generally, contribute a periodic forcing to the circadian clock. This forcing can markedly change clock behavior and degrade clock function. We propose that cyanobacterial clocks have evolved specific features that can mitigate this effect. More broadly, this generically strong coupling to the cell cycle implies important constraints on the design of biological timekeepers if they are to remain accurate in dividing cells.

It is widely accepted that protein levels depend on a cell's gene dosage. As was recently shown experimetnally, a doubling of the number of chromosomal copies of a gene results in an approximate doubling of its mRNA synthesis rate and thus to a corresponding increase in its protein levels [4]. Most often, however, such effects are considered in the context of a change in the number of autosomal gene copies that persists throughout an organism's lifetime [42], as, e.g., in the haploinsufficiency of certain genes [43]. It is less often acknowledged that the number of copies of all genes varies over each cell cycle, despite evidence that these variations have measurable consequences [4, 44–47]. Because of the well-known phenomenon of phase locking of oscillators [48], regular, periodic changes in gene dose are likely to be especially relevant to cellular oscillators that depend on gene expression. A circadian clock that became slaved to the cell cycle, for example, would lose its identity as an autonomous timekeeper, and thus much of its ability to perform its biological function. Here, we show that oscillators built on negative transcriptional feedback—a common motif in both prokaryotic and eukaryotic clocks—are indeed very strongly affected by driving from periodic gene replication events. This immediately raises the question of how real biological clocks are able to function in growing, dividing cells. To address this, we study the circadian clock of the cyanobacterium *Synechococcus elongatus*, which is known to exhibit stable rhythms over a wide range of growth rates [20, 49], but whose clock appears not to regulate DNA replication [3], suggesting exactly the sort of unidirectional forcing of the clock by the cell cycle that might have been expected to impair clock function.

The *S. elongatus* clock combines a negative transcriptional feedback oscillator (the

transcription-translation cycle, or TTC) with a core phosphorylation-based post-translational oscillator (the protein phosphorylation cycle, or PPC). This clock is known to comprise two parts, a phosphorylation-based post-translational oscillator (the protein phosphorylation cycle, or PPC) whose core consists of the proteins KaiA, KaiB, and KaiC, and a more conventional module based on negative transcriptional feedback (the transcription-translation cycle, or TTC). Remarkably, the PPC can be reconstituted *in vitro* with purified proteins [12], allowing detailed study of the mechanisms behind its oscillation. A number of studies have begun to converge on the view that the PPC works by synchronizing the intrinsic phosphorylation cycles of individual KaiC hexamers, primarily through phosphorylation-dependent sequestration of KaiA by KaiC [15, 18, 22, 50–53]. Although many details of the TTC remain murkier, it seems clear that the protein RpaA plays a central role, regulating the expression of clock components in a manner that ultimately reflects the KaiC phosphorylation state [27, 54–57]. Depending on light and nutrient levels, *S. elongatus* can have doubling times ranging from 6 to 72 h [20]; the cell cycle period is thus of the same order as the clock period of roughly 24 h, opening the way for interactions between the two. Indeed, the circadian clock is known to gate mitosis, prohibiting cell division during certain clock phases [3, 36, 37], although in constant light this gating leaves both DNA replication and cell growth essentially unchanged [3]. Conversely, Mori and Johnson argued that cell growth and division don't affect the *S. elongatus* circadian clock [49]. We use mathematical modeling to study the unidirectional forcing of the clock by the cell cycle. We identify specific features of the *S. elongatus* clock that tend to insulate it from entrainment by regular gene replication events. Nonetheless, we argue that, under certain conditions, it should be possible to observe signatures of periodic forcing of the clock by the cell cycle. We further suggest how some of the clock's protective mechanisms might be weakened experimentally, leading to much stronger signatures of its coupling to the cell cycle.

Below, we first model the effects of cell growth and division on a constitutively expressed protein. We show that gene replication, not cell division, is the essential cell-cycle event that influences protein concentrations and that, as long as the constitutively expressed protein is not subject to rapid, active degradation, its concentration varies little over the cell cycle. In contrast, gene replication —which introduces a periodic forcing that drives the circadian clock— can dramatically affect the behavior of a negative transcriptional feedback oscillator (NTFO): the NTFO locks to the cell cycle over a range of cell-division times of many hours and shows erratic behavior outside this regime [41]. Such an oscillator cannot function as a clock. We next ask how the real cyanobacterial clock can be so apparently undisturbed by the cell cycle. We find that incorporating both a PPC and a TTC into the clock significantly weakens coupling to the cell cycle, especially when the clock is read out by a push-pull network that is more sensitive to ratio of concentrations of different phosphorylation states than to their absolute values. The presence of multiple chromosome copies has a still more striking effect: If the cell has 4 copies after division (rather than only 1), as can often be the case in *S. elongatus*, and if these are replicated one after the other [58], then the dose of the clock genes changes much more gradually, and cell cycle effects are almost completely lost. Thus, *S. elongatus* may have evolved to carry multiple, identical chromosome copies in part to insulate its circadian clock from its DNA replication cycles.

2.2. MODELS AND RESULTS

2.2.1. THE CELL CYCLE'S EFFECT ON A CONSTITUTIVELY EXPRESSED GENE IS WEAK

Before turning to the more complex case of a circadian clock, we first investigate how the concentration of a single, constitutively expressed protein varies over a cell cycle [4]. To this end, we add regular, rhythmic DNA replication and mitosis to a simple model of protein production and dilution.

The key quantities in our description are the number of copies $g(t)$ of the gene of interest and the cell volume $V(t)$. These vary periodically in time as sketched in Fig. 2.1A–B, with a period given by the cell division time T_d. We assume for now that there is only one gene copy present immediately after cell division. This copy is replicated at some time before the next division, at which point $g(t)$ jumps from 1 to 2. When the cell divides, the chromosomes are split between the daughter cells, and $g(t)$ returns to 1. The cell volume grows exponentially: $V(t) = V_0 \exp(\mu_\mathrm{d} t)$, with $\mu_\mathrm{d} = \log(2)/T_\mathrm{d}$. When t reaches T_d, division occurs, and $V(t)$ drops back from $2V_0$ to V_0.

The variables $g(t)$ and $V(t)$ define the gene density $G(t) \equiv g(t)/V(t)$. As long as noise and spatial variations are neglected, the behavior of a biochemical network depends only on protein *concentrations*, not separately on protein numbers and cell volume. As a result, the system responds to the protein synthesis rate per unit volume, proportional to $G(t)$, but not to $g(t)$ and $V(t)$ individually (Eq. 2.1, below). Fig. 2.1C shows that $G(t)$ has only a single discontinuity during the cell cycle, corresponding to the doubling of $g(t)$ when the gene is copied; at cell division, both $g(t)$ and $V(t)$ are halved, so their ratio is unchanged. Importantly, then, the meanfield, deterministic dynamics of a biochemical network is sensitive to the timing of DNA replication but not of cell division. This dynamics is likewise unaffected by any gating of cell division by the circadian clock, provided, as is the case in *S. elongatus* [3, 37], that this gating does not affect DNA replication or cell growth. Similarly, regardless of when during the division cycle the gene is copied, the time dependence of $G(t)$ is always the same: It doubles, decays exponentially for a time T_d, then doubles again, etc. The exact moment of gene replication affects only the average value of $G(t)$, which can be absorbed, for modeling purposes, into the parameter β (Eq. 2.1, below). For simplicity, we thus always assume that the gene is replicated exactly at $t = T_\mathrm{d}/2$.

Given the gene density $G(t)$, the concentration $C(t)$ of a constitutively expressed protein evolves as:

$$\frac{dC(t)}{dt} = \beta G(t) - \mu_\mathrm{d} C(t). \tag{2.1}$$

Here, proteins are expressed at a rate β per gene copy and diluted by cell growth at a rate $\mu_\mathrm{d} = \log(2)/T$. We thus assume that, as is true for many bacterial proteins, the protein is not subject to active degradation [59]. Fig. 2.1D shows how $C(t)$ varies over the cell cycle. Remarkably, even though the protein production rate doubles each time the gene is replicated, the protein concentration varies by no more than a few percent: The discrete jumps in protein production are smoothed out by the slow protein dilution. Thus, a protein that is constitutively expressed and not actively degraded is little affected by the cell cycle.

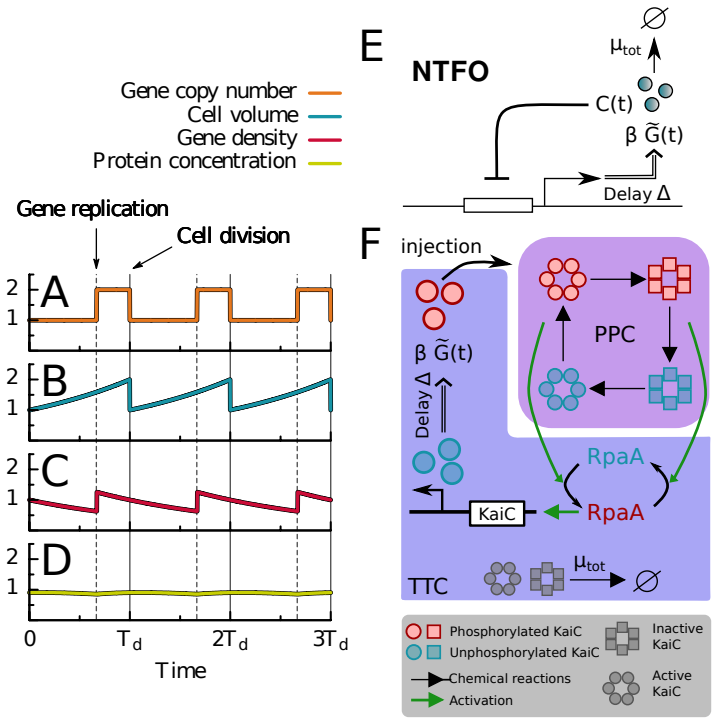

Figure 2.1: DNA replication but not cell division affects average expression levels; for a protein that is constitutively expressed and decays by dilution only, the effect is small. Schematic time courses of the gene copy number $g(t)$ (A), the cell volume $V(t)$ (B), the gene density, $G(t) = g(t)/V(t)$ (C), and the concentration $C(t)$ of a constitutively expressed protein that decays only by dilution (D). Time in units of the cell division time T_d; vertical axes, arbitrary units. The gene density (C) has a discontinuity when the gene is replicated (vertical dotted lines) but not at cell division (vertical solid lines), when both $g(t)$ and $V(t)$ are halved. Even though the protein synthesis rate doubles when the gene is replicated, the maximum deviation of $C(t)$ from its time average is less than 4% (D). (E) The NTFO model: A protein with concentration $C(t)$ represses its own transcription with a delay Δ. (F) Zwicker [19] model for coupled phosphorylation (PPC, purple background) and transcription-translation (TTC, blue background) cycles. KaiC hexamers switch between an active conformational state (circles) in which their phosphorylation level tends to rise and an inactive state (squares) in which it tends to fall. Active KaiC activates RpaA and inactive KaiC inactivates RpaA; active RpaA (red) activates *kaiBC* expression, leading (after a delay) to the injection of fully phosphorylated KaiC (pink) into the PPC.

2.2.2. THE CELL CYCLE STRONGLY PERTURBS BOTH THE PERIOD AND THE AMPLITUDE OF A NEGATIVE TRANSCRIPTIONAL FEEDBACK OSCILLATOR

Although the concentration of a protein that is constitutively expressed does not vary much over the cell cycle, oscillators are known to be far more sensitive to periodic driving

than non-oscillatory systems [48]. We thus next consider a simple model for a clock built on delayed, negative transcriptional feedback (Fig. 2.1E). The model consists of a single variable, $C(t)$, describing the concentration of proteins that inhibit their own production:

$$\frac{dC(t)}{dt} = \beta \tilde{G}(t) \frac{K_c^n}{K_c^n + C(t-\Delta)^n} - \mu_{\text{tot}} C(t). \tag{2.2}$$

We impose a fixed delay Δ between the initiation of transcription and the appearance of functional proteins. Therefore, protein production at time t is proportional to the gene copy number $g(t-\Delta)$ at time $t-\Delta$. These proteins 'arrive' in the cell volume $V(t)$ at time t. The protein synthesis rate per unit volume at time t is thus proportional to the *protein production density* $\tilde{G}(t) \equiv g(t-\Delta)/V(t)$. $\tilde{G}(t)$ is a generalization of the gene density $G(t)$ of the preceding section to the case with a delay Δ and parametrizes the periodic forcing of the NTFO by gene replication. Proteins disappear with a total rate $\mu_{\text{tot}} = \mu_{\text{d}} + \mu_{\text{act}}$, where as before μ_{d} describes dilution due to cell growth, and μ_{act} describes possible active degradation. Including both terms allows us to vary the doubling time T_{d} while holding μ_{tot} constant and hence, in our simulations, to distinguish the trivial influence of the cell cycle on the clock through the dilution rate μ_{d} from other effects.

We next define the peak-to-peak time T_{PtP} as the time between successive peaks in $C(t)$ (see Fig. 2.2 and *Supporting Information* [SI]); T_{PtP} reduces to the period of the circadian clock when oscillations are regular but remains defined when the cell cycle induces more erratic behavior. In Fig. 2.2A we plot the average peak-to-peak time $\langle T_{\text{PtP}} \rangle$ for a range of division times T_{d} at fixed μ_{tot}.

As expected from the general theory of driven oscillators [48], the curve shows two striking features: First, around division times which are fractions or multiples of the clock's intrinsic period of 24 h, the cell cycle determines the period of the clock. Especially around $T_{\text{d}} = 24$ and 48 h, the average peak to peak time is directly proportional to T_{d}. At $T_{\text{d}} = 24$ h (1:1 locking), $\langle T_{\text{PtP}} \rangle = T_{\text{d}}$, and the amplitude of each clock oscillation cycle is the same (Fig. 2.2B). At $T_{\text{d}} = 48$ h (2:1 locking), however, $\langle T_{\text{PtP}} \rangle = T_{\text{d}}/2$, and two full clock cycles are required to make up a single division time. Because these two cycles occur at different gene densities, successive peaks in the trace of $C(t)$ have alternately large and small amplitudes.

Second, the standard deviation of T_{PtP} becomes very large just outside the locking regions. Fig. 2.2C shows that this variability in the phase of $C(t)$ is accompanied by substantial fluctuations in the amplitude for $T_{\text{d}} = 27$ h. Because the difference between T_{d} and the intrinsic clock period is just too large to allow stable locking, the clock constantly tries to lock to the cell cycle, but slips from time to time. As a result, the cell cycle dramatically disrupts the clock. In the *SI* we show that both of these effects survive the introduction of intrinsic noise in chemical reactions and of stochasticity in the timing of DNA replication (Figs. 2.E.1-2.F.1; see also Fig. 2.J.1). Fig. 2.3 qualitatively explains how locking arises in the NTFO.

2.2.3. A PHOSPHORYLATION CYCLE MAKES THE CLOCK MORE ROBUST AGAINST A TIME-VARYING GENE DENSITY

To study how a more realistic clock can become resilient to variability in the gene density, we turn to the *S. elongatus* circadian clock, and more specifically to the model of

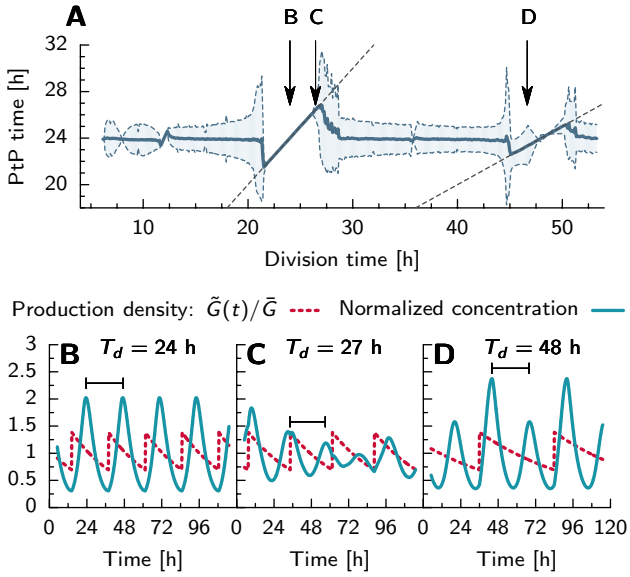

Figure 2.2: Periodic gene replication dramatically affects a negative transcriptional feedback oscillator (NTFO). (A) The average peak-to-peak time $\langle T_{\mathrm{PtP}} \rangle$ (solid curve) versus the cell division time T_{d} at fixed μ_{tot} and β. The shaded region shows the standard deviation of the peak-to-peak times (see *SI text*). Dashed lines indicate regions where the clock locks to the cell cycle with periods in a 1:1 (left) or 2:1 (right) ratio. (Smaller locking regions around $T_{\mathrm{d}} = 6, 12,$ and 36 h are not marked.) (B–D) Protein concentration $C(t)$ (blue solid line) and the protein production density $\tilde{G}(t) = g(t-\Delta)/V(t)$ (red dashed line) for the values of T_{d} indicated by the arrows in (A); horizontal brackets in (B–D) illustrate the definition of the peak-to-peak time T_{PtP}. At $T_{\mathrm{d}} = 24$ h (B), the clock locks firmly to the cell cycle. For $T_{\mathrm{d}} = 27$ h (C), the cell-cycle period is just too large for locking; as a result, the cell cycle dramatically disrupts the clock, leading to a large standard deviation of T_{PtP} (see panel A). At $T_{\mathrm{d}} = 48$ h (D), two oscillation cycles of the NTFO fit exactly in one division time. The larger amplitude oscillation cycle corresponds to cell cycle phases where $\tilde{G}(t)$ is higher and the smaller amplitude to phases where $\tilde{G}(t)$ is lower. Similar results are obtained upon varying T_{d} at constant μ_{act} (Fig. 2.G.1).

Zwicker *et al.* [15, 19] (Fig. 2.1F). This model provides a detailed description of the clock, including the synchronization of the phosphorylation state of different KaiC hexamers via KaiA sequestration and the coupling of the PPC oscillator to the TTC via RpaA. It represents KaiC as a hexamer but does not explicitly take into account that each KaiC monomer has two distinct phosphorylation sites [22, 60]. In the *SI Text*, Fig. 2.K.1, we show that a model based on that of Rust *et al.* [22], which describes KaiC at the level of monomers with two phosphorylation sites, gives similar results. We thus expect that still more elaborate models of the PPC, which include hexameric KaiC with two phosphorylation sites per monomer, such as our new model presented in chapter 4 and [18], will lead to similar results. To include gene replication, we modify the model of [19] so that the delayed negative feedback on KaiC production is modulated by a regularly oscillating

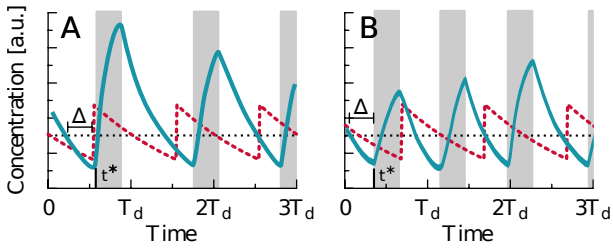

Figure 2.3: Locking mechanism for the NTFO. Shown are time courses of the production density $\tilde{G}(t) = g(t-\Delta)/V(t)$ (dashed red lines) and the protein concentration $C(t)$ (solid blue lines). For clarity, we consider the limit $n \to \infty$, in which the Hill function describing autoregulation (Eq. 2.2) reduces to a step function with repression threshold K_c, denoted by the dotted horizontal line. Shaded regions indicate times when $C(t)$ is rising. The panels correspond to two different initial phase differences between the NTFO and the cell cycle. In each case, when $C(t)$ drops below K_c at time $t^* - \Delta$, protein production starts, but because of the delay Δ, new molecules are injected into the system only at time t^*. (A) The gene has replicated just before $t^* - \Delta$, and $\tilde{G}(t^*)$ is hence large, yielding a large amplitude for the next NTFO cycle. Because the rate of protein decay is independent of $\tilde{G}(t)$, the period of the NTFO cycle is correspondingly long. The subsequent NTFO cycle thus begins at smaller $\tilde{G}(t^*)$, causing it to have a smaller amplitude and a shorter period. (B) The gene has not yet replicated at time $t^* - \Delta$, and $\tilde{G}(t^*)$ is therefore low; consequently, the amplitude and period of the next NTFO cycle are small. The beginning of the subsequent cycle is then shifted towards higher $\tilde{G}(t^*)$, increasing its period. In both cases, the result is that, after a few cell cycles, the period of the NTFO oscillation approaches that of the cell cycle, yielding stable 1:1 locking where the two oscillators have a well-defined phase relation. The largest amplitude and thus longest possible clock period arise when the protein synthesis phase (grey bar) coincides with the maximal $\tilde{G}(t^*)$; if T_d increases beyond this maximal period, locking cannot occur. An analogous loss of locking occurs if T_d decreases below the minimal possible clock period. In either case, the clock shows erratic behavior until T_d approaches values where 1:2 or 2:1 locking is possible.

protein production density $\tilde{G}(t)$ (see *SI Text*). We follow both the total KaiC concentration $C_{tot}(t)$ and the KaiC phosphorylation fraction $p(t) = \sum_{n=1}^{6} nC_n(t)/(6\,C_{tot}(t))$, where C_n is the concentration of n-fold phosphorylated KaiC hexamers.

Fig. 2.4A shows that a model with a PPC coupled to a TTC has a smaller locking window than an NTFO and lacks the large deviations in T_{PtP} just outside the locking region. The *S. elongatus* clock is hence more robust to gene replication than one based only on negative transcriptional feedback.

2.2.4. CLOCK READOUT THROUGH AN RPAA-BASED PUSH-PULL NETWORK FILTERS OUT CELL-CYCLE-DEPENDENT VARIATIONS IN PROTEIN CONCENTRATIONS

Although the variance of T_{PtP} outside of the locking region is relatively small for the combined TTC-PPC model, Fig. 2.4B shows that $C_{tot}(t)$ exhibits strong amplitude fluctuations, mirroring those observed for the NTFO (Fig. 2.2). The phosphorylation fraction $p(t)$, in contrast, is far more resilient, suggesting that the clock encodes temporal infor-

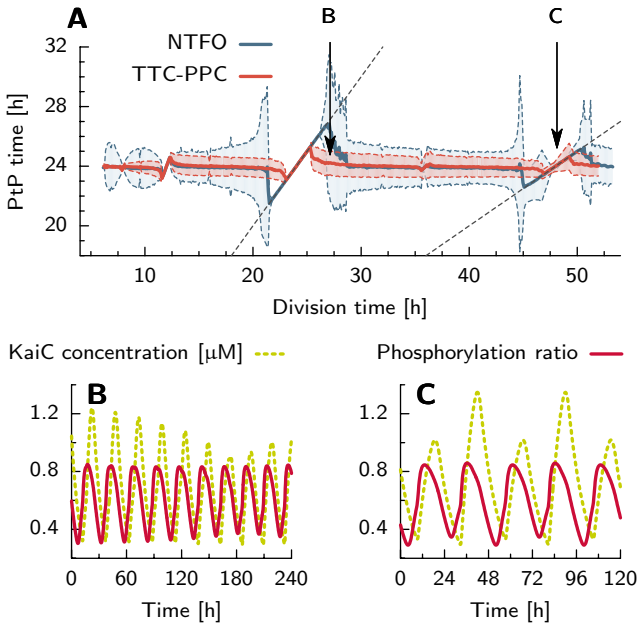

Figure 2.4: A clock with interlocked phosphorylation and transcriptional cycles is more robust against perturbations from periodic gene replication. (A) The average peak-to-peak times $\langle T_{\mathrm{PtP}} \rangle$ of the phosphorylation level $p(t)$ of the coupled TTC-PPC model of the Kai system [19] (red solid curve) and of $C(t)$ of the NTFO (solid blue curve, same as Fig. 2.2A), as a function of the cell division time T_{d}. The shaded regions show the standard deviation of T_{PtP}. Both the widths of the locking regions and the standard deviations of the peak-to-peak time outside the locking regions are smaller for $p(t)$ of the Kai system than for $C(t)$ of the NTFO. Arrows indicate division times for which we show time traces in (B,C). (B) The total KaiC concentration $C_{\mathrm{tot}}(t)$ (dashed line) and $p(t)$ (solid line) at $T_{\mathrm{d}} = 26$ h. Though the amplitude of $C_{\mathrm{tot}}(t)$ is strongly affected by gene replication, the amplitude of $p(t)$ is nearly constant. (C) Plots of $p(t)$ and $C_{\mathrm{tot}}(t)$ at $T_{\mathrm{d}} = 48$ h, where the amplitude of $C_{\mathrm{tot}}(t)$ alternates between a low and a high value depending on the gene copy number in the cell. In contrast, $p(t)$ is almost unaffected by gene replication.

mation more reliably in $p(t)$ than in $C_{\mathrm{tot}}(t)$. Intriguingly, the RpaA-centered push-pull network that transmits this timing signal to downstream genes [27, 54–57, 61] in fact responds primarily to $p(t)$: Because the rates of RpaA phosphorylation and dephosphorylation are indirectly controlled by different KaiC phosphoforms, variations in C_{tot} at fixed p change both rates together, leaving the fraction of phosphorylated RpaA largely unaffected. In contrast, changes in p shift the balance between the two opposing reactions and so modify the RpaA phosphorylation fraction (Fig. 2.H.1 and *SI text*). Thus, not only is the basic PPC-based timekeeping mechanism insulated from variations in protein synthesis, but the readout mechanism selectively follows this more robust signal.

2.2.5. MULTIPLE CHROMOSOME COPIES WEAKEN THE CELL CYCLE'S INFLUENCE ON THE CLOCK

While the PPC reduces gene replication's effect on the clock, it does not eliminate it entirely (Fig. 2.4). What other mechanisms might explain the observed resistance of the *S. elongatus* clock to the cell-cycle? It is known that *S. elongatus* has multiple, identical copies of its chromosome [58, 62–64]. These are not duplicated simultaneously, but rather one at a time, so that DNA replication occurs at a roughly constant rate throughout the cell cycle; furthermore, the timing of chromosome duplication appears to be independent of the phase of the clock [3, 58, 62, 64, 65]. Motivated by this observation, we consider a cell that starts with N chromosomes after division and let $g(t)$ rise to $2N$ in N evenly spaced steps (Fig. 2.5A–B). For larger N, the gene-copy number $g(t)$ increases more gradually, and hence the discrete jumps in the gene density $G(t)$ are considerably smaller. The effect on the clock is dramatic: in both the NTFO (solid line) and the TTC-PPC (dashed line), not only do the locked regions almost disappear for $N = 4$ (Fig. 2.5C), but the variance in the peak-to-peak time T_{PtP} becomes very small (Fig. 2.5D). This latter effect persists even when significant stochastic variability in the rhythm of gene replication (parameterized by the standard deviation σ_{rep} in the replication times —see *SI text*) is introduced. In fact, while *S. elongatus* can have as many as $N \approx 4$ chromosome copies at the beginning of the cell cycle [58, 62–64], these changes are already apparent when N is increased from 1 to 2 (Fig. 2.J.1).

2.3. DISCUSSION

Given the pleiotropic roles of both the cell cycle and the circadian clock, it is natural to ask whether they also influence each other. Our central observation is that such influence need not involve specific interactions between the core genes or proteins of the two systems [32, 39, 40]; rather, the simple fact that the number of cellular copies of a given gene necessarily experiences discrete jumps during DNA replication (Fig. 2.1) implies that clocks must in general feel a periodic driving from the cell cycle [41]. Whereas some genetic circuits can simply average over this time-varying input, oscillators—including biological clocks—are known to be especially sensitive to rhythmic forcing. Indeed, an NTFO either locks to the cell cycle or shows erratic oscillations for a range of doubling times T_{d} (Fig. 2.2), losing its ability to function as a clock in either case.

In light of this strong and detrimental coupling between the cell cycle and a simple transcriptional clock, it is all the more striking that the *S. elongatus* clock is so stable. Our analysis highlights two features of the cyanobacterial clock that are predicted to allow the necessary decoupling from the cell cycle. First, a time-varying gene dosage influences a clock with an autonomous post-translational oscillator less than it does a purely transcriptional clock; even *within* the combined TTC-PPC, the oscillations of the KaiC phosphorylation fraction $p(t)$ are less affected by periodic gene replication than are those of the total KaiC concentration $C_{\mathrm{tot}}(t)$ (Fig. 2.4, Fig. 2.F.1C). Strikingly, the RpaA-based push-pull network that communicates the clock state to the rest of the cell responds to p while ignoring the more strongly fluctuating C_{tot} (somewhat in the spirit of mechanisms that improve the robustness of bacterial chemotaxis to gene expression noise [66]). This filtering function of the push-pull architecture could help explain why the *S. elongatus*

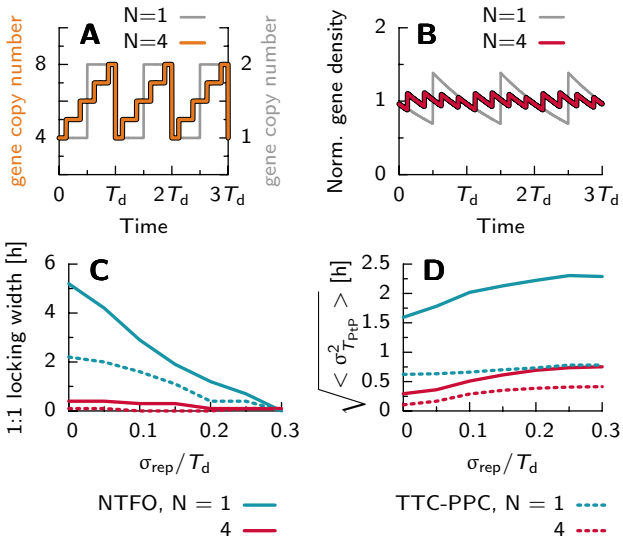

Figure 2.5: Multiple chromosome copies strongly reduce the cell cycle's effect on the circadian clock. (A) Gene copy number $g(t)$ for initial gene copy numbers $N = 4$ (thick curve, left axis) and $N = 1$ (thin curve, right axis) versus time (in units of cell cycle time T_d). The increase in $g(t)$ is more gradual for $N = 4$ than for $N = 1$. (B) The gene density $G(t) = g(t)/V(t)$, normalized to its time average, for $N = 4$ (thick curve) and $N = 1$ (thin curve). At a higher gene copy number, the deviations from the average gene density become smaller. The width of the 1:1 locking region (C) and the square root of the average variance in the peak-to-peak time (D) as a function of the standard deviation in the gene replication time σ_{rep} in a model where the times of replication events vary stochastically about their means (see *SI text*), for the NTFO (solid line) and the TTC-PPC [19] (dashed line). Increasing the chromosome copy number N reduces both the width of the locking region (C) and the variance in the peak-to-peak time (D). In contrast, while increasing σ_{dup} reduces the former, it increases the latter. See also Figs. 2.I.1 and 2.J.1.

clock has a relatively complex output mechanism requiring both CikA and SasA rather than a simpler linear design [67]. More generally, these results underscore the importance of the PPC in providing robustness to the clock [1, 19].

The second feature of the *S. elongatus* clock that we predict mitigates perturbations from the cell cycle is the presence of multiple, identical, asynchronously replicating chromosome copies [58, 62, 64, 65]. This reduces the importance of each individual gene replication event: Rather than seeing a single doubling of the number of gene copies each cell cycle, a cell with many chromosomes instead sees a number of smaller jumps that it can more easily ignore (Fig. 2.5). This adaptation may thus have evolved in part to protect the *S. elongatus* clock from cell cycle effects.

Whereas we have argued that the cell cycle generically affects any transcriptional clock, no comparably general mechanisms exist in the other direction. Moreover, though in many eukaryotic systems the clock is known to regulate key cell-cycle genes [3, 32–38, 68], no similar, specific connections have yet been characterized in *S. elongatus*. In par-

ticular, clock-dependent cell-cycle gating [3], because it acts on cell division but not on growth or DNA replication, does not allow the clock to block the discrete gene replication events that underlie the driving. Nonetheless, since the majority of *S. elongatus* genes shows some degree of clock-dependent expression [69], it is possible that the cyanobacterium's clock does regulate its cell cycle in some as yet undiscovered way. Any such coupling would however have to be weak enough to be consistent with the observation that the rhythm of DNA replication does not depend on clock phase [3, 58, 62, 64, 65]. Because phase locking between two oscillators has strong similarities to the locking of a single oscillator to periodic driving [48], most of our qualitative conclusions would remain unchanged in this case.

To isolate the behavior of the core, autonomous circadian oscillator, studies in the lab are typically performed at constant light levels. In keeping with this tradition, we have limited ourselves here to models of free-running clocks, without any diurnal environmental variation. In nature, however, the circadian clock is exposed to many additional entrainment signals, most notably the 24 h light-dark cycle. In fact, the environmental and cell cycle entrainment signals are intricately intertwined, because DNA replication and the synthesis of most proteins, including clock components, come to a standstill in the dark in a clock-independent fashion [65, 70]. We leave the effects of this complex interplay for future work.

Although we have focused on interactions between the cell cycle and the clock in *S. elongatus*, the basic idea that periodic gene replications must influence biological oscillators is more general and should apply to a wide range of prokaryotic and eukaryotic species. Indeed, cell-cycle-dependent changes in gene copy number have clearly observable effects on gene expression in eukaryotic cells [45], and recent experiments in cultured metazoan cells strongly suggest that the cell cycle exerts a considerable influence on the circadian clock, generally leading to phase locking of the two oscillators [39, 40]. Other generic forms of driving from the cell cycle may also play a role here: for example, in contrast to prokaryotes, eukaryotes typically shut down transcription around mitosis, thereby introducing another source of periodic, cell-cycle dependent variation in protein synthesis [32, 39, 40]. Our analysis thus highlights an important constraint on the design of circadian clocks in organisms from bacteria to humans.

Further, there is no reason for the effects of regular, discrete gene replications to be limited to circadian clocks; they should be observable in any cellular oscillator that depends on transcription and has a period on the same order as that of the cell cycle. Thus, our results may be relevant to phenomena like coupling between the cell cycle and the segmentation clock in vertebrate development [71]. Similarly, in chapter 3 we show that two well-known synthetic circuits [28, 29] can also lock to the cell cycle, and that the strength of locking depends sensitively on the oscillator architecture.

Since we have argued that *S. elongatus* possesses particular adaptations that decouple its circadian clock from the cell cycle, the most obvious experimental test of our ideas would be to observe the consequences of blocking or removing these features. Several strains already exist that might allow just such experiments. Mutants of *S. elongatus* are known with significantly fewer chromosomes per cell than the wildtype [72]; moreover, in some other *Synechococcus* strains, cells are always monoploid [63]. We find that in cells where the number of chromosomes goes from 1 to 2 over the course of a single di-

vision cycle, it should be possible to observe clear signatures of driving by the cell cycle in plots of KaiC's abundance—but not its phosphorylation level—as a function of time (Fig. 2.4). We predict that this effect will be further strengthened if the PPC is removed entirely. It is well-established that this can be accomplished by hyper-phosphorylating KaiC [53, 73]. In all cases, one could study forcing by the cell cycle at a variety of different doubling times. We suggest, however, that a doubling time near 48 hours offers a particularly unambiguous signature of the cell cycle's influence: The KaiC abundance as a function of time should then rise and fall every 24 hours, with successive peaks strictly alternating between higher and lower levels (Fig. 2.4C).

APPENDIX

2.A. OVERVIEW OF MODELS AND NOMENCLATURE

This Supporting Information (SI) discusses three different models of a circadian clock driven by periodic gene replication, which we list here to summarize our naming conventions and the distinctions among the models:

- The Negative Transcriptional Feedback Oscillator, or **NTFO**, consists of a single gene that negatively regulates its own transcription with a delay. Because of gene replication, the number of copies of this gene in each cell doubles over the course of a cell cycle (as described in the next section and in the main text).
- We refer to our central model of the *S. elongatus* circadian clock as either the **TTC-PPC** model or the **TTC-(PPC$_{\text{Zwicker}}$)** model, with the latter name reserved for those parts of this *SI* where there is potential for confusion with the **TTC-(PPC$_{\text{Rust}}$)** model (next item). The **TTC-(PPC$_{\text{Zwicker}}$)** model consists of the protein phosphorylation cycle (PPC) model of Van Zon *et al.* [15] joined to a transcription-translation cycle (TTC) as in the work of Zwicker *et al.* [19] and subject to periodic driving from the cell cycle because of gene replication, as discussed in the main text.
- We also present results for the **TTC-(PPC$_{\text{Rust}}$)** model, which couples the description of the PPC proposed by Rust *et al.* [22] to the same models of the TTC and of forcing from the cell cycle used in the **TTC-(PPC$_{\text{Zwicker}}$)** model.

Of these models, the first two (NTFO and TTC-(PPC$_{\text{Zwicker}}$)) are discussed in the main text, while the third (TTC-(PPC$_{\text{Rust}}$)) is introduced in this Appendix, to demonstrate that our major qualitative conclusions do not depend on our specific assumptions about the PPC.

2.B. THE DETERMINISTIC NEGATIVE TRANSCRIPTIONAL FEEDBACK OSCILLATOR (FIGS. 2.2 AND 2.4)

Here we describe the negative transcriptional feedback oscillator, NTFO, studied in the main text, together with its parameters. The model consists of a single variable, $C(t)$, describing the concentration of proteins that inhibit their own production:

$$\frac{dC(t)}{dt} = \beta \tilde{G}(t) \frac{K_c^n}{K_c^n + C(t-\Delta)^n} - \mu_{\text{tot}} C(t). \tag{2.3}$$

We impose a fixed delay Δ between the initiation of transcription and the appearance of functional proteins. Therefore, protein production at time t is proportional to the gene copy number $g(t-\Delta)$ at time $t-\Delta$. These proteins enter the cell volume $V(t)$ at time t.

Combining these two effects, the protein synthesis rate per unit volume at time t is thus proportional to the *protein production density* $\tilde{G}(t) \equiv g(t-\Delta)/V(t)$. Because the production at time t depends on the state of the promoter at $t-\Delta$, the Hill function describing auto-regulation with Hill coefficient n and concentration of half-maximal repression K_c, is evaluated with the delayed concentration $C(t-\Delta)$. Proteins are degraded with a total rate $\mu_{\rm tot} = \mu_{\rm d} + \mu_{\rm act}$, where $\mu_{\rm d}$ describes dilution due to cell growth, and $\mu_{\rm act}$ describes possible active degradation. Including both terms allows us to vary the doubling time $T_{\rm d}$ while holding $\mu_{\rm tot}$ constant and hence to distinguish the trivial influence of the cell cycle on the clock through the dilution rate $\mu_{\rm d}$ from other effects. This model is a deterministic one, based on mean-field chemical rate equations. A stochastic version that takes into account the intrinsic stochasticity of biochemical reactions is introduced further below.

Parameters used in the simulations are: $\beta = 6.0 \cdot 10^3 \mathrm{h}^{-1}$; $K_c = 1.0\mu\mathrm{M}$; $n = 2$; $\mu_{\rm tot} = 0.2\mathrm{h}^{-1}$; $\Delta = 8\mathrm{h}$. Details regarding the simulations are given in the Methods section of this document. The principal results of this model are presented in Figs. 2 and 4 of the main text.

2.C. THE TTC-(PPC$_{\rm Zwicker}$) MODEL (FIGS. 2.4 AND 2.5)

The TTC-PPC model of the main text is built on the TTC-PPC model of Zwicker *et al.*. It consists of a protein phosphorylation cycle (PPC) combined with a transcription-translation cycle (TTC) [19]. The PPC model is based on that of Van Zon *et al.* [15]. In this model, each KaiC hexamer has an intrinsic tendency to progress through a phosphorylation cycle, while the phosphorylation cycles of the individual hexamers are synchronized via the mechanism of differential affinity: KaiA stimulates KaiC phosphorylation, but the limited supply of KaiA binds preferentially to those KaiC hexamers that are falling behind in the cycle, forcing the front runners to slow down and allowing the laggards to catch up. The model includes the following reactions [15]:

$$\mathrm{C}_i \underset{b_i}{\overset{f_i}{\rightleftarrows}} \widetilde{\mathrm{C}}_i, \qquad \mathrm{C}_i + \mathrm{A} \underset{k_i^{\rm Ab}}{\overset{k_i^{\rm Af}}{\rightleftarrows}} \mathrm{AC}_i \overset{k_{\rm pf}}{\to} \mathrm{C}_{i+1} + \mathrm{A}, \tag{2.4}$$

$$\widetilde{\mathrm{C}}_i + \mathrm{B} \underset{\tilde{k}_i^{\rm Bb}}{\overset{2\tilde{k}_i^{\rm Bf}}{\rightleftarrows}} \mathrm{B}\widetilde{\mathrm{C}}_i, \qquad \mathrm{B}\widetilde{\mathrm{C}}_i + \mathrm{B} \underset{2\tilde{k}_i^{\rm Bb}}{\overset{\tilde{k}_i^{\rm Bf}}{\rightleftarrows}} \mathrm{B}_2\widetilde{\mathrm{C}}_i, \tag{2.5}$$

$$\mathrm{B}_x\widetilde{\mathrm{C}}_i + \mathrm{A} \underset{\tilde{k}_i^{\rm Ab}}{\overset{x\tilde{k}_i^{\rm Af}}{\rightleftarrows}} \mathrm{AB}_x\widetilde{\mathrm{C}}_i, \qquad \mathrm{AB}_2\widetilde{\mathrm{C}}_i + \mathrm{A} \underset{2\tilde{k}_i^{\rm Ab}}{\overset{\tilde{k}_i^{\rm Af}}{\rightleftarrows}} \mathrm{A}_2\mathrm{B}_2\widetilde{\mathrm{C}}_i, \tag{2.6}$$

$$\mathrm{C}_i \underset{k_{\rm dps}}{\overset{k_{\rm ps}}{\rightleftarrows}} \mathrm{C}_{i+1}, \qquad \widetilde{\mathrm{C}}_i \underset{\tilde{k}_{\rm dps}}{\overset{\tilde{k}_{\rm ps}}{\rightleftarrows}} \widetilde{\mathrm{C}}_{i+1}, \tag{2.7}$$

$$\mathrm{B}_x\widetilde{\mathrm{C}}_i \underset{\tilde{k}_{\rm dps}}{\overset{\tilde{k}_{\rm ps}}{\rightleftarrows}} \mathrm{B}_x\widetilde{\mathrm{C}}_{i+1}, \qquad \mathrm{A}_y\mathrm{B}_x\widetilde{\mathrm{C}}_i \underset{\tilde{k}_{\rm dps}}{\overset{\tilde{k}_{\rm ps}}{\rightleftarrows}} \mathrm{A}_y\mathrm{B}_x\widetilde{\mathrm{C}}_{i+1}\,. \tag{2.8}$$

Here, C_i denotes a KaiC hexamer in the active conformational state, in which the number i of phosphorylated monomers tends to increase, and $\widetilde{\mathrm{C}}_i$ denotes a KaiC hexamer in the inactive conformational state in which i tends to decrease; A denotes a KaiA dimer,

and B denotes a KaiB dimer. The reactions $C_i \rightleftarrows \widetilde{C}_i$ in Eq. 2.4 model the conformational transitions between active and inactive KaiC; the second set of reactions in Eq. 2.4 describe phosphorylation of active KaiC that is stimulated by KaiA. The reactions in Eq. 2.5 model the binding of KaiB to inactive KaiC, and those in 2.6 model the sequestration of KaiA by inactive KaiC that is bound to KaiB; note that in this model an inactive KaiC hexamer can bind up to two KaiA dimers. The reactions in Eqs. 2.7 and 2.8 model spontaneous phosphorylation and dephosphorylation of active and inactive KaiC. For a more detailed discussion of the model, we refer to [15].

The TTC and the coupling between the PPC and TTC are described by the following reactions:

$$\widetilde{R} + X \xrightarrow{k_a} R + X, \qquad R + \widetilde{X} \xrightarrow{k_i} \widetilde{R} + \widetilde{X}, \tag{2.9}$$

$$\emptyset \overset{\beta \tilde{G}(t)[R]^4/(K^4+[R]^4)}{\underset{\tau}{\Longrightarrow}} C_6 + 3B, \qquad \emptyset \xrightarrow{\beta_a} A,\ \emptyset \xrightarrow{\beta_r} \widetilde{R}, \tag{2.10}$$

$$R, \widetilde{R}, A, B, AC_i, B_x\widetilde{C}_i, A_yB_x\widetilde{C}_i \xrightarrow{\mu} \emptyset\,. \tag{2.11}$$

Here, R and $\widetilde{R}$ denote the RpaA protein in its active and its inactive form, respectively. The X and $\widetilde{X}$ in Eq. (2.9) denote any of the phosphoforms of KaiC that mediate the activation and repression of RpaA, respectively; as described in more detail in [19], KaiC in the phosphorylation phase activates RpaA, while KaiC in the dephosphorylation phase deactivates it. The double arrow indicates a reaction with a fixed delay τ; we use τ rather than Δ to denote the delay to agree with the notation of [19]. We thus assume that *kaiBC* expression is activated by RpaA and that the activity of RpaA is modulated by the PPC. In contrast, the expression of KaiA and RpaA is taken to occur constitutively. The effect of gene replication is included by making the rate of C_6 production dependent on $\tilde{G}(t)$, just as in the NTFO. We use the parameters given in the supplementary information of [19], except for: $\beta = 1.02 \cdot 10^2 \mathrm{h}^{-1}$ and $\mu_{\mathrm{tot}} = 0.1\mathrm{h}^{-1}$. As for the NTFO, we keep the total degradation rate constant, so that we can distinguish the effects of protein dilution from those due specifically to periodic gene replication events.

2.D. THE TTC-($\mathrm{PPC_{Rust}}$) MODEL (FIG. 2.K.1)

We argue in the main text that entrainment of circadian clocks by the cell cycle should be a robust, generic phenomenon that does not depend on the precise model studied. In support of this claim, we here present a model of the Kai system, the TTC-($\mathrm{PPC_{Rust}}$) model, that combines the TTC of Zwicker *et al.* [19] with a model of the PPC developed by Rust and coworkers [22]; this model is here extended to include a description of the cell cycle with periodic gene replication. The PPC model of Rust *et al.* describes the PPC at the level of KaiC monomers, rather than of hexamers, as in the models of Van Zon *et al.* and Zwicker *et al.* considered in the preceding section and in the main text [15, 19]. The monomers go through a sequence of 4 different phosphorylation states —unphosphorylated (U), phosphorylated only on threonine 432 (T), doubly phosphorylated (D), and phosphorylated only on serine 431 (S)—in a 24 hour cycle [60]. As detailed in the supplementary information of [19], we extended the original Rust model, which contains only the PPC, to include a transcription-translation cycle. Here, we briefly describe the model and discuss how it responds to forcing from the cell cycle.

This PPC is described by the reactions

$$\mathrm{U} \leftrightarrow \mathrm{T}, \qquad \mathrm{T} \leftrightarrow \mathrm{ST}, \qquad \mathrm{ST} \leftrightarrow \mathrm{S}, \qquad \mathrm{S} \leftrightarrow \mathrm{U} \tag{2.12}$$

with reaction rates given by Eq. 5 of the supplementary material of Rust *et al.* [22]. These rates depend on the concentration of free KaiA, which is sequestered by KaiC in the S-state. We model KaiA sequestration explicitly:

$$\mathrm{S} + \mathrm{A} \leftrightarrow \mathrm{AS}, \qquad \mathrm{AS} + \mathrm{A} \leftrightarrow \mathrm{A_2S}. \tag{2.13}$$

Dephosphorylation of KaiC in the S-state is allowed to occur even when KaiA is bound, in which case the KaiA protein is released from the complex. We define the output signal as

$$p(t) = \frac{[\mathrm{T}] + [\mathrm{ST}] + [\mathrm{S}] + [\mathrm{AS}] + [\mathrm{A_2S}]}{[\mathrm{U}] + [\mathrm{T}] + [\mathrm{ST}] + [\mathrm{S}] + [\mathrm{AS}] + [\mathrm{A_2S}]}, \tag{2.14}$$

which resembles the phosphorylation ratio in the case where we cannot distinguish between singly and doubly phosphorylated KaiC. The denominator in the above expression is also the total KaiC monomer concentration. We use the same total protein concentrations as Rust *et al.*, [KaiA] = 1.3 μM (active KaiA monomers) and [KaiC] = 3.4 μM (KaiC monomers).

The TTC is modeled as:

$$\emptyset \overset{\beta\tilde{G}(t)[\mathrm{R}]^4/(K^4+[\mathrm{R}]^4)}{\underset{\tau}{\Longrightarrow}} \mathrm{ST}, \quad \emptyset \xrightarrow{\beta_a} \mathrm{A}, \emptyset \xrightarrow{\beta_r} \widetilde{\mathrm{R}}, \tag{2.15}$$

$$\widetilde{\mathrm{R}} + \mathrm{T} \xrightarrow{k_a} \mathrm{R} + \mathrm{T}, \quad \mathrm{R} + \mathrm{A}_x\mathrm{S} \xrightarrow{k_i} \widetilde{\mathrm{R}} + \mathrm{A}_x\mathrm{S}, \qquad x \in \{0,1,2\}, \tag{2.16}$$

$$\mathrm{A}, \mathrm{R}, \widetilde{\mathrm{R}}, \mathrm{U}, \mathrm{S}, \mathrm{A}_x\mathrm{S}, \mathrm{T}, \mathrm{ST} \xrightarrow{\mu} \emptyset. \tag{2.17}$$

The first line describes protein production and decay; as in Eq. 2.11, the double arrow indicates a delayed reaction. The second line summarizes the RpaA signaling pathway, where KaiC that is phosphorylated at the T site activates RpaA and KaiC that is phosphorylated at the S site represses RpaA activation.

As in our other models, we introduced the effect of gene replication by making the production rate of KaiC proportional to the production density $\tilde{G}(t)$. We use the parameters given in [22] and in section S5.2 of the supplementary information of [19], except for $\beta = 1.29 \cdot 10^3 \mathrm{h}^{-1}$ and $\mu_{\mathrm{tot}} = 0.1 \mathrm{h}^{-1}$.

Fig. 2.K.1 shows the peak-to-peak times of the phosphorylation fraction of this model as a function of the cell-division time T_d. Clearly, this TTC-(PPC$_{\mathrm{Rust}}$) model, which combines a TTC with a PPC based on the model of Rust *et al.* [22], responds to forcing from the cell cycle in essentially the same way as the TTC-(PPC$_{\mathrm{Zwicker}}$) model [19], which couples the same TTC with a PPC based on the model of Van Zon and Zwicker *et al.* [15, 19]. This supports the idea that the benefit of a self-contained protein-modification oscillator is a generic feature of biological clocks.

2.E. THE NEGATIVE TRANSCRIPTIONAL FEEDBACK OSCILLATOR WITH INTRINSIC NOISE (FIG. 2.E.1)

To test the effect on the locking mechanism of intrinsic noise in the chemical reactions that constitute the circadian clock, we simulate the negative transcriptional feedback oscillator (NTFO) using kinetic Monte Carlo simulations of the chemical master equation. Eq. 2.2 of the main text implies that when this intrinsic noise is neglected, the number of proteins N_C in the cell obeys

$$\frac{dN_C(t)}{dt} = \beta g(t-\Delta)\frac{K_C^2}{K_C^2 + \left(\frac{N_C(t-\Delta)}{V(t)}\right)^2} - \mu_{act} N_C(t). \tag{2.18}$$

Here, $V(t)$ is the cell volume, which grows exponentially as described in the main text, and μ_{act} is the active contribution to the total decay rate $\mu_{tot} = \mu_{act} + \mu_d$ in Eq. 2; the term μ_d contributes to the total apparent decay rate for the protein concentration, but not for the protein number. To include the effects of intrinsic noise on the evolution of N_C, we adapted the standard kinetic Monte Carlo algorithm to take into account delayed reactions, volume growth, and gene replication, as described in [19]: The cell volume is increased at discrete time intervals, and reaction propensities are re-calculated after each volume update. After each cell division, the protein number N_C is chosen from a binomial distribution, with N_C halved on average. Gene replication is included through the time dependence of the gene copy number $g(t)$, whose behavior is detailed in the main text.

The parameters used in the simulation are: $\beta = 6.0{\cdot}10^3 \mathrm{h}^{-1}$; $K_c = 1.0\mu$M; $n = 2$; $\Delta = 8$h. The time average $\bar{V}$ of the cell volume is chosen to be $1\mu\mathrm{m}^3$. For different cell division times T_d, we change the active degradation rate such that the total decay rate is kept constant at $0.2\mathrm{h}^{-1}$: $\mu_{act} = 0.2\mathrm{h}^{-1} - \log(2)/T_d$. As in the main text, this allows us to disentangle the effect of gene replication on locking from that of simple protein dilution.

We vary the cell division time T_d from 6 to 52 hours in 0.1 hour intervals and simulate a single trajectory of 10,000 hours for each T_d. From the trajectories, we extract the peak-to-peak times of the oscillations (see Methods in this document). Fig. 2.E.1 shows the average peak-to-peak time of the stochastic NTFO model for different values of the cell division time, with initial gene copy number $N = 1$. Two differences from Fig. 2.2 of the main text, where the system obeys deterministic equations, are worthy of note: First, the 1:1 locking region is much larger when intrinsic noise is included, while the width of the 2:1 locking region has decreased. Intrinsic noise can thus dramatically change the extent of locking. We leave a full investigation of the origins of this effect for future work. Second, because of intrinsic noise, the amplitude is considerably more variable, and the variances in the peak-to-peak times outside of the locking regions are much larger. As in the mean-field model, the variances are very small around $T_d = 24$ hrs, due to locking.

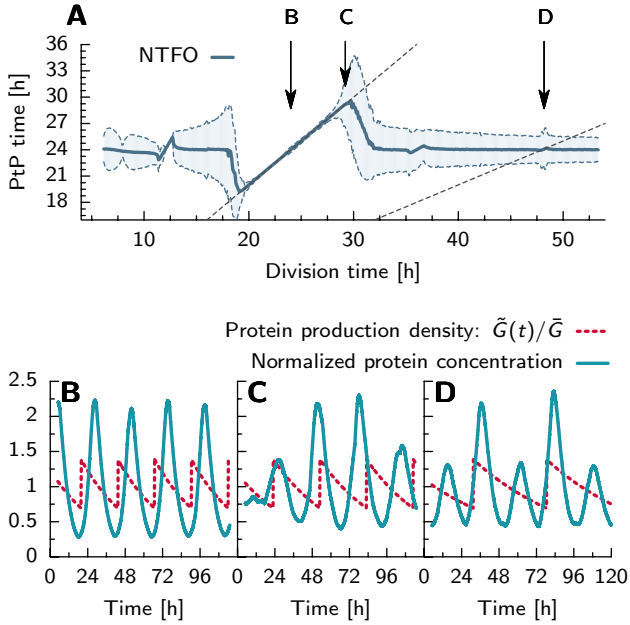

Figure 2.E.1: The effect of intrinsic noise on the locking of the negative transcriptional feedback oscillator (NTFO) to the cell cycle. (A) Average (solid line) and standard deviation (shaded region) of the peak-to-peak time T_{PtP} as a function of the division time T_{d} for an NTFO with intrinsic noise and initial gene copy number $N = 1$. The region of 1:1 locking with the cell cycle (left dashed line) has widened considerably compared to the deterministic case (Fig. 2.2A of the main text), and the standard deviation in T_{PtP} outside the locking region has increased. In contrast, the region of 2:1 locking (right red dashed line) has shrunk almost to nothing. (B–D) Representative time traces for the division times indicated by the arrows in panel A. Shown are the protein concentration $C(t) = N_C(t)/V(t)$ of the NTFO (blue solid line) and the protein production density $\tilde{G}(t)$ (red dashed line), both normalized by their time average values. At a division time of $T_{\mathrm{d}} = 24$h (B), the NTFO is locked to the cell cycle. Because of the intrinsic noise, the amplitude varies slightly from one oscillation cycle to the next. At $T_{\mathrm{d}} = 27$ h (C), just outside the locking region, the oscillator exhibits irregular behavior. At $T_{\mathrm{d}} = 48$ h (D), the NTFO oscillations switch between a small and a large amplitude in successive oscillation cycles, just as in the deterministic case.

2.F. THE EFFECT OF NOISE IN THE TIMING OF GENE REPLICATION EVENTS (FIG. 2.F.1)

In the model of the cell cycle described in the main text, gene replication events occur at perfectly regular, evenly spaced intervals in time. In this section, we explore how stochastic variability in the timing of these replications affects the interaction between the cell cycle and other oscillators. We consider a model in which each gene replicates exactly once in each cell division cycle, at a time t_{g} between 0 and T_{d}, where T_{d} is the length of the cell cycle and does not vary. The times t_{g} are drawn in each cell cycle independently from a Gaussian distribution of mean $\bar{t}_{\mathrm{rep}} = T_{\mathrm{d}}/2$ and variance σ^2_{rep}.

Replication times t'_g that fall outside the interval $[0, T_d)$, are mapped back onto it via $t_g = \mathrm{Mod}(t'_g, T_d)$.

We assume that the standard deviation in the gene replication times, σ_{rep}, is proportional to the cell division time: $\sigma_{rep} \propto T_d$. Because this quantity is, to our knowledge, not known for *S. elongatus*, we varied this quantity between zero and a value that corresponds to replication times being chosen randomly from a uniform distribution, $\sigma_{rep}/T_d = 1/\sqrt{12} \approx 0.3$.

Fig. 2.F.1 shows the effects of introducing this variability on the behavior of our clock models. For the NTFO (Fig. 2.F.1A), the locking regions are reduced in size, but still clearly noticeable. The variance in the peak-to-peak times is typically larger than in deterministic case. For the TTC-($\mathrm{PPC}_{\mathrm{Zwicker}}$) model [19] (Fig. 2.F.1B), the locking regions have almost disappeared, but the variance in T_{PtP} is nearly unaffected by the noise in the timing of gene replication. Panel C shows representative time traces of the total KaiC concentration, $C_{tot}(t)$ and the phosphorylation fraction $p(t)$, for $T_d = 48$h. Clearly, $C_{tot}(t)$ shows much more variability in the height of its peaks than does $p(t)$, further demonstrating the value of a post-translational oscillator in insulating a circadian clock from influences from the cell cycle. We also note that, as in the deterministic limit, the amplitude of the C_{tot} oscillation cycles still (despite the noise) alternates between a high and a low value when gene replications occur only once every 48 hours. It should thus be possible to observe experimentally that $C_{tot}(t)$ is markedly affected by periodic gene replication, while $p(t)$ is much less so.

2.G. ALLOWING THE PROTEIN DECAY RATE TO VARY WITH GROWTH RATE (FIG. 2.G.1)

The total protein decay rate μ_{tot} depends on the rate of active protein degradation μ_{act} and on the rate of dilution due to cell growth μ_d. In the results shown in the main text, we adjusted the active degradation rate with the growth rate so that the total protein decay rate remained constant—this allowed us to zoom in on the effect of locking that is due to periodic gene replication while ignoring confounding changes to clock behavior that might arise because of the variation of μ_{tot}.

It is entirely possible, however, that the real clock system in *S. elongatus* is actually closer to the opposite limit, in which μ_{act} is fixed and μ_d and μ_{tot} vary together with the division time T_d. To investigate how this affects the locking mechanism, we performed simulations in which we kept the active degradation rate μ_{act} constant, but allowed the total degradation rate μ_{tot} to vary with the cell-divisition time: $\mu_{tot} = \mu_{act} + \ln(2)/T_d$; upon varying μ_{tot}, we adjusted the protein synthesis rate β to keep the oscillation period at 24 h. The result is shown in Fig. 2.G.1. Just as in the case in which μ_{tot} is fixed, the effect of locking is very pronounced, both for a simple NTFO (panel A) and for a clock incorporating both a TTC and a PPC (panel B).

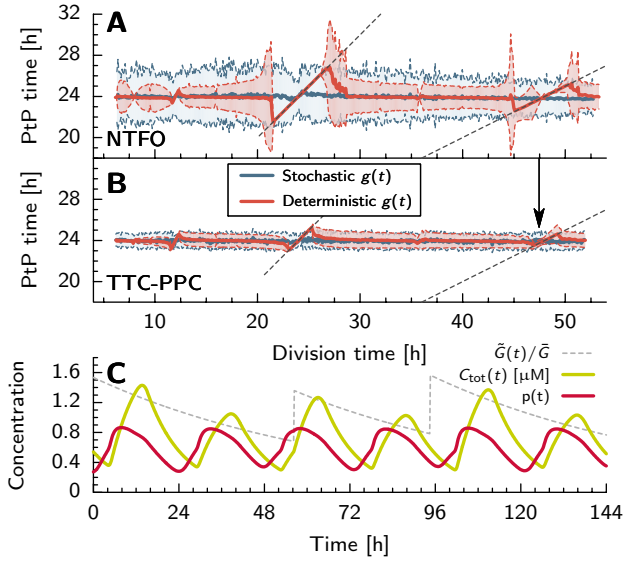

Figure 2.F.1: The effect of stochasticity in the timing of gene replication on locking to the cell cycle. (A) The average and variance of T_{PtP} for an NTFO in which the timing of gene replication is deterministic (red) or is drawn from a Gaussian distribution with a width σ that is 30% of the cell-division time T_{d}(blue). The stochasticity in the replication times decreases the width of the locking regions, but increases the variance in the peak-to-peak times. (B) Same as A, but for T_{PtP} of the phosphorylation fraction $p(t)$ of the TTC-(PPC$_{\mathrm{Zwicker}}$) model [19] (see *SI* section *The TTC-(PPC$_{\mathrm{Zwicker}}$) model*)[19]. Again, stochasticity in the timing of replication reduces locking, but in this case, the increase in the variance of T_{PtP} outside the locking region is much less marked. We attribute this to the ability of the PPC to insulate the clock from variability in gene expression levels. (C) Representative time traces of the production density $\tilde{G}(t)$ (normalized to its time average), the phosphorylation fraction $p(t)$, and the total KaiC concentration $C_{\mathrm{tot}}(t)$ for the Zwicker model for T_{d}=48h. As in the deterministic limit, the amplitude of the $C_{\mathrm{tot}}(t)$ oscillations tends to alternate between a high and a low value, due to gene replication occuring every 48 hours, on average; in contrast, the amplitude of $p(t)$ is relatively constant. The effect of periodic gene replication on $C_{\mathrm{tot}}(t)$ should thus be observable experimentally.

2.H. HOW DOES THE CELL READ OUT THE PHOSPHORYLATION FRACTION INSTEAD OF THE CONCENTRATION OF KAIC? (FIG. 2.H.1)

Both Fig. 2.4 and a comparison of the blue lines in Figs. 2.F.1A,B show that a circadian clock with a post-translational oscillator is much more robust to a time-varying gene density than is an NTFO. Both the peak-to-peak time and the amplitude of the phosphorylation fraction $p(t)$ of the TTC-PPC vary considerably less not just than the corresponding quantities for the NTFO, but even than those for the total protein concentration $C_{\mathrm{tot}}(t)$ of the TTC-PPC itself. To take advantage of the relative stability of the

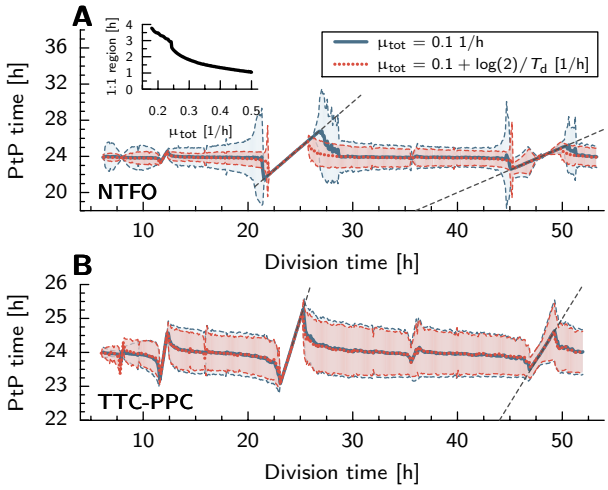

Figure 2.G.1: The effect of allowing the total protein decay rate μ_{tot} to vary with the cell-division time T_{d}, both for an NTFO (A) and for the TTC-(PPC$_{\mathrm{Zwicker}}$) model [19] of a clock, which combines a TTC with a PPC (B; note the vertical scale is different than in panel A). The blue lines correspond to the scenario in which the the total degradation rate is kept constant at $\mu_{\mathrm{tot}} = 0.1/\mathrm{h}$ (as in Figs. 2A and 4A of the main text), and the red lines to the scenario in which the total degradation rate depends on the division time as $\mu_{\mathrm{tot}} = 0.1\mathrm{h}^{-1} + \log(2)/T_{\mathrm{d}}$ [1/h]. When μ_{tot} depends on T_{d}, we adjust the KaiC production rate β such that the intrinsic period of the clock remains 24 hours. For both clock models and both choices of μ_{tot}, the clock tends to lock to the cell cycle. The difference between the results of the two protein-decay scenarios in the case of the NTFO can be understood by noticing that we have chosen our rates so that μ_{tot} never drops below 0.1 h^{-1} in either case, but can become larger than this bound when it is allowed to depend on T_{d}. The protein synthesis rate β then becomes higher than when μ_{tot} is constant. The higher synthesis and decay rate raises the amplitude of the concentration oscillations, making the clock more stable. As the inset indicates, the width of the locking region decreases with increasing μ_{tot} in a similar fashion when μ_{tot} does not depend on T_{d}. The TTC-PPC is almost insensitive to the higher degradation and production rates.

oscillation in $p(t)$, the cell needs to read out the phosphorylation fraction in a way that is insensitive both to the total concentration of KaiC and to the absolute concentrations of its specific phosphorylation states. How does it accomplish this?

To find out, we looked at the architecture of the biochemical network in *S. elongatus* that allows the clock to regulate the transcription of downstream genes. The temporal information encoded in the dynamics of the clock proteins is transmitted via a central node, the response regulator RpaA. This protein can be phosphorylated, and it is known that the phosphorylation level of RpaA controls the expression not only of core clock components, but also of a small set of genes that, in turn, direct genome-wide circadian rhythms [54, 57]. The phosphorylation state of RpaA is regulated by a push-pull network consisting of the histidine kinases SasA, which acts primarily to phosphorylate RpaA, and CikA, whose primary function is to dephosphorylate it [54–56]. The activi-

ties of SasA and CikA are in turn controlled by the different phosphorylation states of KaiC, such that KaiC proteins that are in the phosphorylation phase of the clock tend to push up the phosphorylation level of RpaA [27, 54], while those that are in the clock's dephosphorylation phase tend to pull down RpaA's phosphorylation level [27, 55]. Recent experiments suggest that a structural change in KaiB, which binds KaiC during its dephosphorylation phase, might play a key role in this switch, by simultaneously blocking SasA from binding to KaiC and engaging CikA [27].

To show that such a push-pull network makes the clock readout, the RpaA phosphorylation level, sensitive to the KaiC phosphorylation fraction $p(t)$, but not to its concentration $C_{\mathrm{tot}}(t)$, we adapted the canonical model of Goldbeter and Koshland [61]. A cartoon of the model is shown in Fig. 2.H.1B. It consists of a substrate S, playing the role of RpaA, which can be phosphorylated and dephosphorylated by the antagonistic enzymes K and P, corresponding respectively to SasA and CikA, each of which can be in active (K^* and P^*) or inactive states. Transitions between these states are governed by the time-dependent forward rates $k_K(t)$ and $k_P(t)$, which mimic the effects of time-varying concentrations of different KaiC phosphoforms. In our model, KaiC in the phosphorylation phase increases $k_K(t)$, while KaiC in the dephosphorylation phase increases $k_P(t)$. The substrate-modification reactions follow the standard Michaelis-Menten schemes $K^* + S \rightleftharpoons K^*S \rightarrow K^* + S_p$ and $P^* + S_p \rightleftharpoons P^*S_p \rightarrow P^* + S$.

The key question is whether the phosphorylation level of RpaA depends only on the phosphorylation fraction of KaiC, $p(t)$, or whether it is also significantly affected by the total concentration of KaiC. In terms of the model just presented, we must thus ask whether $[S_p]/[S_{\mathrm{tot}}]$ (the fraction of phosphorylated RpaA) is sensitive only to the *ratio* $k_K(t)/k_P(t)$ or whether it depends on the two rates individually. We begin by considering this question for the case of time-independent rates, in which case the ratio of total concentrations of the active enzyme forms $[K^*]/[P^*]$ is proportional to $k_K(t)/k_P(t)$ and may be used in its place. In qualitative terms, it is easy to imagine that, since K^* and P^* have opposite effects, increasing the concentration of each one by the same factor might speed up the reaction kinetics but would not affect the steady state ratio $[S_p]/[S_{\mathrm{tot}}]$. Indeed, in the regime that the concentrations of the complexes K^*S and P^*S_p are negligible compared to the concentrations of free S and S_p, we can equate the rates of S phosphorylation and dephosphorylation in steady state to arrive at a relation of the form

$$\frac{k_+[K^*][S]}{K_{M,+}+[S]} = \frac{k_-[P^*][S_p]}{K_{M,-}+[S_p]}\,. \tag{2.19}$$

Together with the conservation law $[S]+[S_p]=[S_{\mathrm{tot}}]$, this equation can easily be solved to give $[S_p]/[S_{\mathrm{tot}}]$ as a function of $[K^*]/[P^*]$ only. It is possible, however, that in the RpaA system the concentrations of the intermediate complexes K^*S and P^*S_p cannot be neglected. Fig. 2.H.1A shows that even in this case, to a very good approximation the steady state phosphorylation fraction of S depends only on the ratio $[K^*]/[P^*]$. This mechanism, by which the output of a push-pull network depends on the ratio of the concentrations of the two antagonistic enzymes, but not on their absolute values, has previously been invoked to explain the robustness of the *E. coli* chemotaxis pathway to concerted variations in the expression levels of the chemotaxis proteins [66].

To extend these results to the case of time-varying activation rates that is more di-

rectly relevant to a situation in which the concentrations of the KaiC phosphoforms rise and fall, we allow $k_K(t)$ and $k_P(t)$ to vary with time as shown in Fig. 2.H.1C (upper graph): Similarly to the total KaiC concentration at $T_d = 48$ hours with $N = 1$ (Fig. 2.4C of the main text), each rate changes periodically with a period of 48 hours, with alternating higher and lower peaks. The lower graph of Fig. 2.H.1C , with time traces of $[S_p]$, $[K^*]$ and $[P^*]$, shows that even though the amplitude of $[K^*](t)$ and $[P^*](t)$ varies between oscillation cycles, the amplitude of $S_p(t)$ is constant. Comparable behavior is expected as long as the rates associated with the enzymatic reactions are faster than the timescale of variation of $k_K(t)$ and $k_P(t)$. Provided this is the case, the push-pull system built around RpaA will allow the bacterium to take as its clock readout the phosphorylation fraction $p(t)$—which we have argued is robust to perturbations associated with the cell cycle—rather than the absolute concentration of KaiC or one of its phosphoforms.

2.I. MULTIPLE CHROMOSOMES REDUCE THE COUPLING BETWEEN CLOCK AND CELL CYCLE IN TTC-PPC AND NTFO MODELS (FIG. 2.I.1)

Fig. 2.5C of the main text shows the average peak-to-peak time and its standard deviation for the TTC-($PPC_{Zwicker}$) model [19], as a function of the cell-division time T_d, both for $N = 1$ and for $N = 4$. Panel A of Fig. 2.I.1 shows exactly the same figure, but with underneath it, for $N = 4$, time traces for $p(t)$ and $C_{tot}(t)$ at $T_d = 24.5$ h (panel B) and $T_d = 48$ h (panel C), as also indicated by the arrows in panel A. As discussed in the main text, the effects of the cell cycle have almost completely disappeared, even at $T_d = 24.5$ h, immediately outside the locking regime, where the effects are usually most visible.

Panels D–G of Fig. 2.I.1 show the effect of multiple chromosomes on the locking behavior of the NTFO model. Panel D shows that with $N = 4$ chromosome copies at the beginning of the cell cycle, both the variance and the width of the locking regions have decreased to no more than one hour. The time traces (Fig. 2.I.1E–G) of the protein concentration at cell division times of 24, 27 and 48 hours confirm that the NTFO has indeed become very stable. At $T_d = 27$ hours, where the protein concentration had showed irregular behavior and large amplitude variations for $N = 1$, its behavior has now become much more regular. At a division time of 48 hours, the marked amplitude variations present when $N = 1$ have disappeared. Indeed, the beneficial effect of a higher gene-copy number is larger than that of adding a PPC, as can be seen by comparing Fig. 2.I.1D with Fig. 2.4 of the main text. Fig. 2.J.1 shows that essentially the same effect is also found in the TTC-($PPC_{Zwicker}$) model. Of course, the most stable clock is obtained when a higher gene copy number is combined with a PPC (see Fig. 2.5), suggesting that both are needed for circadian rhythms that are maximally resilient against perturbations from the cell cycle.

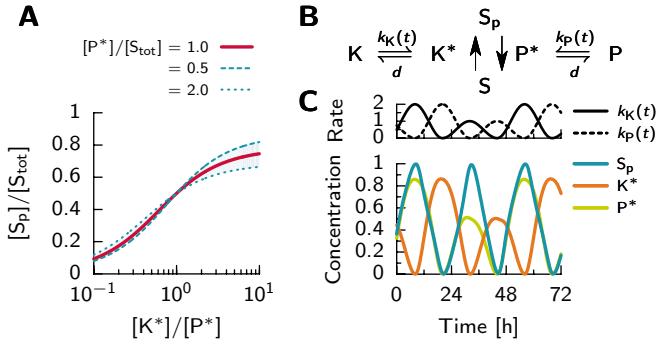

Figure 2.H.1: A push-pull network can read out the phosphorylation fraction $p(t)$ while remaining insensitive to the total concentration $C_{\mathrm{tot}}(t)$ of KaiC. (A) Steady-state output of the push-pull network, the fraction of phosphorylated substrate $[S_p]/[S_{tot}]$, plotted against the ratio of the active kinase concentration, $[K^*]$, to the active phosphatase concentration $[P^*]$; here, $[S_p]/[S_{tot}]$ mimics the phosphorylation fraction of RpaA. In steady state (but not necessarily in the general, time-varying case, see panel B) $[K^*]$ directly reports the concentration of KaiC in the phosphorylation phase of the clock and $[P^*]$ the concentration of KaiC in the clock's dephosphorylation phase. For the solid red line, we change $[K^*]/[P^*]$ from 0.1 to 10, while keeping $[P^*]$ equal to $[S_{tot}]$. The dashed and dotted lines show the result when both the kinase and phosphatase concentrations are halved or doubled, respectively. Because of the push-pull architecture, a change in the total concentration $[K^*]+[P^*]$ at fixed $[K^*]/[P^*]$ has only a small effect on the steady-state level of phosphorylated substrate $[S_p]/[S_{tot}]$; the network predominantly responds to the ratio $[K^*]/[P^*]$. (B) Schematic of our model of a simple push-pull network. The amount of active kinase, $[K^*]$, is controlled by the time-dependent rate $k_K(t)$ of conversion from K to K*, and similarly for $[P^*]$ and $k_P(t)$; we imagine that these rates are proportional to the amount of KaiC in the phosphorylation and dephosphorylation phases of the clock, respectively. The two enzymes return to their inactive states at constant rates d. The interconversion between S and S_p follows the standard Michaelis-Menten reaction scheme. (C) Reading out time-varying rates k_K and k_P. (Top) We let $k_K(t)$ oscillate with a peak-to-peak time of 24 hours, with the amplitude of each consecutive oscillation cycle changing by a factor of two to mimic the variability in the total amount of KaiC when $T_d = 48$ h (Fig. 2.2 of the main text). $k_P(t)$ has the same behavior as $k_K(t)$, but phase shifted by 12 hours. (Bottom) With these time-varying inputs, the active enzyme concentrations $[K^*]$ and $[P^*]$ track the conversion rates $k_K(t)$ and $k_P(t)$, but $[S_p]$ shows an essentially constant amplitude from one cycle to the next. Thus, even with time-varying inputs, the activity of RpaA is sensitive primarily to the ratio $k_K(t)/k_P(t)$, which plays the role of the phosphoryation ratio $p(t)$, not to each rate individually, or by extention to the absolute concentrations of KaiC phosphoforms. In all calculations, $K_M = [S_{tot}]$.

2.J. PHASE DIAGRAMS FOR THE NTFO, TTC-($PPC_{Zwicker}$) AND TTC-(PPC_{Rust}) MODELS (FIG. 2.J.1)

In order to get a better understanding of how the cell cycle perturbs the clock, we have made phase diagrams for both the width of the 1:1 locking region and the average variance of the peak-to-peak times, as a function of the two key variables that affect locking, the number of gene copies N and the standard deviation in their replication times, σ_{rep}, for all models, *i.e.* the NTFO model, the TTC-($PPC_{Zwicker}$) model of Zwicker *et al.* [19]

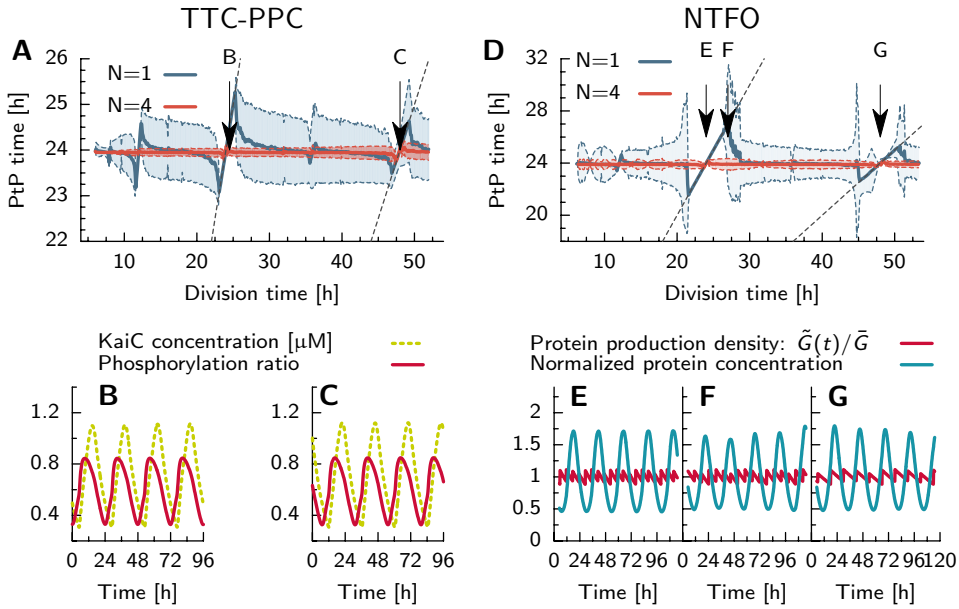

Figure 2.I.1: Multiple chromosome copies reduce the effect of periodic gene replications in both the TTC-($\text{PPC}_{\text{Zwicker}}$) model [19] (panels A–C) and in the simple negative transcriptional feedback oscillator (NTFO) (panels D–G). Panel (A) shows for the TTC-PPC model the average peak-to-peak time as a function of the cell-division time T_d for initial gene copy numbers $N = 1$ (blue) and $N = 4$ (red). This panel is identical to Fig. 2.5C of the main text, except for the arrows indicating the values of T_d for which times traces are shown in (B,C) for $N = 4$. Both the standard deviation in the peak-to-peak times (given by the shaded regions) and the regions where the oscillator is locked to the cell cycle are strongly reduced for $N = 4$. (B) The total KaiC concentration $C_{tot}(t)$ (dashed line) and phosphorylation fraction $p(t)$ (solid line) for the TTC-PPC model at $T_d = 24.5$ h, for $N = 4$. (C) $C_{tot}(t)$ and $p(t)$ at $T_d = 48$ h for the same model, again for $N = 4$. Note that, even at $T_d = 24.5$ h, immediately outside the locking region where the variance in T_{PtP} is generally largest, the effects of the cell cycle have almost completely disappeared. (D) The average peak-to-peak time and its standard deviation for the NTFO model, for $N = 1$ (blue line) and $N = 4$ (red line). It is seen that as in the Zwicker model, multiple chromosome copies dramatically reduce the strength of locking. (E–G) NTFO time traces of the protein-concentration oscillations $C(t)$ (blue lines) and the production density $\tilde{G}(t)$ (red lines), both normalized to their time average values, for $N = 4$, and for cell-division times indicated by the arrows in (D). With $N = 4$, only at, or very close to $T_d = 24$ hr (E), is the NTFO locked to the cell cycle. Even at $T_d = 27$ h (F) and $T_d = 48$ h (G), where for $N = 1$ the NTFO shows irregular behavior and large amplitude variations, respectively (see Fig. 2.2C,D), the time courses are much less perturbed by the cell cycle when $N = 4$. However, albeit greatly reduced, the effect of driving by the cell cycle can nonetheless still be observed, both in the persistence of small regions of locking and in the still appreciable variance in T_{PtP} when the oscillators are not locked (panel D). A PPC must be added to more fully attenuate the cell cycle's influence. (Compare panel D with panel A, taking into account the difference in scale of the y-axis; the standard deviation of T_{PtP} is about 3 times larger for the NTFO than for the full TTC-PPC model.)

and the TTC-(PPC$_{\text{Rust}}$) model [19, 22].

To make phase diagrams as a function N and σ_{rep}, we need to extend our model of stochasticity in the timing of gene replication introduced earlier in this document for $N = 1$, to also allow for a higher gene copy number, $N > 1$. Each cell cycle has N gene replication events which occur at times t_{g}^i within the interval $[0, T_{\text{d}})$. Each replication time is drawn from a Gaussian distribution with a standard deviation σ_{rep} that is proportional to T_{d}, and a mean $\bar{t}_{\text{g}}^i = T_{\text{d}}/N(\frac{1}{2} + i)$, where $i \in \{0, \ldots, N-1\}$ for each cell cycle. Replication times $t_{\text{g}}'^i$ that fall outside the interval $[0, T_{\text{d}})$, are mapped back onto it via $t_{\text{g}}^i = \text{Mod}(t_{\text{g}}'^i, T_{\text{d}})$. Note that this mapping slightly decreases the variance in the replication times; the real variance will therefore be slightly less than σ_{rep}^2 (at most 10%).

We run simulations with values for $\sigma_{\text{rep}}/T_{\text{d}}$ in the range $[0, 0.3]$, which, as mentioned above, corresponds to a scenario where the replication time is essentially chosen at random between 0 and T_{d}. We consider initial gene copy numbers of $N \in \{1, 2, 3, 4\}$.

Panels A, E, and I of Fig. 2.J.1 show the phase diagrams of the width of the 1:1 locking regions as a function of N and $\sigma_{\text{rep}}/T_{\text{d}}$, for the NTFO, TTC-(PPC$_{\text{Zwicker}}$) and TTC-(PPC$_{\text{Rust}}$) models, respectively. The panels underneath, (B, F, J), respectively, show cuts through these phase diagrams. Here, we consider an oscillator to be 1:1 locked to the cell cycle when the difference between its average PtP-time, $\langle T_{\text{PtP}} \rangle$, and the cell-division time T_{d} is less than 0.05 hour: $\langle T_{\text{PtP}}(T_{\text{d}}) \rangle - T_{\text{d}} < 0.05$.

Panels C, G, K of Fig. 2.J.1 show the phase diagrams of the average variance in the peak-to-peak times, $\langle \sigma_{\text{PtP}}^2 \rangle$, again as a function of N and $\sigma_{\text{rep}}/T_{\text{d}}$, for the NTFO, TTC-(PPC$_{\text{Zwicker}}$) and TTC-(PPC$_{\text{Rust}}$) models, respectively. Here, the variance is averaged over a range of cell-division times $6 < T_{\text{d}} < 52$:

$$\langle \sigma_{\text{PtP}}^2 \rangle = \frac{1}{N_{\text{sim}}} \sum_{i=1}^{N_{\text{sim}}} \sigma_{\text{PtP}}^2(T_{\text{d}}^i), \tag{2.20}$$

where N_{sim} is the number of evenly-spaced simulations performed in the range $6 < T_{\text{d}} < 52$ in steps of 0.1 hour, and $\sigma_{\text{PtP}}^2(T_{\text{d}}^i)$ is the variance in the PtP-times for 5000 hours of simulated time at cell-divistion time T_{d}^i. This quantity is a measure for the erratic behavior induced by the coupling to the cell cycle outside the locking region.

It seen that for all models both the width of the 1:1 locking region (Fig. 2.J.1, panels A/B, E/F, I/J) and the variance in the peak-to-peak times (Fig. 2.J.1, panels C/D, G/H, K/L), initially rapidly decreases with N, but then reaches a plateau. Even in the limit that $N \to \infty$, there is still some weak, residual driving by the cell cycle, because N rises linearly while the volume V rises exponentially during the cell cycle, leading to periodic variations in the gene density $G(t) = N(t)/V(t)$. It is also seen that the width of the 1:1 locking region decreases with increasing σ_{rep}, especially when N is small. However, while increasing the noise in the timing of replication reduces the width of the locking region, it also strongly increases the variance in the peak-to-peak times. A reliable clock requires not only a small locking window, but also a small peak-to-peak variance. Clearly, allowing for stochasticity in the timing of replication is not a solution to the locking problem.

Comparing the different models, it is seen that building the clock around a PPC reduces not only the width of the locking region, but also the variance in the peak-to-peak times, both in the TTC-(PPC$_{\text{Zwicker}}$) (Fig. 2.J.1E–H) and in the TTC-(PPC$_{\text{Rust}}$) model

(Fig. 2.J.1I–L). However, with $N = 1$ chromosome copy at the beginning of the cell cycle, the introduction of the PPC is not sufficient to fully eliminate locking. A TTC-PPC model still requires multiple chromosome copies that are replicated asynchronously.

2.K. METHODS

The delay-differential equations (DDE) describing our models in the absence of noise were propagated using the numerical differential equation solver of Mathematica 8 (Wolfram Research). For each value of T_d, we generated a single time trace of 2000 hours. In order to allow the oscillations to settle down to a steady state, we discarded the first 500 hours of each simulation and analyzed the remaining 1500 hours.

To find the peak-to-peak times T_{PtP} in the deterministic DDE simulations (including those with noise in the gene replication times), we used the built-in methods of Mathematica to return all local extrema in the concentration and phosphorylation fraction, respectively; these extrema correspond to the time points t_i where C_i or, respectively, p_i, is higher, in the case of a maximum, or lower, in the case of a minimum, than its two immediate neighbors. As is standard for numerical solution of differential equations, the spacing $t_i - t_{i-1}$ between successive time points was determined adaptively by the algorithm to meet imposed precision bounds but never exceeded 0.2 h. We then checked if a given local minimum was the lowest point within an interval of ± 18 hours centered on the minimum; if so, we defined this point as the global minimum of a single oscillation cycle. If there did exist a local extremum with a lower value, we repeated this procedure around the lower point until we found a point which was the lowest within a time interval of ± 18 hours. The same procedure was followed for the local maxima. The peak-to-peak time was then calculated by subtracting the times of two consecutive minima; we verified that substracting the times of two consecutive *maxima* gave essentially the same results.

To find the peak to peak times in the kinetic Monte Carlo simulation of the NTFO, we record the protein concentration $c_i(t_i)$ every 0.01 h of simulated time. We then use a sliding window of 18 hours over these concentrations to find the extrema of the oscillations. Specifically, to find the time t_j of the next local minimum: starting from the maximum of the preceding oscillation cycle at time t_i, we find the smallest concentration c_j, with $j > i$, in the window $t_j - t_i \leq 18$ h. (Given c_i was a local maxmium, there must exist a concentration $c_j < c_i$.) We then check whether there exists a concentration $c_k < c_j$ for $k > j$ and $t_k - t_j \leq 18$ h. If there is, we replace c_j by c_k and again search for a deeper minimum within 18 hours; otherwise, c_j is the minimum of the next oscillation cycle. A completely analogous procedure is used to identify the maxima of successive oscillation cycles.

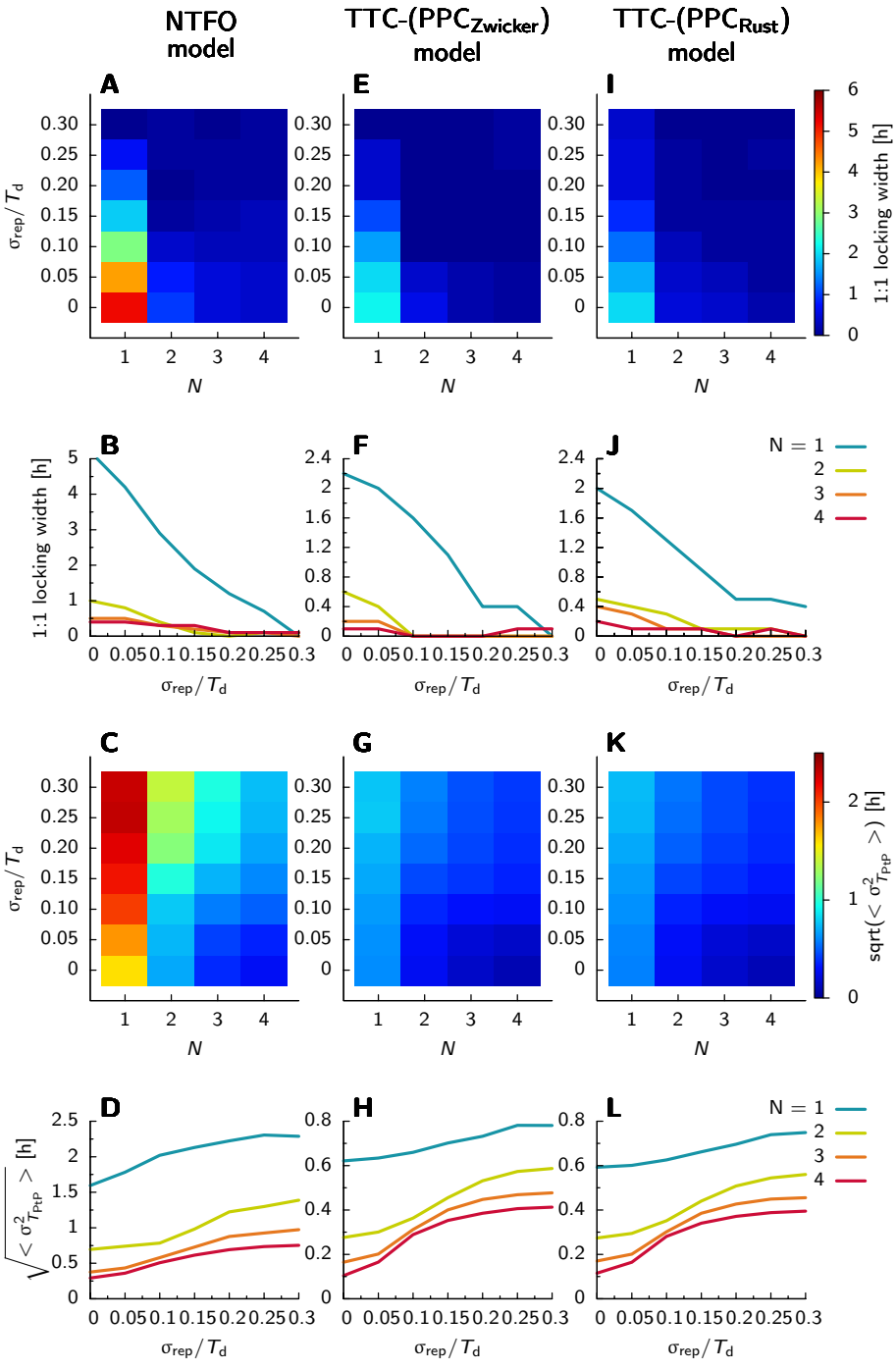
NTFO model
TTC-(PPC$_{Zwicker}$) model
TTC-(PPC$_{Rust}$) model
A
E
I
B
F
J
C
G
K
D
H
L
σ_{rep}/T_d
N
1:1 locking width [h]
N = 1
2
3
4
sqrt($<\sigma^2_{T_{PtP}}>$) [h]
$\sqrt{<\sigma^2_{T_{PtP}}>}$ [h]

Figure 2.J.1 ***(preceding page):*** Heat plots of the width of the 1:1 locking region (panels, A, E, I) and the average standard deviation of the peak-to-peak time (panels C, G, K), as a function of the initial gene copy number N and the standard deviation in the gene replication time $\sigma_{\mathrm{rep}}/T_{\mathrm{d}}$, for the NTFO model (panels A–D), the TTC-(PPC$_{\mathrm{Zwicker}}$) model [19] (panels E–H), and the TTC-(PPC$_{\mathrm{Rust}}$) model [19, 22] (panel I–L). Panels B,D,F,H,J,L show the same data as in the heat plots immediately above them, but as a function of $\sigma_{\mathrm{rep}}/T_{\mathrm{d}}$, for different values of N; *note the difference in scale of the y-axes* in these panels. The major results of panels B, D, F, and H are also summarized in Fig. 2.5C, D of the main text. The average standard deviation in the peak-to-peak time is the standard deviation in the peak-to-peak time averaged over $6 < T_{\mathrm{d}} < 52$; it is a measure for the erratic behavior of the clock outside the locking regions. It is seen that in all models the width of the locking region rapidly decreases with both N and $\sigma_{\mathrm{rep}}/T_{\mathrm{d}}$. However, the average standard deviation in the peak-to-peak time decreases with N, but increases with $\sigma_{\mathrm{rep}}/T_{\mathrm{d}}$. Clearly, while having multiple chromosome copies is a powerful strategy for preventing locking, increasing the stochasticity in the timing of gene replication is not—decreasing locking at the expense of much greater variation in the length of the periodic is unlikely to be functionally advantageous. Comparing the two models with a PPC to the NTFO model shows that adding a PPC to a TTC also decreases both the width of the locking region and the average standard deviation in the peak-to-peak time. Combining both features—a PPC and multiple chromosome copies—gives the strongest reduction in the coupling of the clock to the cell cycle.

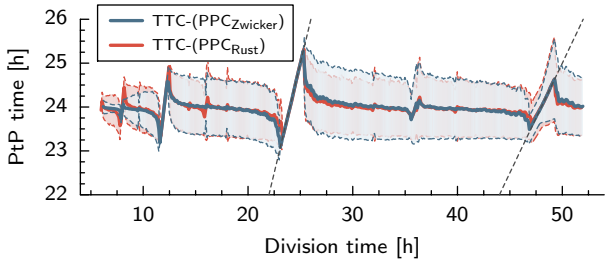

Figure 2.K.1: The TTC-(PPC$_{\mathrm{Rust}}$) model, which combines the TTC of Zwicker *et al.* with the PPC of Rust and coworkers [22], is susceptible to periodic gene replication. The figure shows the average and standard deviation of the peak-to-peak time T_{PtP} of the phosphorylation fraction $p(t)$ for the TTC-(PPC$_{\mathrm{Rust}}$) model and the TTC-(PPC$_{\mathrm{Zwicker}}$) model [19]. Clearly, the two models are similarly affected by the presence of the cell cycle.

3

ROBUSTNESS OF SYNTHETIC OSCILLATORS IN GROWING AND DIVIDING CELLS

Synthetic biology sets out to implement new functions in cells, and to develop a deeper understanding of biological design principles. In 2000, Elowitz and Leibler showed in a seminal paper that by rational design of the reaction network, and using existing biological components, they could create a network that showed periodic gene expression, dubbed the repressilator (Elowitz and Leibler, Nature, 2000). Much attention has gone to what sets the high levels of stochasticity observed in this oscillator, related to intrinsic noise in its components and stochastic expression of its genes. However, as all biological oscillators reside in growing and dividing cells, an important question is what the contribution of the cell cycle to the stochasticity in the synthetic oscillator is. In the previous chapter we showed that discrete replication events can couple the clock to the cell cycle. Here we expand this study to investigate the effects of the cell cycle on the period of two important synthetic oscillators in E. coli: The repressilator and the dual-feedback oscillator by Stricker et. al (Stricker et al.*, 2008). We find that the period of both oscillators can be strongly perturbed by the cell cycle, and that the effect critically depends on the position of the genes on the chromosome. Even in the limit of high levels of noise in the replication times of the genes, both oscillators show clear signatures of locking to the cell cycle. This work enhances our understanding of the design of robust biological oscillators inside growing and diving cells.*

The content of this chapter has been submitted to Physical Review E.

3.1. INTRODUCTION

Synthetic biology strives to implement new functions in living cells, and to develop a deeper understanding of biological design principles, using a modular rational design of biochemical reaction networks [28, 74, 75]. As synthetic biology becomes more mature, the goal is to design robust, stable and tunable networks [76–79], which are resilient to the effects of intrinsic noise and stochastic gene expression [2, 80–83]. Zooming in on oscillators, enhanced robustness has been achieved via the design of the reaction network at the single cell level [23, 29, 84–86], and by connecting multiple cells through quorum sensing [87–89]. However, the direct effect of the cell cycle on biological oscillators has been studied less [39, 40], though cell division introduces noise due to the binomial partitioning of the proteins [90, 91], and, perhaps more importantly, DNA replication can have a marked effect on protein production rates [4, 81].

In this work, we investigate a new aspect of the robustness of synthetic oscillators: We take the biochemical reaction network as given, and want to know what the effect the cell cycle has on the period of the oscillations. In the previous chapter we showed by mathematical modeling that discrete gene replication events can couple a circadian clock to the cell cycle, which means the period of the clock is dictated by the period of the cell cycle [21]. This raises the question of what the effect of periodic gene replications, which are generic to all growing and dividing cells and should thus pertain to any cellular oscillator, are on recently developed synthetic oscillators.

To this end, we study how strongly two canonical synthetic oscillators, both constructed in *E. coli*, can lock to the cell cycle: the repressilator, developed by Elowitz and Leibler [28], and the dual-feedback oscillator, developed by Stricker and coworkers [29]. A recent study, combining modeling with experiments, indepdently found that the dual-feedback oscillator can indeed couple to the cell cycle [92]. Here we study how the coupling strength depends on the noise in gene replication, and more importantly, on the positions of the genes on the DNA. We modify the original computational models of the repressilator and dual-feedback oscillator, to include the periodic doubling of the mRNA production rate with the cell cycle [21]. We consider the scenario that the synthetic oscillators are incorporated into the chromosome, although we will also discuss the fact that in the experiments the oscillators are implemented on plasmids present at high copy number [28, 29]. Under typical slow growth conditions, *E. coli* has one chromosome at the beginning of the cell cycle, in which case the gene copy number goes from 1 to 2 over the course of the cell cycle. At high growth rates, corresponding to cell division times shorter than the replication time of the DNA (on the order of 40 minutes), the chromosome can have multiple replication forks, which means that the gene copy number can be larger. Here, we only consider the regime that the cell division time is on the order of the DNA replication time or longer, such that the gene copy number rises from $N = 1$ at the beginning of the cell cycle to $2N = 2$ at the end. To quantify the sensitivity of the network to the cell cycle, we investigate the effect on the peak-to-peak time in the protein concentrations related to the oscillator, for different periods of the cell cycle.

Contrary to the Kai circadian clock, these two genetic oscillators consist of more than one gene. This introduces new important timescales to the problem: If the genes pertaining to the oscillator are placed at a distance on the chromosome, there is a time delay between when they are replicated. The synthetic oscillators studied here have an intrin-

sic period that is on the order of hours [28, 29], which is similar to the timescale of DNA replication, which takes at least 40 minutes. Consequently, the time delay can, depending on the reaction network, have a strong effect on the period of the oscillations.

We find that both synthetic oscillators can lock to the cell cycle for a wide range of cell division times, but that the effect critically depends on the positioning of the genes on the chromosome: Where the repressilator almost shows no locking when the genes are placed adjacently, the dual-feedback oscillator, to the contrary, experiences the strongest effect in this case, and locking decreases as the genes are placed further apart.

We study the effect of noise in the timing of gene replication on the strength of the coupling to the cell cycle. The noise in the replication time is the result of two stochastic processes: The timing of initiation of DNA replication and the progression of DNA replication. Stochasticity in the initiation of replication has the same effect on all the genes on the chromosome; a fluctuation in the initiation time propagates to the replication times of all the genes, leaving the interval between the gene replication times unchanged. In contrast, stochasticity in replication progression introduces temporal fluctuations in the time between the replication of different genes.

Our simulation results show that, for physiological levels for the noise in the gene replication times, the effect of gene replications on the period of the oscillations are strongly attenuated. However, clear signatures of the cell cycle are observable, especially around the 1:1 locking window. We then address the question which noise source has the strongest effect on attenuating the effects of the cell cycle: Initiation or progression of DNA replication. To find out, we study the effects of the cell cycle in two different scenarios: Either there is noise in the initiation of replication, such that the timing between replicating different genes is fixed, or the noise is limited to the progression of replication such that the initiation time is fixed and the timing between genes is stochastic. Our results reveal that noise in the initiation of DNA replication reduces the effect of locking much more than noise in DNA replication progression. This is because at biologically relavant noise levels, the standard deviation in the initiation of DNA replication is much larger than that in the progression of replication. Nevertheless, even with high levels of noise in the initiation of DNA replication, the effects of locking are still clearly present for cell division times around the oscillator's period. Our results thus predict that synthetic oscillators will be perturbed by the cell cycle in growing and dividing cells, when the oscillators are implemented on the chromosome.

Below, we first give an overview of the models for the repressilator, the dual-feedback oscillator and the models for the cell cycle. First we give a description of a completely deterministic cell cycle, and then introduce stochasticity in the model by making the time DNA replication is initiated and the time it takes to replicate the DNA stochastic variables. To determine how strong the oscillators are coupled to the cell cycle, we study how the period of the oscillators scales with the cell division time.

3.2. THEORY

To study the effect of the cell cycle on the period of synthetic oscillators, we will use the ODE models of the repressilator [28] and dual-feedback oscillator [29], as described in these papers. As we argue in more detail in [21], the key quantity connecting the cell

cycle and the biochemical network is the gene density, $G(t)$, the gene copy number over the cell volume. Because the protein production rate is proportional to the gene copy number, discrete gene replication events cause sudden doubling of the production rate [4]. We include the effects of the discrete gene replication events, by making the increase of the mRNA concentration due to transcription of each gene i, proportional to the gene density $G_i(t) = g_i(t)/V(t)$ [21]. Here g_i is the gene copy number of gene i which switches from 1 to 2 during the cell cycle, and $V(t)$ is the cell volume which exponentially doubles in size during cell division time T_d.

3.2.1. REPRESSILATOR

The repressilator consists of three genes, which sequentially repress each other's expression. As schematically shown in Fig. 3.1A, the first gene represses the expression of the second, which represses the third gene, which in turn represses the expression of the first again [28]. To take into account gene replication, the expression of mRNA is proportional to the gene density $G_i(t)$

$$\begin{aligned} \frac{dm_i(t)}{dt} &= -m_i(t) + \frac{G_i(t)}{\bar{G}_i}\frac{\alpha}{1+(p(t)_j)^n} + \alpha_0 \\ \frac{dp_i(t)}{dt} &= -\mu_p p_i(t) + \gamma m_i(t). \end{aligned} \tag{3.1}$$

Here, m_i and p_i are the concentrations of mRNA and proteins ($i \in \{1,2,3\}$), respectively, both rescaled with the constant of half-maximum repression K_M. The transcription rate is assumed to be proportional to the instantaneous gene density $G_i(t)$; importantly, the gene density can differ between the three genes when they are positioned differently on the chromosome, see Fig. 3.1, panels B and C. $\bar{G}_i$ is the time-averaged gene density, which depends on the phase of the cell cycle at which the gene is duplicated. The mRNA expression has a basal rate α_0 and an enhanced rate α, which is repressed by protein p_j, where $j \in \{3,1,2\}$, with a Hill coefficient n; here, following the original paper [28], time is rescaled in units of the mRNA lifetime and protein concentrations are in units of the concentration necessary for half-maximal repression. In the second equation, μ_p is the protein decay rate over the mRNA decay rate and γ is the translation efficiency, *i.e.* the average number of proteins produced per mRNA molecule. We used the parameters given in Table 3.1.

3.2.2. DUAL-FEEDBACK OSCILLATOR

The dual-feedback oscillator, schematically shown in Fig. 3.1D, consists of two genes, one encoding for an activator and one for a repressor [29]. The activator enhances the expression of both genes, while the repressor represses the expression of both genes. Since the genes have identical promoters, the temporal expression of the two proteins is similar. The model we employ is presented in the SI of [29], but to take into account the periodic variations in the gene density, we have modified the equations describing the

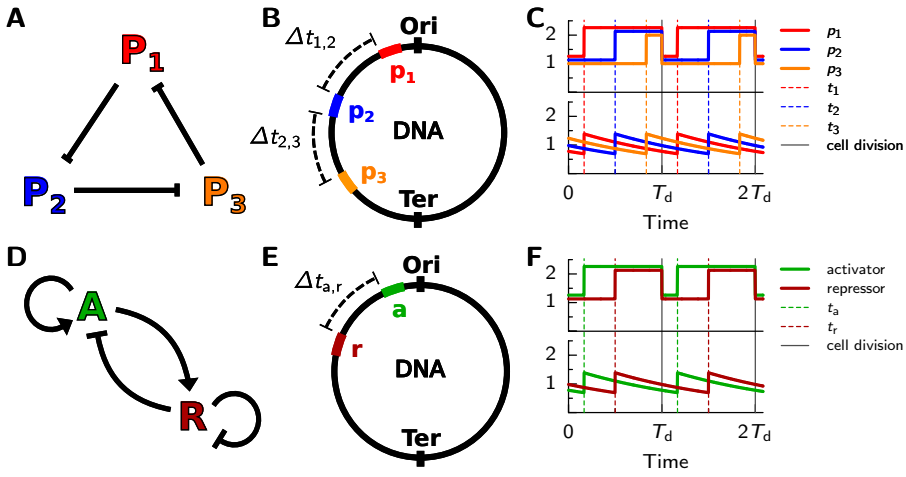

Figure 3.1: Models for the synthetic oscillators and cell cycle. (A) Network architecture of the Repressilator [28]: P_1 represses the production of P_2, P_2 represses P_3 and P_3 represses P_1 again. (B) Illustration of circular chromosome, with the origin (Ori) and termination (Ter) of replication. When the three genes p_1,p_2 and p_3 are placed at a distance on the chromosome, there are temporal delays $\Delta t_{1,2}$ and $\Delta t_{2,3}$, between when the genes are replicated. (C) Gene copy numbers (top) and gene densities (bottom) of the genes p_1 (red), p_2 (blue) and p_3 (orange), respectively. They are replicated at times t_1, t_2 and t_3, respectively, as indicated by the dashed lines. The black vertical lines indicate cell divisions. For the gene copy number, lines are shifted vertically for clarity. (Bottom) Gene densities for each gene, normalized by their average. (D) Network architecture of the dual-feedback oscillator [29]: The activator (A) auto-activates its production and enhances the production of the repressor (R). The repressor auto-represses its production, and suppresses the production of the activator. (E) Schematic of the circular chromosome. The genes for the activator (a) and repressor (r) are placed at different positions on the DNA, such that there is a temporal delay, $\Delta t_{\mathrm{a,r}}$, between there respective replication times. (F) Gene copy numbers (top) and gene densities (bottom) of the genes a (green) and r (red), respectively. Genes a and r are replicated at times, t_{a} and t_{r}, respectively, indicated by the dashed vertical lines. The black vertical lines indicate cell divisions. For the gene copy number, lines are shifted vertically for clarity.

transcription of mRNA of the activator and repressor

$$P^{a/r}_{0,0} \xrightarrow{b_{a/r}G_{a/r}(t)/\bar{G_{a/r}}} P^{a/r}_{0,0} + m_{a/r} \qquad (3.2)$$
$$P^{a/r}_{1,0} \xrightarrow{\alpha b_{a/r}G_{a/r}(t)/\bar{G_{a/r}}} P^{a/r}_{1,0} + m_{a/r}.$$

Here $P^{a/r}_{m,n}$ denotes the promoter of the (a)ctivator/(r)epressor gene, with $m = 0,1$ activator protein and $n = 0$ repressor protein bound to it, respectively. The mRNA $m_{a/r}$ of the activator (a) and repressor (r) is transcribed with a rate $(\alpha)b_{a/r}G(t)$, which depends on the state of the promoter and on the gene density $G_{a/r}(t)$. Parameters are given in Table 3.1. Genes can be placed at a distance from each other on the chromosome, as shown in Fig. 3.1E and F, which introduces a delay between when they are replicated. The intrinsic period of this oscillator without the driving by the gene density is ~ 40 min-

utes. To study the importance of the timing of gene replication, we want an intrinsic period that is longer than the replication time of the DNA, which is also around 40 minutes. To obtain a longer clock period, we use the experimental observation in [29] that the clock period scales with temperature via the Arrhenius law. To this end, we scale all rate constants, k_i, in the model, using

$$k_i = k_{\text{ref}} \exp(-\Theta_{\text{cc}}[1/T - 1/T_{\text{ref}}]), \tag{3.3}$$

where k_{ref} is the rate constant at the reference temperature T_{ref} of 310K and $\Theta_{\text{cc}} \approx 8300$K is a constant. We will evaluate the model at a temperature of 303K where the clock has an intrinsic period of about 73 minutes.

3.2.3. DIFFERENT MODELS FOR GENERATING GENE REPLICATION TIMES

The time at which a gene is replicated depends on the timing of three major events during the cell cycle: The time between the start of the cell cycle and initiation of DNA replication, the replication time of the chromosome and, after this has finished, the time until cell division. As we argued in [21], cell division has a smaller effect on the oscillator as compared to gene replication, as both the cell volume and the gene copy number divide by two at cell division, leaving the important gene density unchanged. Therefore, in our model we assume there is no stochasticity in the division time, and keep T_{d} fixed. Furthermore, we assume that the *E. coli* cells grow slowly, such that the division time is always longer than the DNA replication time. In this case, there are at most two origins of replication per cell, and we do not have to take into account the effects of multiple replication forks [93].

Because it is still poorly understood how the cell coordinates the replication and division cycles, in this work we employ a simple model for the cell cycle. Evidence emerges that initiation of chromosome replication is triggered at a fixed density of the origin of replication (Ori), $\hat{G}_{\text{Ori}}$, independent of cell's division time [94, 95]. Given that the density of the Ori depends on the cell volume $V(t)$, $G_{\text{Ori}} = 1/V(t)$, the time and precision of initiation of DNA replication is set by the evolution of the cell volume and the precision of the sensor for $G_{\text{Ori}}(t)$. Because we consider the slow growth regime where at the beginning of the cell cycle there is only one origin of replication, and because we assume that the initial volume is independent of the growth rate, it follows that the average time at which DNA replication is initiated is at a fixed fraction α_{init} of the division time T_{d}, $\Delta t_{\text{init}} = \alpha_{\text{init}} T_{\text{d}}$, with a standard deviation σ_{init}. We choose, based on data presented in [95], $\alpha_{\text{init}} = 0.2$. The time it takes to replicate the chromosome depends on the speed of the DNA polymerase, which in turn can depend on the cell's physiological state [96]. For simplicity, we assume that the mean time to replicate the whole chromosome is $T_{\text{rep}} = 40$ minutes with a standard deviation given by σ_{rep}, both independent of the cell's division time. In this work we consider two models for the timing of gene replications: One where both the initiation and the progression of DNA replication are deterministic, such that gene replications occur at the same phase each cell cycle and one where we introduce noise in these two processes. The effects of noise in the initiation and progression of DNA replication on the gene replication times is illustrated in Fig. 3.2.

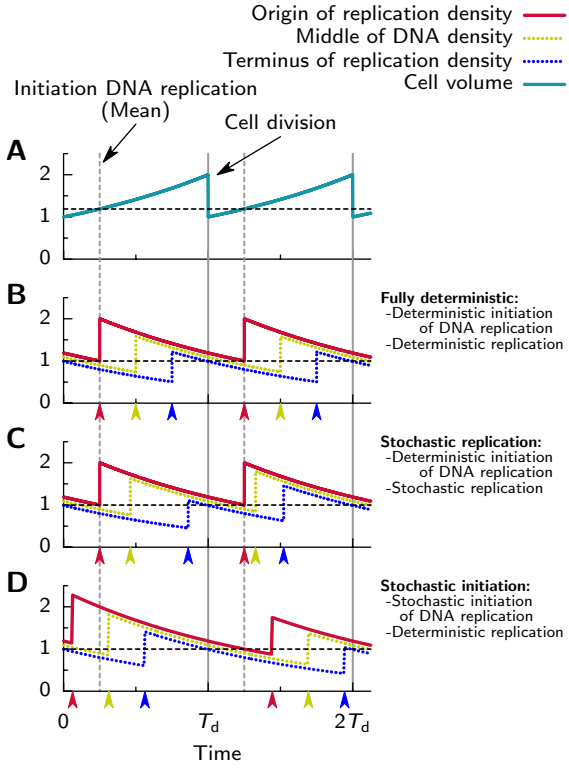

Figure 3.2: Models to determine the gene replication times. (A) Time trace of cell volume, which is a deterministic function of time in all models, where cell division occurs with a period T_d, indicated by the vertical solid gray lines. The vertical dashed gray lines indicate the times at which DNA replication is initiated when the timing of the initiation of replication is deterministic. The horizontal dashed lines show the volume (A) or the concentration of the origin of replication (B-D), at which DNA replication, on average, initiates. (B-D) Time traces of the density of the origin of replication of the chromosome (red solid line), a gene precisely in between the origin and terminus of replication (green dotted line) and the terminus of replication (blue dotted line). Arrows below the x-axis indicate the replication times of these sites. Note that the gene densities show no discontinuity at cell division. All gene densities are normalized by the critical density for replication initiation. (B) Fully deterministic model. Initiation of replication and the replication of the two genes occur at fixed times each cell cycle. (C) When there is stochasticity in DNA replication progression, the timing between initiation of replication and the replication of genes further along the DNA becomes stochastic. (D) When the initiation of replication is stochastic, but the replication rate is constant, all replication events move in concert, and the time between initiation and replication of the genes is fixed.

DETERMINISTIC MODEL

The first model is completely deterministic. Indeed, when we assume the evolution of the cell volume, $V(t)$, to be deterministic and that DNA replication initiates exactly when $G_{\mathrm{Ori}}(t) = \hat{G}_{\mathrm{Ori}}$, then the evolution of $G_{\mathrm{Ori}}(t)$ becomes fully deterministic. Clearly, since both the initiation and the progression of DNA replication are deterministic, the respective genes are copied at the same times each cell cycle (See Fig. 3.2A and B). Furthermore, in our model the first gene of the oscillator is next to the origin of replication, such that the time this gene is replicated, $t_1 = \Delta t_{\mathrm{init}} = \alpha_{\mathrm{init}} T_{\mathrm{d}}$. Note that it is not important when exactly during the cell cycle the gene is replicated, as it only changes the gene density by a prefactor, which we compensate for by normalizing $G_i(t)$ by its mean $\bar{G}_i$. However, as we will see, the time between the replication of the different genes is important. Genes can be placed apart on the DNA which introduces a time delay, $\Delta t_{i,j}$, between when the genes i and j are copied, respectively. The times during the cell cycle when the genes p_1, p_2 and p_3 are replicated, for the repressilator, and the activator and repressor genes for the dual-feedback oscillator, are

$$\begin{aligned} t_1 &= \Delta t_{\mathrm{init}} & t_{\mathrm{a}} &= \Delta t_{\mathrm{init}} \\ t_2 &= t_1 + \Delta t_{1,2} & t_{\mathrm{r}} &= t_{\mathrm{a}} + \Delta t_{\mathrm{a,r}} \\ t_3 &= t_2 + \Delta t_{2,3} \end{aligned} \tag{3.4}$$

STOCHASTIC MODEL: NOISE IN THE INITIATION AND PROGRESSION OF DNA REPLICATION

For the second model, we again assume that the evolution of the cell volume is deterministic, but turn replication progression and replication initiation into stochastic processes. Due to stochasticity in the progression of DNA replication, the time interval between the gene replication events becomes stochastic, as illustrated in Fig. 3.2C. We assume the time it takes to replicate the full chromosome follows a Gaussian distribution with a mean T_{rep}=40 min, and standard deviation σ_{rep} that is proportional to replication time T_{rep}. When the standard deviation in the DNA replication time is the result of many independent stochastic steps, the time between replicating genes i and j, $\delta\tau_{i,j}$, which on average takes a time $\Delta t_{i,j}$, will therefore also be Gaussian distributed with a standard deviation of $\sqrt{\Delta t_{i,j}/T_{\mathrm{d}}}\,\sigma_{\mathrm{rep}}$.

Stochasticity in the initiation of replication affects the replication times of all genes equally; indeed, the time between copying two different genes, $\Delta t_{i,j}$, is constant, as is shown in Fig. 3.2D. This stochasticity in the timing of the initiation can come from the sensing limit of measuring $G^{\mathrm{Ori}}(t)$, or because of stochasticity in the evolution of the cell volume (which, however, we assume to progress deterministically in this scenario). In our model, the time of initiation of DNA replication, $\delta\tau_{\mathrm{init}}$, is a stochastic variable drawn from a Gaussian probability distribution with a mean $\alpha_{\mathrm{init}} T_{\mathrm{d}}$ with a standard deviation σ_{init}. Assuming the standard deviation in measuring $G_{\mathrm{Ori}}(t)$, $\sigma_{G_{\mathrm{Ori}}}$, is small, the standard deviation in the initiation time is

$$\sigma_{\mathrm{init}} = \left| \frac{d(\Delta t_{\mathrm{init}})}{dG_{\mathrm{Ori}}} \bigg|_{G_{\mathrm{Ori}}=\hat{G}_{\mathrm{Ori}}} \right| \sigma_{G_{\mathrm{Ori}}}. \tag{3.5}$$

DNA replication is initiated when $G_{\mathrm{Ori}} = V_0^{-1}\exp(-\ln(2)/T_{\mathrm{d}}\,\Delta t_{\mathrm{init}}) = \hat{G}_{\mathrm{Ori}}$, where V_0 is the cell volume after cell division. Solving this equation for the initiation time gives

$\Delta t_{\text{init}} = -T_{\text{d}} \ln(\hat{G}_{\text{Ori}} V_0)/\ln(2)$. Then, from Eq. 3.5 it follows that the standard deviation in the initiation time is $\sigma_{\text{init}} \sim T_{\text{d}} \sigma_{\text{G}_{\text{Ori}}}$. Therefore, in our model, the standard deviation in the initiation time is proportional to T_{d}.

Assuming that the two stochastic processes are independent, the replication times of the genes for the repressilator and dual-feedback oscillator become, respectively

$$\begin{aligned} t_1 &= \delta\tau_{\text{init}} & t_{\text{a}} &= \delta\tau_{\text{init}} \\ t_2 &= t_1 + \delta\tau_{1,2} & t_{\text{r}} &= t_{\text{a}} + \delta\tau_{\text{a,r}} \\ t_3 &= t_2 + \delta\tau_{2,3} \end{aligned} \tag{3.6}$$

Because in our model, the division time is fixed each cell cycle, we have to constrain the values of the replication times to lie within the finite interval $[0, T_{\text{d}}]$. First we choose $\delta\tau_{\text{init}}$, and constrain it to lie within $[0, (T_{\text{d}} - T_{\text{rep}})]$. Then we draw a value for $\delta\tau_{1,2}$ and constrain it to lie within the interval that is symmetric around its mean value $\Delta\tau_{1,2}$, $[0, 2\Delta t_{1,2}]$. Similarly, we draw a value for $\delta\tau_{2,3}$ constrained to the interval $[0, 2\Delta t_{2,3}]$. For the dual-feedback oscillator, the times $\delta\tau_{a,r}$ are constrained to the interval $[0, 2\Delta t_{\text{a,r}}]$. We map values that lie outside these intervals back on it by mirroring these values across the nearest boundary of the domain.

Recent single cell experiments revealed the coefficient of variation (CV) in the time of initiation of DNA replication, $\text{CV}_{\text{init}} = 0.7$, and in the time of replicating the DNA, $\text{CV}_{\text{rep}} = 0.16$, in slow growing *E. coli* cells [96]. Given our models for stochasticity in replication times (including the fact that the initiation times are constrained to lie in the windows discussed above), we find that standard deviations of $\sigma_{\text{init}} = 0.2 T_{\text{d}}$ and $\sigma_{\text{rep}} = 0.35 T_{\text{rep}}$ give similar coefficients of variation. All parameters are listed in Table 3.1.

3.3. RESULTS

Here we study how the peak-to-peak times of the oscillations of the repressilator and the dual-feedback oscillator depend on the cell division time. Furthermore, we illuminate the effects of the position of the genes on the DNA and the role of stochasticity in the replication times.

3.3.1. REPRESSILATOR

We first consider the scenario in which the three genes are close together on the chromosome, such that, to a good approximation, they are replicated at the same time, and the timing of DNA replication is fixed. In Fig. 3.1, panel A, we show the mean peak-to-peak time, T_{PtP}, in the concentration of P_1, for different cell division times, T_{d}. Clearly, locking is not very strong: The locking windows—the range of cell division times where the mean peak-to-peak time of the repressilator is equal to a multiple of T_{d}—are very small. The only effect of locking is that in these very small windows the variance in the peak-to-peak time is strongly reduced. The reason why locking is weak is that while the genes are replicated at the same time, they are expressed at different times. This means that gene replication has a different effect on the expression level of each of the three genes. Hence, even when the cell cycle period T_{d} is approximately equal to the the oscillator's intrinsic period T_{int}, $T_{\text{d}} \approx T_{\text{int}}$, the oscillation of each protein concentration has a

different amplitude, as shown in Fig. 3.1B. This makes it harder for all three protein oscillations to get the same period as that of the cell cycle, and become locked to it. Interestingly, Fig. 3.1C shows that when the cell-cycle time is twice the intrinsic clock period, the pattern of alternating smaller and larger oscillation amplitudes can still be observed for each of the respective protein concentration profiles. This observation can be used to detect the effect of periodic gene replication experimentally.

We now consider a scenario in which the different genes are replicated at different times during the cell cycle, which corresponds to a situation where the genes are located at different positions on the chromosome. We assume that the gene for protein p_1 is close to the origin of replication, such that it is copied at the moment DNA replication is initiated. We consider two scenarios for the order of the genes on the DNA. In the first scenario, panel D, genes are placed on the DNA in order of their interaction in the biochemical reaction network, p_1, p_2, p_3 (see Fig. 3.1A): The gene for p_2 is copied a time $\Delta t_{1,2}$ after p_1, and p_3 a time $\Delta t_{2,3}$ after p_2. In the second scenario, panel E, genes are in order of maximal expression: p_3, p_2, p_1 (see Fig. 3.1,B and C), which corresponds to negative values of $\Delta t_{1,2}$ and $\Delta t_{2,3}$. Throughout this work, we will use the condition $\Delta t_{1,2} = \Delta t_{2,3}$. Interestingly, while the locking regions are very small when the genes are replicated at the same time ($d_2 = 0$, panel A, gray lines in panels D and E), replicating them at different times introduces marked locking: both for $\Delta t_{1,2} > 0$ (panel D) and $\Delta t_{1,2} < 0$ (panel E) strong locking is observed. Even more strikingly, the 1:1 locking region is largest when genes are replicated in order of maximal expression, and when the distance between them is the largest (panel E). This can be understood by noting that when genes are replicated in the order of maximal expression, shifting the phase of the clock with respect to that of the cell cycle has then the strongest effect on the amplitude and hence the period of the clock oscillations, which underlies the phenomenon of locking, as explained in chapter 2.

To see if locking persists in the presence of physiologically levels of noise in gene replication times, we change the gene replication times t_1, t_2 and t_3 into stochastic variables via Eq. 3.6. Our results reveal that both when $\Delta t_{1,2} > 0$ (Fig. 3.2, panel A) and when $\Delta t_{1,2} < 0$ (panel B), the coupling of the repressilator to the cell cycle is strongly attenuated. However, the effects of the cell cycle are still clearly observable around the 1:1 locking region and when $T_\mathrm{d} = 0.5 T_\mathrm{int}$. For division times longer than the oscillators intrinsic period, all signatures of locking have disappeared.

3.3.2. DUAL-FEEDBACK OSCILLATOR

Fig. 3.3A shows strong locking of the dual-feedback oscillator to the cell cycle. We assume here that the genes are located next to each other on the chromosome, so that their time-varying gene-densities are the same. Clearly, the widths of the locking regions are very large; they are even larger than those observed for our simple negative feedback oscillator studied in chapter 2. In Fig. 3.3B we show a time trace of the irregular oscillations around a cell-division time of $T_\mathrm{d} = 98$ minutes. Fig. 3.3C shows that the amplitude of the oscillations alternates between a high and a low value when the cell-division time T_d is about twice the intrinsic clock period of $T_\mathrm{int} = 74$ minutes, due to periodic gene replication every other clock period. We thus conclude that also the dual-feedback oscillator can strongly lock to the cell cycle and that this effect should be observable experimen-

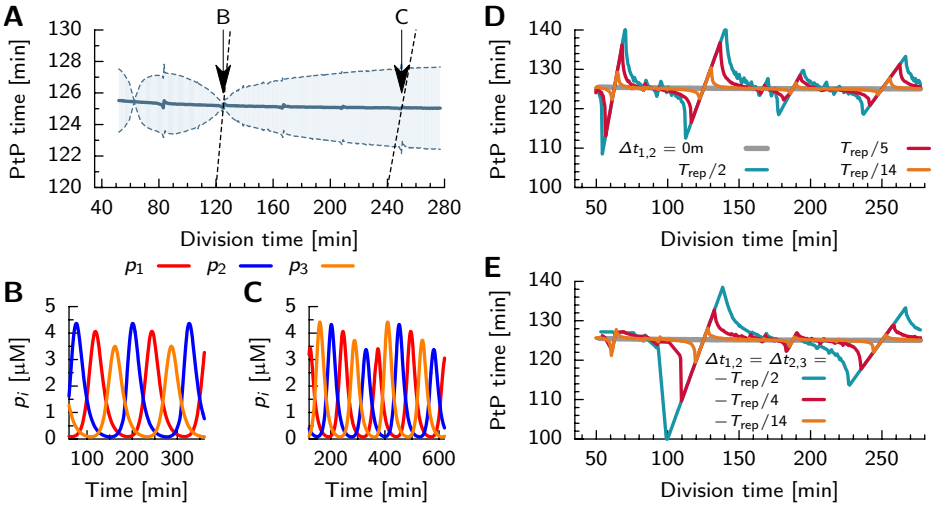

Figure 3.1: The repressilator [28] can strongly lock to the cell cycle, and the strength of locking depends sensitively on how the genes are positioned on the DNA. (A) Average (solid line) and standard deviation (shaded region) of the peak-to-peak time T_{PtP} as a function of the division time. The repressilator has an intrinsic period of $T_{\mathrm{int}} = 125$ minutes and the three genes are replicated simultaneously. The locking regions around T_{int} and $2T_{\mathrm{int}}$ are almost absent. (B and C) Representative time traces of the concentrations of the three repressilator proteins, $p_1(t)$ (red), $p_2(t)$ (blue) and $p_3(t)$ (orange), for the cell-division times indicated by the arrows in panel A. (B) When $T_{\mathrm{d}} = T_{\mathrm{int}}$, the oscillations are very regular (almost no variance in the PtP-times), but each protein concentration has a different amplitude. (C) At $T_{\mathrm{d}} = 2T_{\mathrm{int}}$, all three protein concentrations switch between a small and a large amplitude in successive oscillation cycles. Panels D and E show the effect of varying the timing of replication of the three genes. For clarity, we only show the average peak-to-peak time as a function of T_{d}, not the standard deviation. We assume that the first gene of the repressilator is replicated at the time DNA replication starts, and that the mean time intervals between copying the genes are $\Delta t_{1,2}$ and $\Delta t_{2,3} = \Delta t_{1,2}$. The gray line gives the situation where all genes are replicated simultaneously, $\Delta t_{1,2} = 0$. Other values of $\Delta t_{1,2}$ are given in the legend, and are written as a fraction of the mean DNA replication time T_{rep}. Panel D shows scenarios for which the genes are on the DNA in order of their interaction, p_1, p_2, p_3, and panel E shows situations where the genes are positioned in order of their maximal expression during the oscillator's period, p_3, p_2, p_1. Remarkably, for all $\Delta t_{1,2} \neq 0$, there is significant locking. For the scenarios in panel E, the width of the 1:1 locking region is the largest; this is because the temporal order the genes' replication during the cell cycle is the same as that of their expression in the oscillator. Clearly, the timing of gene replication can markedly affect locking, which means that the spatial distribution of the genes over the chromosome can be of critical importance in the interaction between the clock and the cell cycle.

tally.

Fig. 3.3D,E shows the result of varying the moment of gene replication for the two genes. Again, in this model, the first gene of the oscillator is placed next to the origin of replication such that it is replicated at initiation of DNA replication, and the second gene is replicated with a mean delay $\Delta t_{\mathrm{a,r}}$ later. For positive $\Delta t_{\mathrm{a,r}}$, the activator is repli-

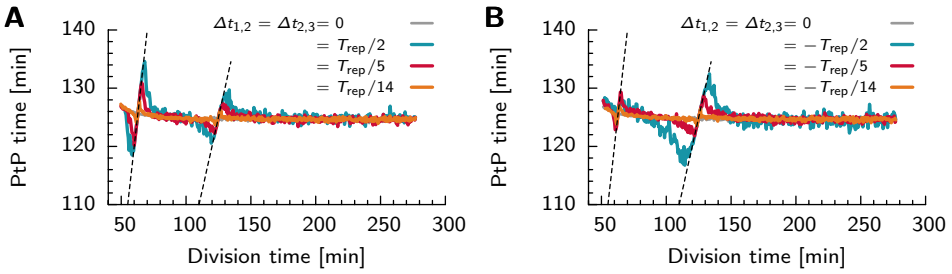

Figure 3.2: In the repressilator, locking persists in the presence of physiological levels of noise in the gene replication times. In all panels, solid lines show the peak-to-peak time in the activator concentration for different periods of the cell division time T_d. Standard deviation in T_{PtP} omitted for clarity, but is similar in all panels. Legends are defined in Fig. 3.1. We used the physiologically motivated values for the standard deviations in the timing of the initiation, $\sigma_{init} = 0.2T_d$, and the progression, $\sigma_{rep} = 0.35T_{rep}$, of DNA replication. The two panels show a different order of the genes p_1, p_2 and p_3 with $\Delta t_{1,2} > 0$ (A), and $\Delta t_{1,2} < 0$ (B). Clearly, at these noise levels, locking is strongly reduced compared to the deterministic case (See Fig. 3.1, panels D and E), but still clearly observable around $T_d = T_{int}$.

cated before the repressor, and negative $\Delta t_{a,r}$, vice versa. We vary the time delay between the replication of the two genes, as $\Delta t_{a,r} = 0$, $T_{rep}/8$, $T_{rep}/2$ and T_{rep} (panel D) and minus these values (panel E), where T_{rep} is the mean replication time of the DNA. It is seen that in both scenarios the strength of locking decreases with increasing the distance between the genes on the DNA: The strongest entrainment is observed when the genes are replicated at the same time during the cell cycle (gray lines), in stark contrast to the behavior of the repressilator. While in the repressilator the locking increases with the distance between the genes, the dual-feedback oscillator shows the opposite behavior. Interestingly, though, in the dual-feedback oscillator locking still persists when the genes are placed at maximum distance from each other.

To see if locking persists in the presence of noise in the timing of gene replications, we changed the time of replication of both genes, t_a and t_r, into stochastic variables via Eq. 3.6. The noise strongly attenuates the effects of the cell cycle, both for positive (Fig. 3.4 panel A) and negative (panel B) $\Delta t_{a,r}$, as compare against deterministic result of Fig. 3.3. However, the peak-to-peak times of the dual-feedback oscillator are still perturbed around the 1:1 locking region, especially in the case $\Delta t_{a,r} > 0$.

3.3.3. WHAT ATTENUATES THE EFFECTS OF THE CELL CYCLE MORE: STOCHASTICITY IN THE INITIATION OR PROGRESSION OF DNA REPLICATION?

Comparing Fig. 3.1 with Fig. 3.2 for the repressilator and Fig. 3.3 with Fig. 3.4 for the dual-feedback oscillator, it is clear that noise in gene replication times has a significant effect on the coupling between the cell cycle and these synthetic oscillators. In our model, noise in the replication times is the result of noise in the initiation and in the progression of DNA replication. We want to know which of these two sources of stochasticity is key

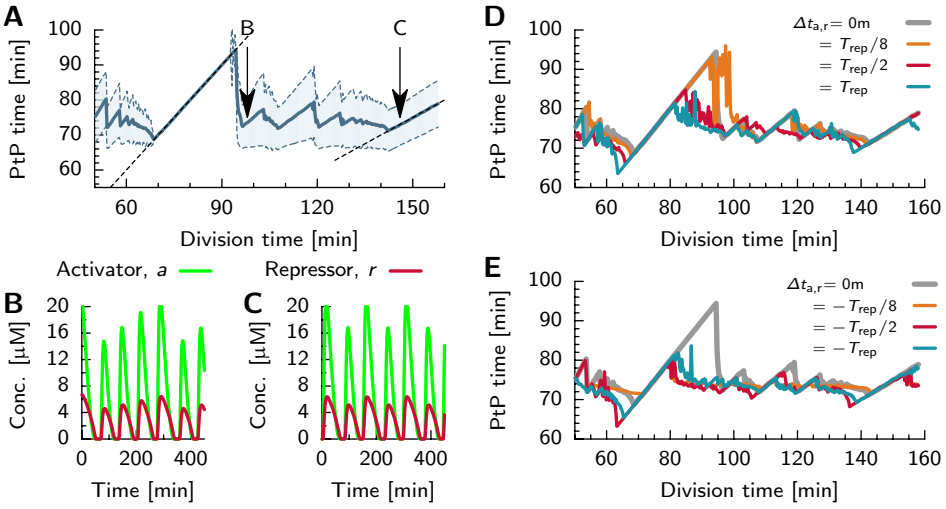

Figure 3.3: The dual-feedback oscillator [29] can strongly lock to the cell cycle, and the strength of locking depends on the temporal order in which the genes are replicated during the cell cycle. The intrinsic period of the oscillator $T_{\text{int}} = 73$ minutes. (A) Average (solid line) and standard deviation (shaded region) of the peak-to-peak time T_{PtP} as a function of the division time T_{d} when both genes are replicated simultaneously. There is a wide region of cell division times (around $T_{\text{d}} = T_{\text{int}}$) where the oscillator has a T_{PtP} equal to the the cell cycle (left dashed line). (B and C) Representative time traces for the division times indicated by the arrows in panel A. Shown are the activator and repressor concentrations $a(t)$ (green line) and $r(t)$ (green line), respectively. At a cell-division time of $T_{\text{d}} = 98$ min (B), just outside the region where the oscillator is locked to the cell cycle, the time traces show very irregular behavior resulting in a large variance in the PtP times. At $T_{\text{d}} = 2T_{\text{int}}$ (C), the oscillations switch between a small and a large amplitude in successive oscillation cycles, a signature of the periodic gene replications. (D and E) The effect of the order of gene replication during the cell cycle. For clarity, only the average peak-to-peak time as a function of T_{d} is shown, not the standard deviation. We assume one of the two genes is always replicated at the start of DNA replication, and that the other gene is replicated with a delay $\Delta t_{\text{a,r}}$. In both figures, the gray line gives the situation where the genes are replicated simultaneously, $\Delta t_{\text{a,r}} = 0$. Other values of $\Delta t_{\text{a,r}}$ are given in the legend, and are written as as a fraction of the DNA replication time T_{rep}. (D) Positive $\Delta t_{\text{a,r}}$; the repressor gene is replicated after the activator gene. (E) Negative $\Delta t_{\text{a,r}}$; the repressor gene is replicated before the activator gene. Both in panels D and E, the 1:1 locking region decreases as $\Delta t_{\text{a,r}}$ increases. Comparing the behavior of the dual-feedback oscillator, which exhibits the strongest entrainment when the genes are replicated simultaneously, with that of the repressilator, which shows the weakest coupling when the genes are replicated together, illustrates that the influence of the cell cycle on the clock depends in a non-trivial way on the architecture of the clock and on the nature of the driving signal.

for reducing the coupling between the cell cycle and the oscillator.

To find out whether the initiation or the progression of DNA replication is more important for attenuating the effects of gene replications, we studied two models for the noise in the replication times. Either there is only noise in the progression of replication, such that the time intervals between replicating different genes, $\delta\tau_{1,2}$ and $\delta\tau_{\text{a,r}}$,

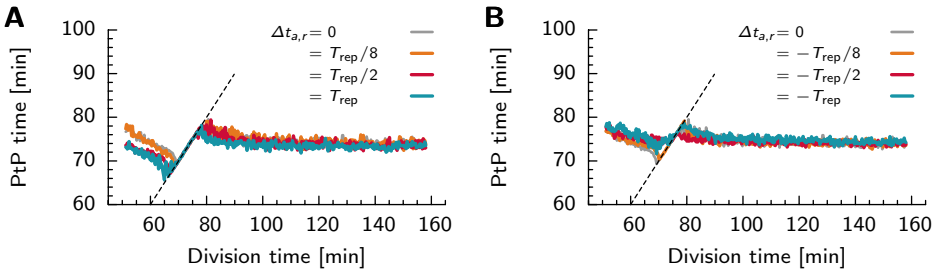

Figure 3.4: In the dual-feedback oscillator, locking persists in the presence of physiological levels of noise in the gene replication times. In all panels, solid lines show the peak-to-peak time in the activator concentration for different periods of the cell division time T_d. Standard deviation in T_{PtP} omitted for clarity, but is similar in all panels. Legends are defined in Fig. 3.3. We used the physiologically motivated values for the standard deviations in the timing of the initiation, $\sigma_{init} = 0.2T_d$, and the progression, $\sigma_{rep} = 0.35T_{rep}$, of DNA replication. The two panels show a different order of the activator and repressor gene with $\Delta t_{a,r} > 0$ (A), and $\Delta t_{a,r} < 0$ (B). As observed for the repressilator, locking is strongly reduced compared to the deterministic case (See Fig. 3.3, panels D and E), but still clearly observable around $T_d = T_{int}$.

are stochastic variables but the time of initiation of DNA replication is deterministic, $\Delta t_{init} = \alpha_{init} T_d$ (See Fig. 3.2C). Or, the initiation of DNA replication, $\delta\tau_{inti}$, is stochastic but the progression of replication is deterministic such that $\Delta t_{1,2}$ and $\Delta t_{a,r}$ are fixed each cell cycle (See Fig. 3.2D). We will use the same values for the standard deviations σ_{init} and σ_{rep} of the two noise sources as before.

In Fig. 3.5 we show the effects of the cell cycle on the period of the repressilator when there is only noise in the progression of replication, panels A and B, or when there is only noise in the initiation of DNA replication, panels C and D. Clearly, when there is only noise from replication progression, both for positive (panel A) and negative (panel B) $\delta\tau_{1,2}$, the width of the locking regions are almost the same as compared to the deterministic case (See Fig. 3.1, panels D and E). The effects of the cell cycle are not significantly attenuated by the noise in DNA replication progression. However, when the noise is due to the initiation of replication (panels C and D for positive and negative $\delta\tau_{1,2}$, respectively), all signatures of coupling disappear for $T_d > T_{int}$, and the width of the 1:1 locking region is strongly reduced compared to the case of a deterministic cell cycle. We conclude that the decrease in locking to the cell cycle is predominantly due to the stochasticity in the initiation time of DNA replication.

For the dual-feedback oscillator we obtain similar results. In Fig. 3.6 we show the effects of the cell cycle on the period of the dual-feedback oscillator when there is only noise in the progression of replication, panels A and B, or when there is only noise in the initiation of DNA replication, panels C and D. When there is only noise due to the progression of DNA replication, both for positive (panel A) and negative (panel B) $\delta\tau_{a,r}$, strong signatures of locking persists, especially around $T_d = T_{int}$ and $T_d = 2T_{int}$. Again, stochasticity in DNA replication progression does not attenuate the coupling to the cell cycle much. When the source of noise is due to stochasticity in the initiation of DNA replication, almost all effects of the cell cycle on the peak-to-peak time of the dual-

feedback oscillator have disappeared; only when $T_d = T_{int}$ locking can still be observed. Clearly, also for the dual-feedback oscillator the initiation of DNA replication has the biggest effect on the coupling between the cell cycle and the oscillator.

We observe that, both for the repressilator and the dual-feedback oscillator, the initiation of DNA replication is dominant in attenuating the effects of the cell cycle. Why is this the case? An oscillator couples to the cell cycle by maintaining a specific phase relation between the phase of the oscillator and that of the gene density, as explained in Fig. 2.3. When the standard deviation in the replication times is of the same order as the intrinsic period of the oscillator, it becomes impossible to maintain this phase relation, and the oscillator can not couple to the cell cycle. Because in our model, the standard deviation in the initiation of replication is proportional to T_d, while the standard deviation in replication *progression* is constant, initiation of DNA replication will be the dominant source of noise when $T_d > T_{int}$. Indeed, for $T_d > T_{int}$, the stochasticity in the initiation of DNA replication will be so large, that the clock no longer couples to the cell cycle (See Fig. 3.2 and Fig. 3.4). For $T_d \leq T_{int}$, the stochasticity in the initiation of DNA replication is much smaller. Moreover, the noise in DNA replication progression is so small that the coupling of the clock to the cell cycle is not much weakened by it (See Fig. 3.5A,B and Fig. 3.6A,B). This explains why for $T_d \leq T_{int}$, noise in DNA replication does not appreciably attenuate the locking of the clock to the cell cycle.

3.4. DISCUSSION

Discrete gene replication events, present in all cells, can have marked effects on the period of circadian clocks [21]. We wanted to know how gene replications affect the robustness of two renowned synthetic oscillators build in *E. coli*: The repressilator by Elowitz *et. al.* [28] and the dual feedback oscillator by Stricker *et. al* [29]. Using computational modeling, we show how the peak-to-peak time of the oscillators depend on the cell division time, the position of the genes on the DNA and the noise in the gene replication times.

We find that both synthetic oscillators can strongly lock to the cell cycle, where the oscillator's peak-to-peak time is equal to a multiple of the cell division time, over a wide range of division times. Remarkably, the effect strongly depends on how the genes of the oscillator are located on the chromosome. The distance between the genes introduces a temporal delay between the moments at which the different genes of the oscillators are replicated, which affects the period of the oscillations. Increasing the distance between genes has an opposite effect on the two oscillators: Whereas the repressilator exhibits almost no locking when the genes are positioned close together yet strong coupling over a wide range of T_d when the temporal delay is increased, the dual-feedback oscillator shows the strongest coupling to the cell cycle at negligible temporal delay between gene replications. For both models, the signature of the gene replication events should be clearly visible in the amplitude of the time traces of the protein concentrations.

It is well known that the timing of key events during the cell cycle, such as the start of DNA replication, the duration of chromosome replication and cell division, exhibit high levels of stochasticity [97, 98], which will propagate to the replication times of the oscillator's genes. To investigate how strong noise in the timing of gene replication affects

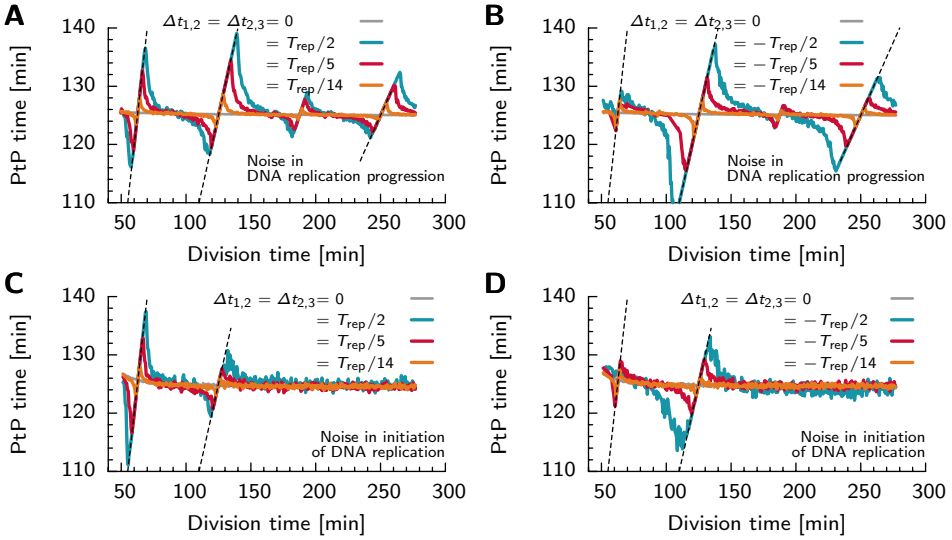

Figure 3.5: In the repressilator, stochasticity in the initiation of DNA replication plays the dominant role in attenuating the effects of gene replications. Top row, panels A and B, show results with only noise in the progression of DNA replication, $\sigma_{rep} = 0.35T_{rep}$, and the bottom row, panels C and D, corresponds to the situation where there is only noise in replication initiation, $\sigma_{init} = 0.2T_d$ (see Fig. 3.2). In both panels, solid lines show the peak-to-peak time T_{PtP} in the oscillations of P_1 as a function of the cell division time T_d. Standard deviation in T_{PtP} omitted for clarity, but is similar in all panels. Legends are defined in Fig. 3.1. (A,B) When there is noise in the time intervals between the gene replication events, but the initiation of DNA replication is fixed, locking seems little affected compared to the deterministic case (See Fig. 3.1, panels D and E). (C,D) When there is noise in the initiation of DNA replication, but the time between replications is fixed, the effects of the cell cycle almost disappear for division times $T_d > T_{int}$, in both ways of ordering the genes. However, strong locking persists at the 1:1 locking region and for $T_d < T_{int}$. Comparing with panels A and B, noise in the initiation of DNA replication seems to be more effective in protecting the clock against the cell cycle.

the oscillator's coupling to the cell cycle, we introduced two noise sources in our model of the cell cycle: one in the time of when DNA replication is initiated and one in the time it takes to replicate the chromosome. Using physiologically relevant values for the standard deviations in the timing, we found that noise in gene replication times strongly attenuates the effects of the cell cycle. However, observable signatures of locking remain for division times equal and shorter than the oscillator's intrinsic period. For these cells, the standard deviation in gene replication times becomes smaller than the oscillator's intrinsic period, making it possible for the clock to lock to a certain phase of the gene density, which sets the peak-to-peak time. We then asked which of these two sources is more important in attenuating the coupling between the cell cycle and the oscillator. To this end, we made two models for stochasticity in the replication times: One with only noise in replication progression and the other with only noise in the time of replication initiation. We found that noise in the initiation of DNA replication has a stronger effect

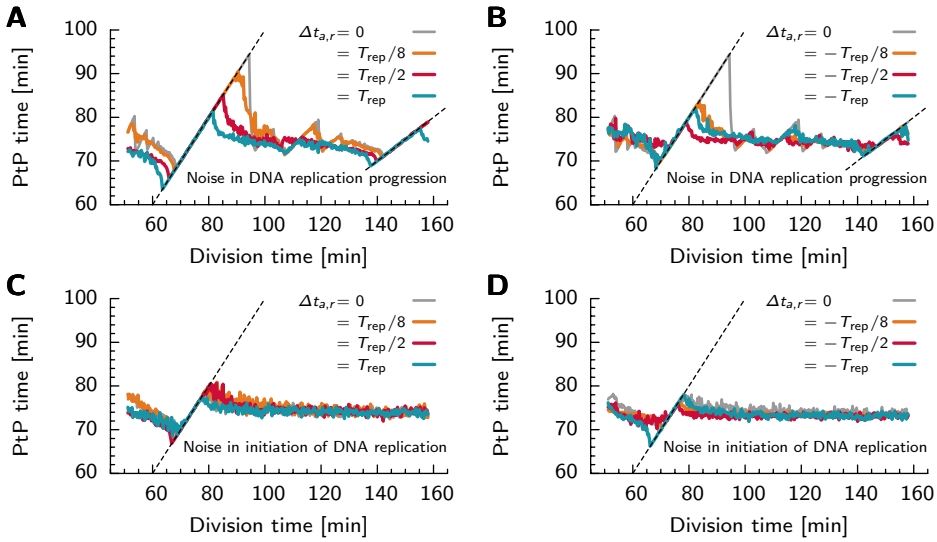

Figure 3.6: In the dual-feedback oscillator, stochasticity in the initiation of DNA replication plays the dominant role in attenuating the effects of gene replications. In both panels, solid lines show the peak-to-peak time in the activator concentration for different periods of the cell division time T_d. Standard deviation in T_{PtP} omitted for clarity, but is similar in all panels. Legends are defined in Fig. 3.3. We compare a scenario with only noise in DNA replication progression, with a standard deviation $\sigma_{rep} = 0.35T_{rep}$, panels A and B, to a scenario with only noise in the initiation of DNA replication, with a standard deviation $\sigma_{init} = 0.2T_d$, panels C and D. (A,B) As observed for the repressilator, with noise in replication progression but not in replication initiation, locking is little affected, compared to the deterministic case (See Fig. 3.3, panels D and E). (C,D) In the opposite scenario, with noise in the initiation of DNA replication but not in the progression of replication, most signatures of locking disappear, both when the activator or repressor gene is replicated first. Only around $T_d = T_{int}$, locking persists. Clearly, comparing with panels A and B, noise in the initiation of DNA replication has a stronger attenuating effect on locking.

than that in the progression of DNA replication. The reason is that, at physiologically motivated values, the standard deviation in the time of replication initiation is much larger than the standard deviation in the time of replicating the chromosome. We thus conclude that the initiation of DNA replication is mainly responsible for attenuating the effects of the gene replications on the repressilator and dual-feedback oscillator.

Throughout this work, we assume the genes reside on the bacterial chromosome. Importantly, however, the synthetic oscillators were originally constructed on plasmids, which are often present in large copy numbers ranging from 10-100. Moreover, experiments indicate that these plasmids are copied at random times during the major part of the cell cycle [99]. Based on our observation that multiple chromosome copies that are replicated asynchronously strongly reduce the strength of locking [21] (chapter 2), we expect that, at these high plasmid copy numbers, the synthetic oscillators exhibit no clear signatures of locking. Indeed, the original study on the dual-feedback oscillator does not report any effects from the cell cycle, even when the growth rate is comparable to the os-

cillator's intrinsic period where locking is expected to occur [29]. Remarkably, however, signatures of locking were observed for the dual-feedback oscillator in the experiments of [92], even though also in these experiments the genes reside on plasmids [92]. It is hard to explain what underlies the effect of the cell cycle in these experiments [92]. In any case, our analysis predicts that the effects will be much stronger when the genes are put on the chromosome. Conversely, in order to prevent locking, it seems beneficial to construct the oscillator on high copy number plasmids.

The genes of biological oscillators such as circadian clocks do reside on the chromosome, and the period of these oscillators are often unaffected by the cell cycle [100]. So, in line with Feynman's famous adage *"What I cannot create, I do not understand"*, as long as we can not create a robust oscillator that resides on chromosomes in growing and dividing cells, we do not fully understand all the challenges involved. In recent years, synthetic networks are increasingly being constructed directly onto the chromosome [79]. Therefore, implementing the synthetic oscillators on the bacterial chromosome allows us to investigate the mechanisms that are critically important for robust oscillations inside growing and dividing cells. For example, the dual-feedback oscillator studied in this work, based on a coupled positive and negative feedback architecture regulating gene expression, has been predicted to produce robust oscillations [23, 86]: The amplitude and period do not critically depend on specific parameter values, and oscillations persist in a wide range of temperatures and growth media [29, 89]. However, these models do not take the effect of gene replications into account, and in the experiments the genes reside on high copy-number plasmids potentially abolishing any effect of the cell cycle. Our results suggest that the relatively simple design of the dual-feedback oscillator implemented on the chromosome might not be very robust in growing and dividing cells, since its period scales with that of the cell cycle. Comparing the unstable synthetic oscillators with their evolved stable counterparts found in e.g. *S. elongatus* and *N. crassa*, could elucidate why the latter feature a remarkably more complex reaction network, including, for example, post-translational modification of the proteins [21, 101].

3.5. METHODS

Both the models of the repressilator and dual-feedback oscillator are described with ordinary differential equations, and propagated using Mathematica 8 (Wolfram Research). For each value of T_{d}, we generate a single time trace of about 200 oscillations for the repressilator and 100 oscillations for the dual-feedback oscillator. In order to allow the oscillations to settle down to a steady state, we discard the first 10 oscillations in the system.

To simulate the (stochastic) gene replication events, for each gene n in the model, we generate a list of replication times, τ_i^n, using Eqs. 3.4-3.6. The gene copy number for this gene, $g^n(t)$, equals 1 when $t < \tau_i^n$, and 2 when $t > \tau_i^n$, modulo T_{d}. The discrete gene replication events enter the models via the gene density, $G^n(t) = g^n(t)/V(t)$, where $V(t) = \exp(\ln(2)/T_{\mathrm{d}} \bmod(t, T_{\mathrm{d}}))$ is the cell volume [21].

To find the peak-to-peak times, T_{PtP}, in the ODE simulations (including those with noise in the gene replication times), we use the built-in methods of Mathematica to return all local extrema in the concentration of p_1 (repressilator) or the activator (dual-

Parameter	Value	Definition and motivation
α_{init}	0.2	Fraction of T_{d} when replication starts [95].
T_{rep}	40 min	Mean DNA replication time in *E. coli*.
$\Delta t_{1,2}/T_{\text{rep}}$	$0, \frac{1}{14}, \frac{1}{5}, \frac{1}{2}$	Time between gene replications (repressilator)
$\Delta t_{\text{a,r}}/T_{\text{rep}}$	$0, \frac{1}{8}, \frac{1}{2}, 1$	Time between gene replications (dual-feedback)
σ_{rep}	0.35 T_{rep}	SD in DNA replication progression [96].
σ_{init}	0.20 T_{d}	SD in initiation of DNA replication [96].
arab	0.70%	Arabinose level in dual-feedback oscillator.
[IPTG]	2nM	IPTG concentration in dual-feedback oscillator.

Table 3.1: Parameters used in the models. For the repressilator, we used the parameters given in Box 1 in [28]. For the dual-feedback oscillator we used the parameters given in the SI of [29]. SD stands for standard deviation.

feedback oscillator). These extrema correspond to the time points t_i where the concentration is higher, in the case of a maximum, or lower, in the case of a minimum, than its two immediate neighbors. As is standard for numerical solution of differential equations, the spacing $t_i - t_{i-1}$ between successive time points is determined adaptively by the algorithm to meet imposed precision bounds but never exceeded 0.2 h. We then checked if a given local minimum is the lowest point within an interval of $\pm$ 3/4 the oscillator's intrinsic period, T_{int}, centered on the minimum; if so, we define this point as the global minimum of a single oscillation cycle. If there exist a local extremum with a lower value, we repeated this procedure around the lower point until we found a point which was the lowest within a time interval of $\pm$ 3/4T_{int}. The same procedure is used for finding the local maxima of the oscillations. The peak-to-peak time is then calculated by subtracting the times of two consecutive minima; we verified that subtracting the times of two consecutive *maxima* gave essentially the same results.

4

A THERMODYNAMICALLY CONSISTENT MODEL OF THE POST-TRANSLATIONAL KAI CIRCADIAN CLOCK

The principal pacemaker of the circadian clock of the cyanobacterium S. elongatus is a protein phosphorylation cycle consisting of three proteins, KaiA, KaiB and KaiC. KaiC forms a homohexamer, with each monomer consisting of two domains, CI and CII. Both domains can bind and hydrolyze ATP, but only the CII domain can be phosphorylated, at two residues, in a well-defined sequence. While this system has been studied extensively, how the clock is driven thermodynamically has remained elusive. Inspired by recent experimental observations and building on ideas from previous mathematical models, we present a new, thermodynamically consistent, statistical-mechanical model of the clock. At its heart are two main ideas: i) ATP hydrolysis in the CI domain provides the thermodynamic driving force for the clock, switching KaiC between an active conformational state in which its phosphorylation level tends to rise and an inactive one in which it tends to fall; ii) phosphorylation of the CII domain provides the timer for the hydrolysis in the CI domain. The model also naturally explains how KaiA, by acting as a nucleotide exchange factor, can stimulate phosphorylation of KaiC, and how the differential affinity of KaiA for the different KaiC phosphoforms generates the characteristic temporal order of KaiC phosphorylation. As the phosphorylation level in the CII domain rises, the release of ADP from CI slows down, making the inactive conformational state of KaiC more stable. In the inactive state, KaiC binds KaiB, which not only stabilizes this state further, but also leads to the sequestration of KaiA, and hence to KaiC dephosphorylation. Using a dedicated kinetic Monte Carlo algorithm, we show that the model can describe a wealth of experimental data.

The content of this chapter has been submitted to PLoS Computational Biology.

4.1. INTRODUCTION

Circadian clocks, which allow organisms to anticipate changes in day and night, are a fascinating example of biological rhythms. One of the most studied and best characterized models of circadian oscillations is the cyanobacterium *Synechococcus elongatus*. It is now known that it combines a protein synthesis cycle [11, 102, 103] with a protein phosphorylation cycle [70], and in 2005 the latter was reconstituted in the test tube [12]. This stimulated a detailed characterization of its design principles, in a fruitful collaboration between experiments and modeling [14–17, 19, 21, 22, 104–108]. Yet, it has remained an open question how the clock is driven thermodynamically.

The central components of the protein phosphorylation cycle are three proteins, KaiA, KaiB, and KaiC. KaiC forms a hexamer with two phosphorylation sites per monomer, serine 431 and threonine 432 [11], which are phosphorylated and dephosphorylated in a well-defined temporal order [22, 60]. KaiA stimulates phosphorylation [109–111], while KaiB negates the effect of KaiA [110–113]. Modeling in combination with experiments indicate that the oscillations of the individual hexamers are synchronized via the mechanism of differential affinity: while KaiA stimulates phosphorylation, the limited supply of KaiA binds preferentially to those KaiC hexamers that are falling behind in the cycle, forcing the front runners to slow down and allowing the laggards to catch up [15, 22]. The mechanism of differential affinity appears to be active not only during the phosphorylation phase of the cycle [15, 18], but also during the dephoshorylation phase, when KaiC sequesters KaiA via the binding of KaiB [15, 22]. Monomer exchange between hexamers, observed in experiments [14], is an alternative synchronization mechanism. However, theoretical studies by us and others show that monomer exchange is not critical for stable oscillations [15, 22, 105].

While it is clear that the clock is driven by the turnover of ATP [114, 115], how fuel turnover drives the phosphorylation oscillations is still unclear. In previous models [15, 18, 26], phosphorylation is driven by ATP hydrolysis, while dephosphoryation proceeds via the spontaneous release of the phosphate groups from the threonine and serine residues. Intriguingly, however, recent experiments have revealed that during the dephosphorylation phase of the clock, ATP is regenerated [116, 117]: the phosphate groups on the serine and threonine residues are transferred back to ADP. If phosphorylation and dephosphorylation do not cause any net turnover of ATP, what then drives the clock? Clocks are necessarily dissipative, entailing a net turnover of fuel molecules per cycle.

KaiC consists of two highly homologous domains, called the CI and the CII domain [118]. Both domains can bind and hydrolyze ATP [119, 120], but only the CII domain can be phosphorylated [121]. The ATP regeneration experiments indicate that the phosphorylation and dephosphorylation of CII proceeds via the transfer of phosphate groups between the threonine/serine residues and the nucleotide bound to CII [116, 117], leaving open the possibility that there is no net turnover of ATP on the CII domain.

Here, we argue that the hydrolysis of ATP in the CI domain is the principal energetic driver of the clock. We present a new mathematical model of the post-translational Kai circadian clock in *S. elongatus*, which is based on the idea that ATP hydrolysis in CI drives a conformational switch of KaiC. Previously, it has been predicted that ATP hydrolysis plays an important role in driving conformational transitions [26, 122] and that these transitions are vital to generating the oscillations [15, 18], predictions that have found

experimental support [26, 122–126]. Our model, however, goes farther, predicting that ATP hydrolysis in the CI domain is the place where detailed balance must be broken in order to generate sustained oscillations. Our model is inspired by that of Van Zon *et al.* [15]. KaiC switches between an active conformation in which the phosphorylation level tends to rise, and an inactive one in which it tends to fall [15]. The model describes how KaiA binds to the CII domain of KaiC in the active conformation, and how KaiA can then drive phosphorylation by acting as a nucleotide exchange factor [127], stimulating the exchange of ADP for ATP. The model predicts that as the phosphorylation level in the CII domain rises, the release of ADP from CI slows down. The ADP-bound state makes the inactive conformation of KaiC more stable, causing the hexamer to flip to the inactive state and triggering dephosphorylation. In our model, ATP hydrolysis in the CI domain thus provides the thermodynamic driving force for the oscillations, while the phosphorylation in the CII domain provides the timer for the hydrolysis in the CI domain.

While the coupling between ATP hydrolysis in the CI domain and phosphorylation in the CII domain is the main feature of the new model, leading to novel testable predictions, the model can also describe a wealth of additional experimental data. The differential affinity of KaiA for the different KaiC phosphoforms naturally explains the characteristic sequence in which the threonine and serine sites are phosphorylated [22, 60]. In addition, while the slow release of ADP from CI triggers a switch between the active and inactive KaiC conformations, the binding of KaiB to CI stabilizes the inactive state further, and leads, as in previous models [15, 18, 22], to the sequestration of KaiA, necessary for synchronizing the oscillations. The model predicts that the slow binding of KaiB, as observed experimentally [27], introduces a delay between the moment that a given KaiC hexamer reaches its point of maximum phosphorylation, and hence no longer needs KaiA to progress along the phosphorylation cycle, and the moment that the same KaiC actually sequesters KaiA. In our model, this delay is essential because it allows the laggards (hexamers with a phosphorylation level lower than the mean) to reach the top of the cycle before the front runners (hexamers with a phosphorylation level higher than the mean) take away KaiA. Lastly, the model can explain the experimental observation that the oscillation period is robust to variations in steady-state ATP/ADP levels, while the system can be entrained by transient changes in this ratio, which is one of the important mechanisms for coupling the clock to light [30].

4.2. MODEL OVERVIEW

Our model of the in-vitro Kai circadian clock [12] builds on the hexamer model developed by Van Zon and coworkers [15]. But in contrast to that model, and following the models developed by Rust *et al.* [18, 22, 26], it explicitly keeps track of the two phosphorylation sites on each of the monomers, as well as their nucleotide binding states. The purpose of this section is to give an overview of the new model and its state variables, and to provide background information on ideas from previous models and their experimental justification. The new ingredients of the model, as well as their experimental movitation, are discussed only briefly; they are discussed in much more detail in the sections below.

KaiC Monomers Our model follows the phosphorylation and nucleotide-binding

state of each of the six monomers inside a hexamer. Each monomer consists of two domains: The CI and the CII domain. The CII domain has two phosphorylation sites, the threonine and the serine site, resulting in four different phosphorylation states [22, 60]: unphosphorylated (U), phosphorylated only on serine (S), phosphorylated only on threonine (T), and phosphorylated on both serine and threonine (D). Furthermore, both the CI and CII domains have a nucleotide binding pocket which can be in one of two possible states: Either there is adenosine triphosphate (ATP) or adenosine diphosphate (ADP) bound to it. The unbound state is ignored, because nucleotide binding is much faster than nucleotide dissociation, as described below. The state variables of the monomers are given in Table 4.1.

In the next sections, we describe in detail how ATP hydrolysis in CI drives the conformational switch of KaiC and how phosphorylation in CII controls hydrolysis in CI.

KaiC Hexamers Van Zon *et al.* postulated that KaiC can be in either an active (A) or inactive (I) conformation [15]. Experiments probing the exposure of the C terminal tails of KaiC and the stacking interactions between the CI and CII domains indeed provide evidence that KaiC can exist in multiple conformational states [122, 123, 125, 126]. We follow Van Zon *et al.*, and assume in the spirit of the Monod-Wyman-Changeux (MWC) model [128] that the CI and CII domains of all the monomers in a hexamer switch conformation in concert, such that we can speak of the hexamer as either being in the active or inactive state. Following Van Zon *et al.* our model does not include monomer exchange, which does not appear to be essential [15, 22, 105].

KaiB binding The phosphorylation behavior of KaiC in the presence of KaiB, but not KaiA, is highly similar to that of KaiC alone [22, 111, 112]. This observation indicates that KaiB does not directly affect the phosphorylation and dephosphorylation rates. Following Van Zon *et al.* [15], we assume instead that KaiB plays the following dual role: i) KaiB binding increases the stability of the inactive state by binding to the CI domain of inactive KaiC; experimental observations support the idea that the binding of KaiB to KaiC depends on the conformational state of KaiC [122, 125, 126]; moreover, the experiments show that KaiB binding peaks in the dephosphorylation phase of the cycle, when KaiC is in the inactive conformational state [22, 50, 129]; ii) KaiB associated with the CI domain of inactive KaiC strongly binds the limiting pool of KaiA, thereby sequestering it. In our model, we do not explicitly keep track of the KaiB concentration. Recent experiments show that KaiB binds KaiC as a monomer [126], but is energetically most stable as a tetramer [27]. Because of the equilibrium between the terameric and monomeric state of KaiB, the tetrameric state will act as a reservoir stabilizing the KaiB monomer concentration, making the latter less dependent on the total KaiB concentration [126]. Moreover, as predicted by Van Zon *et al.* [15], as long as KaiC can bind enough KaiB to sequester KaiA effectively, the concentration of KaiB does not affect the amplitude and period of the oscillations [130]. We therefore do not explicitly model the concentration of KaiB, but rather include it in the definition of the effective rate constant for KaiB-KaiC binding.

KaiA binding Experiments have unambiguously demonstrated that KaiA stimulates the phosphorylation of KaiC [11, 110–112]. Moreover, they indicate that in the absence of KaiB, KaiA binds to the CII domain [131]. Inspired by the recent observation that KaiA acts as a nucleotide exchange factor [127], our new model describes how KaiA bound

to CII is able to drive phosphorylation by controlling the nucleotide exchange rate. In the presence of KaiB, KaiA can also bind to the CI domain of inactive KaiC [122]. We do keep track of the KaiA dimers in the solution, and explicitly model their binding to the CII domain and their sequestration on the CI domain via KaiB. The interactions are always described as bimolecular reactions between KaiC hexamers and KaiA dimers or KaiB monomers. For simplicity, and lack of experimental evidence suggesting otherwise, the binding of KaiA or KaiB to KaiC always affects all monomers in the hexamer equally. In our model, a single KaiA dimer can bind to the CII domain of the hexamer, six KaiB monomers can bind to the CI domain of a KaiC hexamer and six KaiA dimers can, in turn, be sequestered by the CI domains of the hexamer in the inactive state.

State variables and parameters The state variables describing the hexamer and possible values are summarized in Table 4.1. The parameters of the model were obtained by fitting the predictions of the model to experimental data, as explained in the sections below. In this procedure, the parameters were "hand tuned"—we did not follow a systematic, formal, fitting procedure.

Monomer	
Variable	States
Phosphorylation	U,T,D,S
CI Binding pocket	ATP,ADP
CII Binding pocket	ATP,ADP

Hexamer	
Variable	States
Conformation	Active/Inactive
$n^{\text{CII·KaiA}}$	$0,1$
$n^{\text{CI·KaiB}}$	$0-6$
$n^{\text{CI·KaiA}}$	$0-6$

Table 4.1: Monomer and hexamer state variables, with possible values. Variables $n^{\text{CI·KaiA}}$ and $n^{\text{CII·KaiA}}$ count the number of KaiA dimers bound to the CI and CII domain, respectively, and $n^{\text{CI·KaiB}}$ counts the number of KaiB monomers bound to CI.

Chapter overview. We have split the explanation of the full model into two parts: First we give a detailed description of the phosphorylation dynamics in the Kai system which primarily concerns the CII domain of KaiC and its interaction with KaiA. In the next part, we describe the power cycle in the CI domain, the connection between the CI and CII domain and the binding and unbinding kinetics of KaiB and the subsequent sequestration of KaiA by the CI domain. Then we present the results for the model, again split into two sections: One relating to the phosphorylation dynamics and one to the power cycle in the CI domain.

4.3. MODEL OF THE KAIC PHOSPHORYLATION DYNAMICS

Here we give a detailed description of how the phosphotranfer reactions, the ratio of ATP to ADP in the binding pockets and differential affinity of KaiA together give rise to the ordered phosphorylation of the serine and threonine sites in the CII domain. In steps, we present the foundations of our model together with the experimental results that underlie it and give a detailed mathematical description of the resulting free energies and reaction rates.

4.3.1. PHOSPHORYLATION AND DEPHOSPHORYLATION ONLY OCCUR VIA PHOSPHOTRANSFER WITH ATP AND ADP

Recent experiments unexpectedly showed that during the dephosphorylation of KaiC, the inorganic phosphate group on the serine and threonine sites of KaiC is transfered to the ADP in the binding pocket of the CII domain, effectively regenerating the ATP that was used for phosphorylation [116, 117]. We hypothesize that in our model, dephosphorylation without a nucleotide as an intermediate does not occur. Therefore, the phosphotransfer reactions are the only pathways for phosphorylation and dephosphorylation of KaiC,

$$\begin{aligned} \mathrm{U\cdot ATP} &\underset{k^0_{\mathrm{TU}}}{\overset{k^0_{\mathrm{UT}}}{\rightleftharpoons}} \mathrm{T\cdot ADP}, \qquad \mathrm{T\cdot ATP} \underset{k^0_{\mathrm{DT}}}{\overset{k^0_{\mathrm{TD}}}{\rightleftharpoons}} \mathrm{D\cdot ADP}, \\ \mathrm{U\cdot ATP} &\underset{k^0_{\mathrm{SU}}}{\overset{k^0_{\mathrm{US}}}{\rightleftharpoons}} \mathrm{S\cdot ADP}, \qquad \mathrm{S\cdot ATP} \underset{k^0_{\mathrm{DS}}}{\overset{k^0_{\mathrm{SD}}}{\rightleftharpoons}} \mathrm{D\cdot ADP}. \end{aligned} \tag{4.1}$$

Here, U,T,D and S correspond to the phosphorylation state of the monomer and ATP and ADP denote the state of the CII nucleotide binding pocket. k^0_{XY} are the phosphotransfer rate constants when KaiA is not bound to CII. Since these rates are independent of the state of the other monomers, the monomers in a hexamer are phosphorylated in a random order. As is clear from Eq. 4.1, the phosphorylation dynamics critically depends on the state of the nucleotide binding pocket of the CII domain: With ATP in the binding pocket, KaiC can only be phosphorylated and with ADP in the binding pocket KaiC can only be dephosphorylated. Therefore, we explicitly keep track of the state of the nucleotide binding pocket adjacent to the serine and threonine sites of each monomer.

4.3.2. KAIA ACTS AS A NUCLEOTIDE EXCHANGE FACTOR ON THE BINDING POCKETS OF THE CII DOMAIN

Nucleotide exchange rate in absence of KaiA. Since the nucleotide binding rates are much faster than the dissociation rates, the unbound state can be neglected [116, 119, 120], and the nucleotide binding pocket will alternate only between the ATP and ADP bound state, both on CI and CII. Assuming that the association rates for ATP and ADP are diffusion limited and similar, the dynamics of nucleotide exchange will solely be governed by the nucleotide dissociation rates, which for the CII domain, discussed here, are denoted by $k_{\mathrm{off}}^{\mathrm{CII\cdot ATP}}$ and $k_{\mathrm{off}}^{\mathrm{CII\cdot ADP}}$, respectively. Next to nucleotide exchange, ATP can also be converted to ADP via hydrolysis with a rate $k_{\mathrm{hyd}}^{\mathrm{CII}}$ [114, 116]. Since hydrolysis is a strong downhill reaction under experimental conditions, we neglect the reverse reaction of the hydrolysis pathway, such that there is no ATP production by spontaneous binding of a phosphate group to ADP. The relative affinity between ATP and ADP for the nucleotide binding pocket of the CII domain is given by $K^{\mathrm{CII}}_{\mathrm{ATP/ADP}} = K_{\mathrm{d}}^{\mathrm{CII\cdot ATP}} / K_{\mathrm{d}}^{\mathrm{CII\cdot ADP}} = k_{\mathrm{off}}^{\mathrm{CII\cdot ATP}} / k_{\mathrm{off}}^{\mathrm{CII\cdot ADP}}$. Assuming that the nucleotide exchange and hydrolysis pathways are independent, we can simply add their reaction rate constants, such that the rates for

changing between the ATP and ADP bound states become

$$k^{\mathrm{CII}}_{\mathrm{ATP}\to\mathrm{ADP}} = k^{\mathrm{CII}}_{\mathrm{hyd}} + (1-\alpha_{\mathrm{ATP}})\, k^{\mathrm{CII\cdot ATP}}_{\mathrm{off}}, \tag{4.2}$$

$$k^{\mathrm{CII}}_{\mathrm{ADP}\to\mathrm{ATP}} = \alpha_{\mathrm{ATP}}\, k^{\mathrm{CII\cdot ADP}}_{\mathrm{off}}. \tag{4.3}$$

Here, α_{ATP} is the fraction, [ATP]/([ATP] + [ADP]), of ATP nucleotides in the solution. The rate from ATP to ADP (Eq. 4.2) is the sum of the hydrolysis rate plus the rate of dissociating ATP times the probability of immediately binding an ADP, which due to the equal association rates for ADP and ATP binding, is simply given by the fraction of ADP in the bulk $(1-\alpha_{\mathrm{ATP}})$. The reverse rate (Eq. 4.3) is given by the rate of dissociating ADP times the probability, α_{ATP}, of binding an ATP nucleotide.

KaiA speeds up nucleotide exchange on CII domain. It is well known that KaiA stimulates the phosphorylation of KaiC [11, 110], and that without KaiA, KaiC dephosphorylates [111, 112]. Recent experiments showed that KaiA increases the fraction of ATP in the nucleotide binding pockets and thereby stimulates phosphorylation [127, 132]. Without KaiA bound to the CII domain, the exchange rates between the nucleotide binding pocket and the bulk are very low, such that eventually all ATP molecules bound to the binding pockets of the CII domain are hydrolyzed [116, 127]. Given these observations, in our model KaiA will act as a nucleotide exchange factor increasing the dissociation rates of both ATP and ADP in equal amounts, such that the relative affinity, $K^{\mathrm{CII}}_{\mathrm{ATP/ADP}}$, is unchanged.

We model the interaction between KaiA and the CII domain on three simplifying assumptions: First, only one KaiA dimer can bind to the CII domain of a hexamer, although a higher stoichiometry has been observed [52, 133]. Second, when KaiA is bound to CII, it enhances the nucleotide exchange rates in all the monomer binding pockets equally. Third, when KaiA is not bound, the nucleotide dissociation rates are zero. Therefore, the nucleotide dissociation rates given in Eqs. 4.2 and 4.3, are given by $k^{\mathrm{CII\cdot ADP}}_{\mathrm{off,KaiA}}\, K^{\mathrm{CII}}_{\mathrm{ATP/ADP}}$ and $k^{\mathrm{CII\cdot ADP}}_{\mathrm{off,KaiA}}$, respectively, when KaiA is bound to CI, and equal zero when KaiA is absent. The hydrolysis rate, $k^{\mathrm{CII}}_{\mathrm{hyd}}$, does not depend on whether KaiA is bound to the CII domain, in the interest of simplicity.

KaiA stimulates phosphorylation by speeding up nucleotide exchange. Fig. 4.1 illustrates how KaiA can enhance the ATP fraction in the nucleotide binding pocket of the CII domain. KaiA does not change the affinity of CII for ATP and ADP, and hence also leaves their relative affinity unchanged. However, it does increase the binding and unbinding rates with the same magnitude. Moreover, the binding of the nucleotides is coupled to the non-equilibrium process of ATP hydrolysis, which breaks detailed balance. The result is that in the absence of KaiA, the hydrolysis rate dominates over the nucleotide exchange rates, driving the binding pockets towards the ADP state. In the presence of KaiA, exchange rates are larger than the hydrolysis rate, and because these rates favor ATP over ADP, the binding pocket is predominantly bound to ATP when KaiA is present. By increasing the occupation of ATP of the binding pocket, KaiA not only enhances phosphorylation but also blocks dephosphorylation since there is no ADP. Blocking dephosphorylation, which was implicitly present in previous models [15, 18, 22], is important, as it prevents futile phosphorylation cycles.

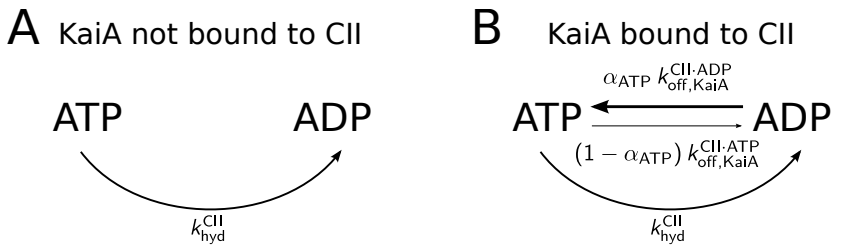

Figure 4.1: KaiA regulates the fraction of ATP in the CII binding pockets by increasing the nucleotide dissociation rates, in combination with the ATPase activity in the CII domain. A) Without KaiA bound to CII, the nucleotide dissociation rates are identically zero, and there is no nucleotide exchange with the bulk. Since ATP is hydrolyzed at a constant rate, $k_{\mathrm{hyd}}^{\mathrm{CII}}$, eventually all the binding pockets will be occupied by ADP. B) When KaiA is bound to the CII domain, it increases the dissociation rates of ADP and ATP, $k_{\mathrm{off}}^{\mathrm{CII\cdot ATP}}$ and $k_{\mathrm{off}}^{\mathrm{CII\cdot ADP}}$, respectively, while leaving the affinities for ATP and ADP unchanged. Now, ADP in the CII domain is replaced by ATP at a rate that is faster than that at which ATP is hydrolyzed; this indeed increases the fraction of ATP in the binding pockets. For simplicity, we assumed equal diffusion limited association rates for ATP and ADP, such that the probability of binding ATP after ADP has dissociated is equal to α_{ATP}. The rate of the reverse pathway is proportional to the bulk ADP fraction, $1-\alpha_{\mathrm{ATP}}$.

4.3.3. DIFFERENTIAL AFFINITY OF THE CII DOMAIN FOR KAIA CHANGES THE PHOSPHOTRANSFER RATES

It was predicted from theoretical arguments [15], and later confirmed by experiments [18, 53, 134], that the affinity of KaiC for KaiA depends on the phosphorylation state of the hexamer. By measuring the phosphorylation speed of KaiC, starting with different initial phosphorylation levels, Rust *et. al.* observed that the rate of KaiC phosphorylation decreases as the fraction of S and D phosphorylated KaiC monomers increases [18]. This suggests that KaiA has a high affinity when KaiC is in the U and T state, and a low affinity when KaiC is in the S and D state, leading to the mechanism of differential affinity [15].

Differential affinity means that the binding and unbinding rates of KaiA to the CII domain of KaiC depend on the phosphorylation state of the hexamer. The observation that KaiA predominantly binds to CII during the phosphorylation phase of the cycle indicates that, in addition, KaiA has a higher affinity for the active conformational state. In our model, when all the monomers of an active hexamer are in the unphosphorylated U state, KaiA will bind and unbind with the rates $k_{\mathrm{on,0}}^{\mathrm{CII\cdot KaiA}}$ and $k_{\mathrm{off,0}}^{\mathrm{CII\cdot KaiA}}$, respectively. The subsequent phosphorylation of KaiC changes the binding free energy $\Delta G_{\mathrm{bind}}^{\mathrm{CII\cdot KaiA}}$: this indeed underlies the mechanism of differential affinity. Assuming each monomer adds linearly to $\Delta G_{\mathrm{bind}}^{\mathrm{CII\cdot KaiA}}$, the change in the binding free energy between KaiA and CII becomes

$$\Delta G_{\mathrm{bind}}^{\mathrm{CII\cdot KaiA}} = \sum_{i=1}^{6} \delta g_{\mathrm{bind}}^{\mathrm{CII\cdot KaiA}}(X_i) + h_{\mathrm{Inactive}}\, \delta g_{\mathrm{A,I}}^{\mathrm{CII\cdot KaiA}}. \tag{4.4}$$

Here, $\delta g_{\mathrm{bind}}^{\mathrm{CII\cdot KaiA}}(X_i)$ is the contribution of each monomer in phosphorylation state $X_i \in \{\mathrm{U,T,D,S}\}$ to the binding free-energy. KaiA bound to the CII domain stabilizes the ac-

tive conformational state with a fixed free-energy difference $\delta g_{\mathrm{A,I}}^{\mathrm{CII\cdot KaiA}}$ and h_{Inactive} is an indicator function that is one when the hexamer is in the inactive state and zero otherwise. Note that the stabilization of the active state with respect to the inactive one does not depend on the phosphorylation state, and therefore the hexamer's conformation will not affect the phosphotransfer dynamics [22, 26]. Given the effect of differential affinity on the binding free energy, $\Delta G_{\mathrm{bind}}^{\mathrm{CII\cdot KaiA}}$, detailed balance dictates that the association and dissociation rates of KaiA become, respectively,

$$k_{\mathrm{on}}^{\mathrm{CII\cdot KaiA}} = k_{\mathrm{on},0}^{\mathrm{CII\cdot KaiA}} \exp\left(-(1-\lambda)\,\Delta G_{\mathrm{bind}}^{\mathrm{CII\cdot KaiA}}\right) \tag{4.5}$$

$$k_{\mathrm{off}}^{\mathrm{CII\cdot KaiA}} = k_{\mathrm{off},0}^{\mathrm{CII\cdot KaiA}} \exp\left(\lambda\,\Delta G_{\mathrm{bind}}^{\mathrm{CII\cdot KaiA}}\right). \tag{4.6}$$

Assuming changes in binding free-energy have an equal effect on the association and dissociation rate, $\lambda = 1/2$.

KaiA binding changes the phosphortansfer rates We model phosphorylation and dephosphorylation, with and without KaiA, via a microscopically reversible phosphotranfer reaction. Then, as an inevitable consequence of differential affinity, detailed balance implies that the binding of KaiA must also influence the phosphotransfer rates [135]. We denote the free-energy difference between phosphorylation states X and Y, with X,Y$\in$ {U, T, D, S}, when KaiA is not bound to KaiC as $\delta g_{\mathrm{XY}}^{0}$ and when KaiA is bound as $\delta g_{\mathrm{XY}}^{\mathrm{KaiA}}$. Detailed balance then implies that the difference between these two, $\delta g_{\mathrm{XY}}^{0} - \delta g_{\mathrm{XY}}^{\mathrm{KaiA}}$, is equal to the change in the binding free-energy of KaiA that results from a change in the phosphorylation state X_i to Y_i of monomer i:

$$\delta g_{\mathrm{XY}}^{0} - \delta g_{\mathrm{XY}}^{\mathrm{KaiA}} = \Delta G_{\mathrm{bind}}^{\mathrm{CII\cdot KaiA}}(\mathrm{Y}_i) - \Delta G_{\mathrm{bind}}^{\mathrm{CII\cdot KaiA}}(\mathrm{X}_i) \tag{4.7}$$

$$= \delta g_{\mathrm{bind}}^{\mathrm{CII\cdot KaiA}}(\mathrm{Y}) - \delta g_{\mathrm{bind}}^{\mathrm{CII\cdot KaiA}}(\mathrm{X}), \tag{4.8}$$

where $\delta g_{\mathrm{XY}}^{0} = -\log\left(k_{\mathrm{XY}}^{0}/k_{\mathrm{YX}}^{0}\right)$, with k_{XY}^{0} given by Eq. 4.1, and $\delta g_{\mathrm{XY}}^{\mathrm{KaiA}} = -\log\left(k_{\mathrm{XY}}^{\mathrm{KaiA}}/k_{\mathrm{YX}}^{\mathrm{KaiA}}\right)$. The second line follows after substituting Eq. 4.4 for $\Delta G_{\mathrm{bind}}^{\mathrm{CII\cdot KaiA}}$. The subscript i in Eq. 4.7 is to emphasize that we compare hexamers that differ in the phosphorylation state of one of their monomers. Finally, we can write the phosphotransfer rate constants between states X and Y, for the situation where KaiA is bound to CII, as

$$k_{\mathrm{XY}}^{\mathrm{KaiA}} = k_{\mathrm{XY}}^{0} \exp\left(-\frac{1}{2}\left[\delta g_{\mathrm{bind}}^{\mathrm{CII\cdot KaiA}}(\mathrm{Y}) - \delta g_{\mathrm{bind}}^{\mathrm{CII\cdot KaiA}}(\mathrm{X})\right]\right), \tag{4.9}$$

and the reverse reaction rate by interchanging labels X and Y. This equation indeed shows that the phosphotranfer rates in the presence of KaiA, $k_{\mathrm{XY}}^{\mathrm{KaiA}}$, depend on how the phosphorylation states change the affinity for KaiA and the energy levels of the monomers, depicted in Fig. 4.2.

4.3.4. ORDERED PHOSPHORYLATION OF THE S AND T SITES DUE TO PHOSPHOTRANSFER DYNAMICS, NUCLEOTIDE EXCHANGE AND DIFFERENTIAL AFFINITY

It is well known that the S and T sites in each monomer are sequentially phosphorylated through the cycle: U → T → D → S → U [22, 60]. In previous models, this order

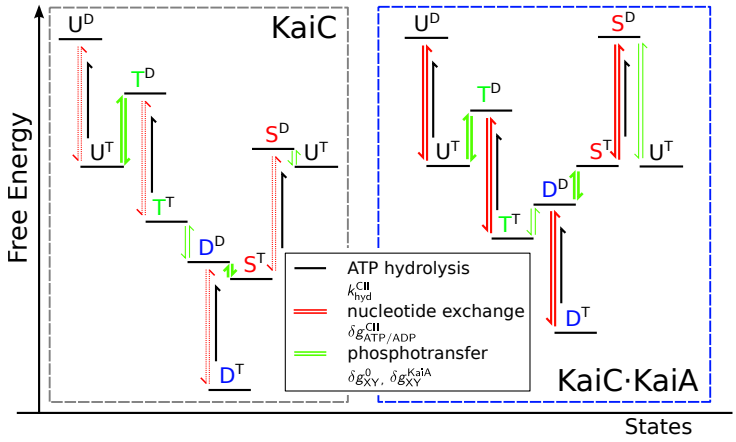

Figure 4.2: Proposed free energy landscape of the CII domain, without (left gray box) and with (right blue box) KaiA bound to the CII domain, for equal concentrations of ATP and ADP in the bulk. Superscripts denote the nucleotide bound state of CII. The landscape results from hand-tuning the phosphotransfer and nucleotide exchange parameters as well as the free-energy changes due to KaiA binding to give best agreement with the (de)phosphorylation assays performed in the SI of [26]. The free energy differences between phosphorylation states are defined by the detailed-balance relations, $\delta g^0_{XY} = -\log\left(k^0_{XY}/k^0_{YX}\right)$ and $\delta g^{KaiA}_{XY} = -\log\left(k^{KaiA}_{XY}/k^{KaiA}_{YX}\right)$ without and with KaiA bound to CII (connected by green arrows), respectively. Thick and thin green arrows indicate high and low phosphotransfer rates, respectively. Furthermore, $\delta g^{CII}_{ATP/ADP} = \log\left(K^{CII}_{ATP/ADP}\right)$, gives the free-energy difference between nucleotide bound states (connected by red arrows). Note that the relative nucleotide affinity is independent of the KaiA-binding state or phosphorylation state. Dotted red arrows indicate the nucleotides exchange rates are identically zero when KaiA is not bound to CII, and solid red arrows in the right panel indicate high nucleotide exchange rates when KaiA is bound. The black arrows denote the irreversible hydrolysis of ATP. For hydrolysis, the magnitude of the arrow does not give the free energy change of the system since that involves the contribution from the phosphate release, which we do not take into account. Note that without hydrolysis, the total free-energy drop upon going through the cycle, $U^T \to T^D \to T^T \to D^D \to S^T \to S^D \to U^T$, is zero.

was imposed by choosing different effective rate constants between the phosphorylation states, in situations of KaiA bound to KaiC, and KaiA not bound to KaiC [15, 18, 22]. However, given the properties of the phosphotransfer reactions and the effects of KaiA as described above, we can now have a more detailed understanding of what makes the phosphorylation cycle go round.

At the start of the phosphorylation phase of the oscillation, the majority of monomers are in the non-phosphorylated state (U), with an ADP in the nucleotide binding pocket and a KaiA bound to the CII domain. The binding of KaiA will enhance the nucleotide exchange rate which increases the fraction of ATP in the binding pocket and consequently forces the phosphotransfer reactions in Eq. 4.1 towards the phosphorylated state while in the mean time blocking the reverse reaction. The next question is why the threonine

residue is phosphorylated before the serine residue. Here, differential affinity plays a key role because KaiA binding lowers the free-energy of the T-state and increases that of the S phosphorylation state, as is shown in Fig. 4.2. Together with the fact that the phosphotransfer rate constants of the T site are much faster than those of the S site [22, 117], the threonine residues are phosphorylated before the serine residues. Differential affinity thus has two important effects: Not only does it help to synchronize the KaiC hexamers as found in [15], but it also enforces the correct order of phosphorylation. Since a T-phosphorylated hexamer will still have a high affinity for KaiA, the ATP fraction in the binding pockets will remain high such that eventually both the serine and threonine sites are phosphorylated and the monomers arrive in the D state.

In the dephosphorylation phase of the oscillation, when all KaiA is sequestered and therefore no nucleotide exchange is possible in the CII domain, the ATP in the CII binding pockets will eventually be hydrolyzed. With ADP in the binding pocket, phosphotransfer reactions can occur causing dephosphorylation [116]. Without KaiA bound to CII, the serine residue becomes energetically favorable over the threonine residue again and because phosphotransfer with the threonine residue is faster than with the serine residue, (meaning that the D $\rightleftharpoons$ S transitions are faster than the D $\rightleftharpoons$ T transitions) the majority of the D phosphorylated monomers will proceed to the S state instead of the T state. The S-site will slowly further dephosphorylate to the U state. This shows how differential affinity and nucleotide exchange together give rise to the ordered phosphorylation of the monomers.

4.3.5. ADP IN SOLUTION SLOWS DOWN PHOSPHORYLATION

Experiments show that the fraction of ATP in solution, α_{ATP} =[ATP]/([ATP]+[ADP]), has a significant effect on the phosphorylation speed and the amplitude of the oscillations in the in-vitro system [7, 26, 30]. They also show that this sensitivity to the ATP fraction is the primary input for entraining the oscillator to the daily day-night cycle [7, 26, 30]. As explained above, our model exhibits this sensitivity because the binding probabilities for ATP and ADP to the binding pocket of the CII domain are directly proportional to the ATP and ADP fraction, respectively, as given in Eq. 4.2 and Eq. 4.3.

4.4. MODEL OF THE KAIC POWER CYCLE

In the previus section we discussed the phosphorylation dynamics of KaiC and the interaction between its CII domain and KaiA, and how these effects combine to generate the ordered phosphorylation of the threonine and serine sites in KaiC. The CI domain does not seem to play a crucial role here, in particular since the phosphorylation dynamics in the presence of KaiA only, is unaffected in a KaiC mutant where hydrolysis in the CI domain is deactivated [26]. This raises the question of what role the CI domain fulfills in the Kai oscillator. Here we describe how hydrolysis in CI together with the binding of KaiB to the CI domain drives the conformational switch of the hexamer and how the slow binding of KaiB, together with the subsequent sequestration of KaiA, helps to synchronize the ensemble of KaiC hexamers.

4.4.1. NUCLEOTIDE EXCHANGE IN AND KAIB BINDING TO THE CI DOMAIN DRIVES THE CONFORMATIONAL SWITCH OF KAIC

KaiB and ADP binding to CI is cooperative. Experiments show that the binding of KaiB requires catalytic activity of the CI domain, since a mutant that lacks the hydrolysis site does not bind KaiB [26]. Furthermore, when KaiB is added to a solution with only KaiC and ATP, the ATPase rate drops significantly [114]. Because KaiA is not present in both experiments, the ATPase activity in the CII domain is negligible as explained in the section on KaiA acting as a nucleotide exchange factor, such that the change in the ATPase rate can be attributed to changes in the CI domain. Given these results, it seems likely that the affinity of KaiB for KaiC depends on ADP in the CI binding pockets created by ATP hydrolysis, and that, vice versa, KaiB binding stabilizes the binding of ADP [126].

In our model, the conformational switch from the active to inactive state depends on ATP hydrolysis in the CI domain and the binding of KaiB to the CI domain. Specifically, both KaiB and ADP binding to CI stabilize the inactive conformational state. It is a characteristic of the MWC model that this introduces an effective cooperativity between KaiB and ADP: KaiB binding enhances the probability that KaiC is in the inactive state, in which ADP will then remain bound more strongly; conversely, ADP in CI will increase the likelihood that KaiC is in the inactive state, in which it will bind KaiB more strongly.

KaiB and ADP binding to CI stabilize the inactive conformational state. For simplicity, there is no direct cooperativity between ADP and KaiB binding, such that the free-energy difference between the active and inactive conformation of the hexamer is proportional to the number of ADP nucleotides, $n^{\text{CI·ADP}}$, and KaiB monomers, $n^{\text{CI·KaiB}}$,

$$\Delta G_{\text{A,I}}^{\text{hex}} = \left(n^{\text{CI·ADP}} - n_0^{\text{CI·ADP}}\right) \delta g_{\text{A,I}}^{\text{ATP,ADP}} + n^{\text{CI·KaiB}} \delta g_{\text{A,I}}^{\text{CI·KaiB}}. \tag{4.10}$$

Here, $\delta g_{\text{A,I}}^{\text{ATP,ADP}} = \delta g_{\text{I}}^{\text{ATP,ADP}} - \delta g_{\text{A}}^{\text{ATP,ADP}}$, is the difference in the free-energy increase upon converting one ATP into an ADP in the CI domain, between the inactive and active conformational state of the hexamer. Since the experiments indicate that the stability of the inactive state increases with the number of bound ADP molecules as discussed above, ADP needs to have a higher affinity for the CI domain in the inactive state as compared to the active state. On the other hand, the exchange of ADP for ATP should be energetically favorable, $\delta g_{\text{I}}^{\text{ATP,ADP}} > 0$, such that ADP can be exchanged spontaneously at the end of the phosphorylation cycle. These two conditions can be satisfied by choosing $\delta g_{\text{A}}^{\text{ATP,ADP}} > \delta g_{\text{I}}^{\text{ATP,ADP}} > 0$ such that $\delta g_{\text{A,I}}^{\text{ATP,ADP}} < 0$. These conditions indeed ensure that ATP hydrolysis stabilizes the inactive state, while still allowing for spontaneous ADP release at the end of the cycle. The free-energy difference between the active and inactive state when all CI binding pockets have ATP bound, is set in Eq. 4.10 via the parameter $n_0^{\text{CI·ADP}}$, and determines the threshold number of bound ADP molecules that are required to make the inactive state more stable that the active one.

Finally, the free-energy contribution to $\Delta G_{\text{A,I}}^{\text{hex}}$, from binding a single KaiB monomer follows from the dissociation constants for KaiB binding to the active and inactive state, $K_{\text{d,A}}^{\text{CI·KaiB}}$ and $K_{\text{d,I}}^{\text{CI·KaiB}}$, respectively, via the detailed balance relation $\delta g_{\text{A,I}}^{\text{CI·KaiB}} = -\log\left(K_{\text{d,A}}^{\text{CI·KaiB}} / K_{\text{d,I}}^{\text{CI·KaiB}}\right)$. A higher affinity of KaiB for inactive KaiC than for active KaiC, $K_{\text{d,A}}^{\text{CI·KaiB}} / K_{\text{d,I}}^{\text{CI·KaiB}} > 1$, means that KaiB binding stabilizes the inactive state,

$\delta g_{\mathrm{A,I}}^{\mathrm{CI\cdot KaiB}} < 0$. Eq. 4.10 thus shows how ADP and KaiB binding stabilize the inactive conformational state of KaiC.

4.4.2. TIMING OF THE CONFORMATIONAL SWITCH IS DETERMINED BY PHOSPHORYLATION OF CII DOMAIN, WHICH SETS ADP DISSOCIATION RATE IN CI DOMAIN

The switch from the active to the inactive state is driven energetically by hydrolysis of ATP in the CI domain and the subsequent binding of KaiB. But what exactly sets the timing of this switch?

Phosphorylation of the CII domain controls ADP dissociation rate in CI domain. Interestingly, experiments show that the binding of KaiB requires not only the hydrolysis of ATP in CI, but also that the CII domain is phosphorylated at least on the serine residue [26]. The latter observation might be the result of a direct interaction of KaiB with the CII domain, but could also be due to an indirect effect, in which the likelihood that CI is bound to ADP (which enhances KaiB binding), depends on the phosphorylation state of CII. The latter hypothesis is supported by the experimental observation that hyperphosphorylated KaiC has a lower ATPase activity and a higher fraction of ADP in the binding pockets, as compared to non-phosphorylated KaiC [114, 115, 127]. As mentioned before, the lower measured ATPase activity in hyperphosphorylated KaiC must, because of the absence of KaiA in these experiments, be attributed to the CI domain, and not to changes in the CII domain. The lower ATPase rate is the result of a lower hydrolysis rate and/or a lower ADP dissociation rate. However, a lower hydrolysis rate with a constant ADP dissociation rate would lead to a lower ADP fraction in the binding pockets, in contrast to what has been observed experimentally [116, 127]. We thus conclude that the phosphorylation state of CII determines the ATPase rate of CI through the ADP dissociation rate: As serine residues on CII become phosphorylated, the ADP dissociation rate at the CI domain decreases. These arguments indicate that the regulatory mechanism that controls the timing of the conformational switch is the dependence of the dissociation rate of ADP in the CI domain on the phosphorylation state of the CII domain.

The ADP-CI association and dissociation rates change, but their ratio, the affinity, does not. A question not yet answered is whether the phosphorylation of the CII domain changes the magnitude of the ADP binding rates, or also the affinity, which depends on the ratio of the association and dissociation rates. Recent experiments allow us to answer this question. These experiments show that both the phosphorylation and dephosphorylation rates are unchanged in a KaiC mutant where ATP hydrolysis in the CI domain is deactivated, which decreases the fraction of bound ADP [26]. Detailed balance would entail that if phosphorylation of CII were to stabilize ADP in CI, then vice versa ADP would stabilize the phosphorylated state; KaiC hexamers with fewer ADP molecules in CI, such as the KaiC mutant in [26], would then dephosphorylate faster. Fig. 4.1A illustrates how the detailed balance condition changes the phosphotranfer rates in this case. Since the experiments show that the (de)phosphorylation rates are unchanged [26], we must conclude that the affinity of ADP for CI does not depend on the phosphorylation state of the CII domain. Hence not only the dissociation rate of ADP changes, but also the association rate changes by the same factor, leaving the affinity unchanged.

Phosphorylation of CII controls ADP dissociation from CI via transition state. To explain the dependence of the ADP fraction in the CI domain on the phosphorylation state of the CII domain, we envision a model in which the phosphorylation state of CII affects a short-lived transition state for the dissociation of ADP from the nucleotide binding pocket of CI. In our model, the activation energy for ADP dissociation in each monomer, $\Delta G_{\text{act}}^{\text{CI·ADP}}$, depends linearly on the phosphorylation state X_i of all monomers in the hexamer

$$\Delta G_{\text{act}}^{\text{CI·ADP}} = \sum_{i=1}^{6} \delta g_{\text{act,A/I}}^{\text{CI·ADP}}(X_i), \tag{4.11}$$

where $\delta g_{\text{act,A/I}}^{\text{CI·ADP}}$ is the contribution of a single monomer on the activation energy in the active (A) or inactive (I) conformational state. The ADP dissociation rate is then given by

$$k_{\text{off}}^{\text{CI·ADP}} = k_{\text{off,0}}^{\text{CI·ADP}} \exp\left(-\Delta G_{\text{act}}^{\text{CI·ADP}}\right), \tag{4.12}$$

where $k_{\text{off,0}}^{\text{CI·ADP}}$ is the off-rate when the hexamer is in the active state with all monomers in the U-state.

Fig. 4.1B shows the ADP dissociation rate as a function of the number of monomers in the S state for a hexamer in the inactive conformation, assuming that the other monomers are in the U state, which is typically the case during the dephosphorylation phase of the cycle. The energy values $\delta g_{\text{act,I}}^{\text{CI·ADP}}(\text{X})$ determine when the rate of ADP dissociation (typically leading to ATP binding) will be higher than the ATP hydrolysis rate (leading to the ADP bound state). We choose the energy values such that the crossover in the rates happens between 1 and 2 monomers in the S-state. Hence, when a hexamer has fewer than 2 monomers in the S-state, ADP will be released and ATP becomes bound, and the hexamer will switch back to the active conformation, completing the cycle.

ADP-CI association rate is very low and ADP arises only via hydrolysis of bound ATP. Because ADP in the CI binding pocket is energetically very unfavorable [118–120], and the affinity of KaiC for KaiB does not seem to depend on the bulk ATP fraction [7, 26], the association rate of ADP from the bulk to the CI binding pocket will be very low. Therefore, in our model, regardless of the bulk ATP fraction, ADP can only appear in the CI binding pocket through the hydrolysis of ATP:

$$k_{\text{ATP}\to\text{ADP}}^{\text{CI}} = k_{\text{hyd}}^{\text{CI}}. \tag{4.13}$$

Moreover, since the ADP association rate is assumed to be zero, after ADP dissociation the pocket will always bind ATP. Since ATP association is much faster than nucleotide dissociation [127], as exploited in the section on phosphorylation dynamics, the rate of exchanging ADP for ATP is simply given by the ADP dissociation rate:

$$k_{\text{ADP}\to\text{ATP}}^{\text{CI}} = k_{\text{off}}^{\text{CI·ADP}}, \tag{4.14}$$

with the latter given by Eq. 4.12.

4.4.3. T AND S PHOSPHORYLATION STATE HAVE AN ANTAGONISTIC EFFECT ON THE ADP FRACTION IN CI

Experiments show that a phosphomimetic of the T state of KaiC has a higher ATPase activity compared to unphosphorylated KaiC and that a phosphomimetic of the S state has

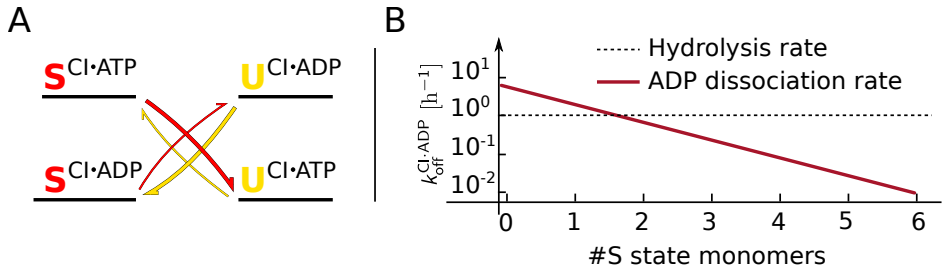

Figure 4.1: The phosphorylation state of the CII domain regulates the ADP fraction in the CI domain by changing the ADP release rate, but not its affinity. (A) Cartoon that illustrates why the experiments rule out the latter "affinity" scenario. The cartoon shows a free-energy landscape for this affinity scenario in which the phosphorylation state of CII determines the affinity of ADP for CI; specifically, the cartoon illustrates the case where a monomer in the S state stabilizes ADP in the binding pocket of the CI domain. The superscripts denote the state of the CI binding pocket. The S-phosphorylated state lowers the free energy of the CI·ADP state compared to the U state, thereby decreasing the ADP-off rate in the CI binding pocket. It follows from detailed balance that when the CI domain has ADP bound, the phosphotransfer from S to U is uphill in free energy, which decreases the phosphotransfer rate. The diagonal arrows give the phosphotransfer reactions from U to S (yellow) and S to U (red). When the CI domain is predominantly in the ATP state, as was done in [26] by lowering the hydrolysis rate with a mutation, the affinity scenario, as illustrated in the cartoon, would predict that the dephosphorylation rate from S to U increases. However, the experiments show no significant change in phosphotransfer rates, making this affinity scenario unlikely. Based on these arguments, we predict that phoshorylation of CII does not change the affinity, but does change the absolutes rates of ADP dissociation and association. (B) ADP dissociation rate as a function of the number of S state monomers in the hexamer, assuming the other monomers are in the U state. Between 1 and 2 monomers in the S state, the dissociation rate is higher than the hydrolysis rate which puts the binding pockets of CI domain predominantly in the ATP state, driving the hexamer to the active conformational state.

a lowered ATPase activity [136]. This is likely the result of an antagonistic effect of phosphorylation of the threonine and serine sites on the dissociation rate of ADP. Specifically, in our model monomers in the T state will lower the activation energy for ADP release, $\delta g_{\mathrm{act}}^{\mathrm{CI\cdot ADP}}(\mathrm{T}) < 0$, while monomers in the S state increase the activation energy, $\delta g_{\mathrm{act}}^{\mathrm{CI\cdot ADP}}(\mathrm{S}) > 0$. Because the S and T sites are orderly phosphorylated, their antagonistic effect on the dissociation rate of ADP will create a sharp transition between the phase in which the ADP fraction in CI is low and that in which it it is high [18]. Furthermore, just like in the push-pull network studied by Goldbeter *et al.* [61], the ATP fraction in the CI domain will depend on the *difference* between the monomers in the S and T state, and not on their absolute number. This makes the regulation of the CI domain less sensitive to the absolute phosphorylation level, which depends on the bulk ATP fraction [7, 26]. The effects on this we explore in detail in chapter 5.

4.4.4. NUCLEOTIDES, KAIB AND KAIA BINDING TO THE CI DOMAIN STABILIZES THE INACTIVE STATE

Detailed balance implies that the different affinities of the binding partners for the active and inactive state of KaiC, is reflected in the free-energy difference between the two conformations, $\Delta G^{\text{hex}}_{\text{A,I}}$. In our model, we assume there is no cooperative binding to KaiC in a given conformational state (although the MWC model introduces an effective cooperativity as explained in the theory section on the power cycle), such that we can split the dependence on each binding partner in independent terms,

$$\begin{aligned}\Delta G^{\text{hex}}_{\text{A,I}} &= \left(n^{\text{CI·ADP}} - n_0^{\text{CI·ADP}}\right)\delta g^{\text{ATP,ADP}}_{\text{A,I}} + n^{\text{CI·KaiB}}\,\delta g^{\text{CI·KaiB}}_{\text{A,I}} \\ &\quad + n^{\text{CI·KaiA}}\,\delta g^{\text{CI·KaiA}}_{\text{A,I}} + n^{\text{CII.KaiA}}\,\delta g^{\text{CII·KaiA}}_{\text{A,I}}. \end{aligned} \tag{4.15}$$

Each contribution is directly proportional to the the number of ADP nucleotides, KaiB monomers and KaiA dimers, $n^{\text{CI·ADP}}$, $n^{\text{CI·KaiB}}$, and $n^{\text{CI·KaiA}}$, respectively, bound to the CI domain. The last contribution is proportional to the number of KaiA dimers, $n^{\text{CII·KaiA}}$, bound to the CII domain. The terms depending on the number of ADP and KaiB proteins bound are explained in the theory section. As discussed in more detail in the next section, KaiA can only bind to the CI domain when 6 KaiB monomers are bound to CI; moreover, 6 KaiA dimers can then be sequestered. The free-energy contribution for binding a single KaiA dimer to CI results from the detailed-balance relation, $\delta g^{\text{CI·KaiA}}_{\text{A,I}} = -\log\left(K^{\text{CI·KaiA}}_{\text{d,A}}/K^{\text{CI·KaiA}}_{\text{d,I}}\right)$, where $K^{\text{CI·KaiA}}_{\text{d,A}}$ and $K^{\text{CI·KaiA}}_{\text{d,I}}$ are the dissociation constants for the active and inactive state of KaiC, respectively. A similar relation holds for KaiA binding to CII.

Because KaiA and KaiB have a higher affinity for inactive KaiC than for active KaiC, their binding to the CI domain will stabilize the inactive state. This, together with CI-ATP stabilizing the active state and CI-ADP stabilizing the inactive one, creates a hysteresis loop, as illustrated in Fig. 4.2. Importantly, while KaiA and KaiB binding to CI stabilizes the inactive state with respect to the active one, the system is designed such that when no ADP is bound to CI, the active conformation with KaiA and KaiB bound is *more* stable than the inactive one with KaiA and KaiB bound to it. This is critical, because it ensures that when, during the dephosphorylation phase, the system eventually reaches the point where all ADP has been released and the number of CI-bound ADP molecules has reached zero, the hexamer flips back from the inactive to the active state. In this active state, KaiA and KaiB will then spontaneously dissociate from KaiC, because this conformational state has a low affinity for KaiA and KaiB. Given Eq. 4.15, the requirement that the active state with KaiA and KaiB bound is always more stable than the inactive one when no ADP is bound to CI, entails that $n_0^{\text{CI·ADP}}\,\delta g^{\text{ATP,ADP}}_{\text{A,I}} \leq n^{\text{CI·KaiA}}_{\text{max}}\,\delta g^{\text{CI·KaiA}}_{\text{A,I}} + n^{\text{CI·KaiB}}_{\text{max}}\,\delta g^{\text{CI·KaiB}}_{\text{A,I}}$.

Given $\Delta G^{\text{hex}}_{\text{A,I}}$, the rates of switching from the active to the inactive state, $k^{\text{conf}}_{\text{f}}$, and vice versa, $k^{\text{conf}}_{\text{b}}$, become $k^{\text{conf}}_{\text{f}} = k_0^{\text{conf}}\exp\left(-\Delta G^{\text{hex}}_{\text{A,I}}/2\right)$ and $k^{\text{conf}}_{\text{b}} = k_0^{\text{conf}}\exp\left(\Delta G^{\text{hex}}_{\text{A,I}}/2\right)$, respectively. The prefactor k_0^{conf} sets the timescale of switching.

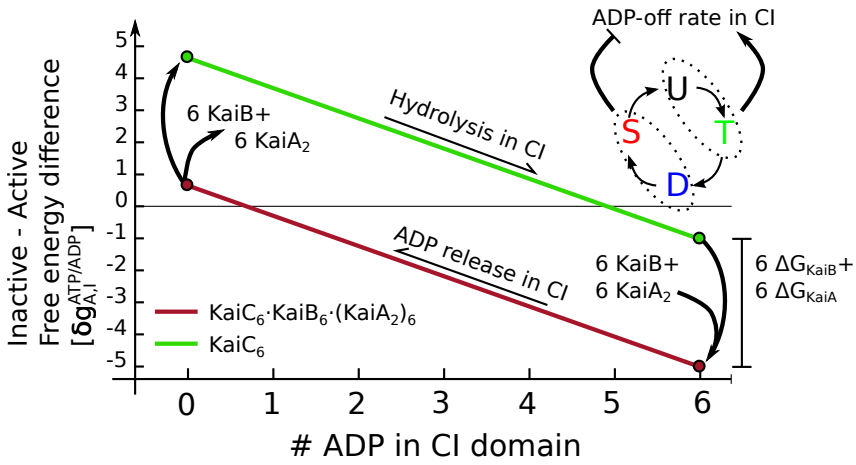

Figure 4.2: Hydrolysis of ATP in the CI domain, and the subsequent binding op KaiB and KaiA, forms a hysteresis loop in the free-energy difference between the active and inactive state. Starting in the active state with no ADP bound to CI, the free-energy difference linearly decreases as ATP is hydrolyzed to ADP, as indicated by the green line. The number of ADP in CI is set by the competition between a fixed hydrolysis rate and a variable ADP off-rate, which is set by the phosphorylation state of the CII domains of the whole hexamer. Phosphorylation of the T site initially enhances the ADP off-rate, but the subsequent phosphorylation of the S-site will decrease the dissociation of ADP. The antagonistic effect of the cyclically phosphorylated T and S sites on the ADP dissociation rate causes a sharp transition in the CI binding pocket from ATP dominated to ADP dominated [18]. When there are 5 or 6 ADP nucleotides in the CI domain, the hexamer will flip to the inactive state, increasing the affinity for KaiB, which will then slowly bind inactive KaiC. When 6 KaiB monomers are bound to the hexamer, up to six free KaiA dimers will be sequestered from solution. This complex of KaiB and KaiA on the CI domain stabilizes the inactive state of KaiC further (red line). When all KaiA is sequestered, the dephosphorylation of the S-site in the CII domain will increase the nucleotide exchange rate in the CI domain, decreasing the ADP level and increasing the free-energy difference between the active and inactive state. This allows the hexamer to flip back to the active state. In the active state, KaiA and KaiB dissociate from KaiC, because they have a low affinity for CI in the active conformation.

4.4.5. SLOW KAIB BINDING SETS A TIME DELAY BETWEEN THE PHOSPHORYLATION AND DEPHOSPHORYLATION PHASE

Experiments show the binding of KaiB to the KaiC hexamer is slow due to a combination of a slow conformational switch of KaiB monomers in solution, and a slow reaction step attributed to the CI domain of KaiC [26, 27]. Theoretical modeling suggests that the Kai oscillator requires a time delay between the phase of phosphorylation and the phase of KaiA sequestration and dephosphorylation, in order to generate stable oscillations [15, 18, 19] and a period that is fairly insensitive to changes in the bulk ATP fraction [26]. We introduce this delay by assuming KaiC can not sequester KaiA before a full ring of 6 KaiB monomers has formed on the CI domain [126, 137], and that the rate of KaiB binding to KaiC is slow and independent of the number of KaiB proteins already bound.

In this way we simulate the slow appearance of KaiB monomers in the bulk that have a binding competent conformation [27]. As explained in the model overview, section 4.2, we do not explicitly keep track of KaiB, but coarse grain the KaiB concentration in the association rates motivated by the observation that the absolute concentration of KaiB has little influence on the amplitude and period of the oscillation [15, 130].

After a KaiB ring has formed, KaiA will immediately be sequestered from solution due to a very high on-rate, with a maximum of 6 KaiA dimers per hexamer [52, 53]. The affinity of KaiA for the hexamer with a KaiB ring depends on the conformational state of the hexamer: Only in the inactive state KaiA stays bound to KaiC-KaiB. In the active state of KaiC, the CI domain has a lower affinity for both KaiA and KaiB, as discussed in the previous section. Hence, after KaiC has flipped to the active state, KaiA and KaiB will be released, and the cycle starts over.

4.4.6. SUMMARY OF THE CYCLE DYNAMICS

Hydrolysis of ATP in the CI domain drives the conformational transition from the active state to the inactive one, because ADP in the CI domain stabilizes the inactive state with respect to the active one, see Fig. 4.2. This allows the ensemble of KaiC hexamers to switch from a phosphorylation phase with free KaiA in solution to a dephosphorylation phase where all KaiA is sequestered by KaiC. In the inactive state, the CI domain of KaiC has a high affinity for both KaiA and KaiB, meaning that the complex $\mathrm{KaiC}_6 \cdot \mathrm{KaiB}_6 \cdot (\mathrm{KaiA}_2)_6$, is energetically very stable.

While all KaiA is sequestered, the KaiC ensemble will dephosphorylate leaving most of the monomers in the U or S state. When the number of S phosphorylated sites drops below a threshold, the energy barrier for ADP release from CI will become sufficiently small, such that ADP will dissociate even though ADP is stabilized by the binding of KaiA and KaiB. Without ADP bound to the CI domain the hexamer returns to the active state, which has a lower affinity for KaiA and KaiB. The sequestered proteins are then immediately released such that the cycle can start over again with an ensemble of hexamers with monomers in the U state.

4.5. RESULTS ON PHOSPHORYLATION DYNAMICS

To test the validity of our model of the CII domain and to find the correct parameter values, we compare with the rich body of quantitative experimental results on the in-vitro Kai system. As is done in these experiments, we will study the behavior of different combinations of the main actors: KaiA, KaiB and KaiC and the ATP fraction, α_{ATP}. First, we will study the dephosphorylation dynamics of phosphorylated KaiC in the absence of KaiA and KaiB. We will investigate the dependence on the ATP fraction and compare to experimental results. Furthermore, we test our hypothesis that phosphotransfer from the threonine and serine sites to ADP is the only possible pathway for dephosphorylation, by comparing to a model where release of the phosphate into the bulk is possible. Next, we study the effects of KaiA on the phosphorylation dynamics and the influence of the bulk ATP fraction on the speed and steady state level of phosphorylation. Lastly, we study if the ordered phosphorylation dynamics of the serine (S) and threonine (T) residues persists when the system of KaiA and KaiC has reached steady state. In the

subsequent section, we address the role of KaiB and the oscillatory dynamics. All simulations were performed at the experimental standard concentrations of 0.6μM for KaiA and KaiC, which corresponds to simulating 720 KaiC hexamers and 720 KaiA dimers in a volume of 2 cubic micron.

4.5.1. KaiC DEPHOSPHORYLATION SPEED IS SET BY PHOSPHOTRANSFER RATES

Dephosphorylation via phosphotransfer can reproduce experiments We start with simulating an ensemble of KaiC hexamers in a 100% ATP solution, and compare with the experimental results presented in the supplementary information of [26]. Initially, KaiC is highly phosphorylated. As there is no KaiA (and also no KaiB), this experiment allows us to distinguish the rate constants related to phosphotransfer dynamics given in Eq. 4.1, from the effects related to the interaction with KaiA. To obtain the rapid decay of the T and D phosphorylated states and the transient peak in the S state, as shown in experiments, the phosphotransfer rates relating to the threonine site have to be significantly faster than the rates relating to the serine site. Furthermore, dephosphorylation is downhill in free energy. Otherwise, the ATP regenerated in the dephosphorylation reaction would phosphorylate KaiC again. The free-energy landscape of the phosphorylation states combined with the nucleotide binding pockets, where KaiA is not bound, is drawn in Fig. 4.2, left panel. We choose the magnitude of the rates such that we can reproduce the time and height of the maximum in the concentration of S-phosphorylated KaiC as well as its subsequent decay. Since dephosphorylation can only occur after the ATP in the binding pocket has been hydrolyzed, the hydrolysis rate constant sets an upper bound on the speed, which we set to $k_{\mathrm{hyd}}^{\mathrm{CII}} = 1\,\mathrm{h}^{-1}$, similar to what has been found in [116]. Fig. 4.1A shows that our model, with the phosphotranfer parameters of Table 4.1, reproduces the dephosphorylation dynamics of Fig. S2 of [26]. Next, we test if in our model the dephosphorylation speed is independent of the bulk ATP fraction, $\alpha_{\mathrm{ATP}} = [\mathrm{ATP}]/([\mathrm{ATP}] + [\mathrm{ADP}])$, as was found in experiments [7]. Eq. 4.1C shows that this is indeed the case. This is because the ATP hydrolysis rate is higher than the ATP dissociation rate and the phosphotransfer rate is faster than the ADP dissociation rate. This ensures that during dephosphorylation the nucleotides are not released—if this would happen, the subsequent nucleotide binding and hence the dephosphorylation rate would depend on α_{ATP}. When the above requirements are fulfilled, KaiC predominantly dephosphorylates via $\mathrm{D{\cdot}ATP} \xrightarrow{k_{\mathrm{hyd}}^{\mathrm{CII}}} \mathrm{D{\cdot}ADP} \xrightarrow{k_{\mathrm{DS}}^{0}} \mathrm{S{\cdot}ATP} \xrightarrow{k_{\mathrm{hyd}}^{\mathrm{CII}}} \mathrm{S{\cdot}ADP} \xrightarrow{k_{\mathrm{SU}}^{0}} \mathrm{U{\cdot}ATP}$.

Dephosphorylation does not occur via phosphate release. Importantly, if dephosphorylation were to occur through the direct release of inorganic phosphate groups into the bulk, then the dephosphorylation speed would *also* trivially be independent of the bulk ATP fraction. We therefore set out to test the hypothesis that all phosphorylation and dephosphorylation occurs only through phosphotransfer [116, 117]. To this end, we compare our model with an alternative one in which also direct exchange of the phosphate group with the bulk is possible. Specifically, we simulated the experiments performed in [116], in which they track radioactively labeled phosphate groups starting bound to the serine and threonine sites of KaiC, in a solution with only non-radioactive ATP. Here, KaiC dephosphorylates while producing a transient population of radioac-

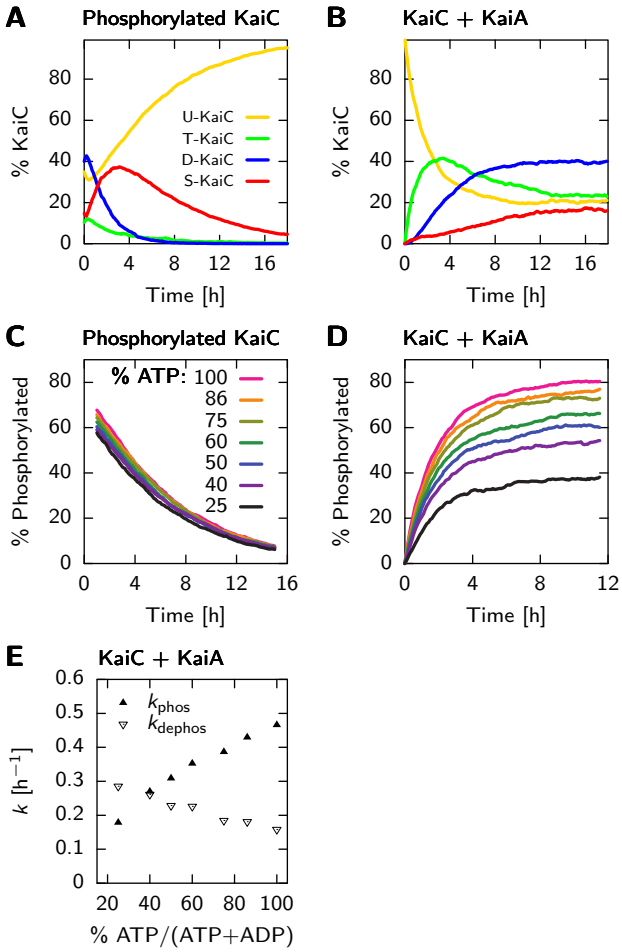

Figure 4.1: The free-energy landscape in Fig. 4.2, accurately describes the phosphorylation and dephosphorylation dynamics in the CII domain, both with and without KaiA and for a wide range of different ATP to ADP fractions in the solution, α_{ATP}. (A) Dephosphorylation dynamics, starting with phosphorylated KaiC and (B) phosphorylation dynamics, starting with unphosphorylated KaiC and KaiA. (C) Dephosphorylation and (D) phosphorylation of KaiC under similar conditions as in panels A and B, respectively, but now for different bulk fractions of ATP, α_{ATP}. Consistent with experiments, the rate of dephosphorylation is independent of α_{ATP}, but the phosphorylation-rate does depend strongly on this fraction. (E) Rates of (de)phosphorylation in panel D, found by fitting Eq. 4.16 to the first four hours of phosphorylation data.

tive ATP*, where at the maximum around 20% of the radioactive phosphates are bound to a nucleotide. The inorganic phosphate groups, Pi*, only appear in the bulk after a marked delay. Fig. 4.2A shows that our model, where dephosphorylation can only occur via phosphotransfer of the phosphate to the ADP and the subsequent hydrolysis of the

Parameters relating to the CII domain		
Parameter	Value	Explanation
Phosphotransfer		
k^0_{UT}	0.50 h^{-1}	U·ATP→T·ADP
k^0_{TU}	1.78 h^{-1}	T·ADP→U·ATP
k^0_{TD}	0.40 h^{-1}	T·ATP→D·ADP
k^0_{DT}	0.20 h^{-1}	D·ADP→T·ATP
k^0_{SD}	1.50 h^{-1}	S·ATP→D·ADP
k^0_{DS}	2.00 h^{-1}	D·ADP→S·ATP
k^0_{US}	0.15 h^{-1}	U·ATP→S·ADP
k^0_{SU}	0.20 h^{-1}	S·ADP→U·ATP
Nucleotide binding pocket		
k^{CII}_{hyd}	1.00 h^{-1}	ATP hydrolysis rate
$k^{CII\cdot ADP}_{off,KaiA}$	6.00 h^{-1}	ADP off-rate with KaiA bound
$K^{CII}_{ATP/ADP}$	0.10	Relative affinity for ATP and ADP
KaiA affinity		
$k^{CII\cdot KaiA}_{on,0}$	1.00 mM h^{-1}	KaiA on-rate for CII domain
$k^{CII\cdot KaiA}_{off,0}$	1.00 h^{-1}	KaiA off-rate for CII domain
$\delta g^{CII\cdot KaiA}_{bind}(U)$	0.00 kT	Free energy effect of U-monomer
$\delta g^{CII\cdot KaiA}_{bind}(T)$	-0.30 kT	Free energy effect of T-monomer
$\delta g^{CII\cdot KaiA}_{bind}(D)$	1.00 kT	Free energy effect of D-monomer
$\delta g^{CII\cdot KaiA}_{bind}(S)$	2.00 kT	Free energy effect of S-monomer
$\Delta G^{CII\cdot KaiA}_{A,I}$	10 kT	Free energy effect of conformation

Table 4.1: Model parameters relating to the CII domain are introduced in the theory section on phosphorylation dynamics and their values are motivated in the results section. Energies are given in units of kT, were k is Boltzmann's constant and T the temperature.

ATP, is in good quantitative agreement with the results of Fig. 2 in [116]: Both the magnitude and timing of the peak in the radioactive ATP, ATP*, and the delay in the appearance of radioactive phosphate groups in the bulk, Pi*, are in good agreement.

To test the effects of direct exchange of phosphate with the bulk, we add the reaction, $X\cdot N \rightleftharpoons Y\cdot N + Pi$, to our model. Here X and Y are connected phosphorylation states and N is the state of the nucleotide binding pocket, which can be either ATP or ADP, but now does not change state during dephosphorylation. We changed the rate constants such that dephosphorylation occurs equally through both pathways and the overall speed is comparable to the original model. Fig. 4.2B shows that in this scenario, the time traces are qualitatively different from the experiment. Because the phosphate groups are directly exchanged with the bulk, the delay in [Pi*] has disappeared and the magnitude of the peak in radioactive ATP is less than 10%. The discrepancy between the experimental data and our simulations thus confirms our hypothesis that the direct

exchange with the bulk of phosphates is negligible.

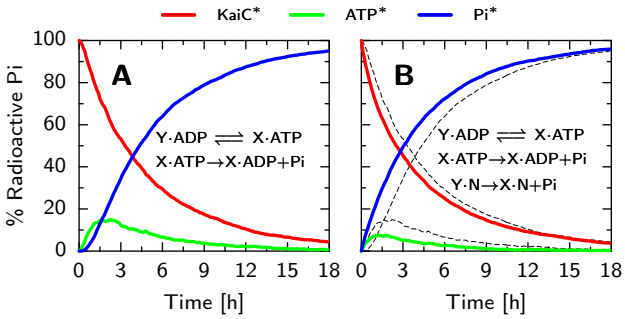

Figure 4.2: Tracking radioactive phosphate groups shows that phosphotransfer between KaiC and ADP is the dominant pathway during dephosphorylation of KaiC. Initial conditions are the same as in Fig. 4.1A, but now the phosphate groups are 'radioactive' such that they can be tracked from their initial position on the KaiC (KaiC*, red line), to the radioactive ATP (ATP*, green line) and in solution (Pi*, blue line). Panels A show time traces of this dephosphorylation assay for our model, while panel B shows the results for an alternative model in which spontaneous dephosphorylation is also possible (for comparison, the results of our model, presented in panel A, are shown as thin dashed lines). In our model, panel A, dephosphorylation can only occur via the transfer of the phosphate group on KaiC to the ADP in the CII nucleotide binding pocket. The alternative model, panel B, represents a scenario with two dephosphorylation pathways: one via the transfer to ADP, and one via direct exchange with the bulk. We chose rates such that dephosphorylation occurs equally through both pathways, and the total dephosphorylation speed is similar to our model (panel A). Comparing with the radioactive phosphate tracking experiment in [116], Fig 2, shows that the scenario where dephosphorylation can only occur via ADP (panel A) agrees best with the data. In particular, the onset of the free Pi concentration shows a temporal delay in panel A that is absent in panel B. Furthermore, the level of radioactive ATP is the highest in panel A, which compares best to the 20% peak height in [116].

4.5.2. KAIA SETS PHOSPHORYLATION DYNAMICS

To address the effect of KaiA binding to the CII domain on the phosphorylation free-energy landscape shown in Eq. 4.4, we compared with the experimental time traces in Fig. S2 of [26]. In these experiments, they start with unphosphorylated KaiC together with KaiA and track the fractions of T,D and S phosphorylated KaiC. This allows us to constrain the change in free energy of the phosphorylation states due to KaiA binding, $\delta g_{\mathrm{bind}}^{\mathrm{CII\cdot KaiA}}(\mathrm{X})$, introduced in Eq. 4.4. The affinity of KaiA for unphosphorylated KaiC is about 1nM, as reported in [53]. To make the large overshoot of T-phosphorylated KaiC possible, KaiA should lower the free-energy of the T state, $\delta g_{\mathrm{bind}}^{\mathrm{CII\cdot KaiA}}(\mathrm{T}) < 0$, while at the same time blocking the transition to the S-phosphorylation state, $\delta g_{\mathrm{bind}}^{\mathrm{CII\cdot KaiA}}(\mathrm{S}) > 0$. Experiments starting with radioactively labeled nucleotides in the binding pockets of KaiC in solution with KaiA, show that ADP bound to the CII domain has a very high off-rate [127]. Therefore, the ADP dissociation rate with KaiA bound to the CII domain has a high value of $k_{\mathrm{off,KaiA}}^{\mathrm{CII\cdot ADP}} = 6.0\mathrm{h}^{-1}$. Taken together, we find that with the parameters for KaiA binding presented in Table 4.1, we can reproduce the phosphorylation dynamics as

shown in Fig. 4.1B.

Model can reproduce dependence phosphorylation on ATP fraction. Next we checked the effect of the ATP fraction, α_{ATP}, on the speed and steady state level of the total phosphorylation fraction and compare with experiments in [7]. As the sensitivity of phosphorylation to the bulk ATP fraction is set by the relative affinity for ATP and ADP, $K^{\mathrm{CII}}_{\mathrm{ATP/ADP}}$, we can use the data in [7] to constrain this parameter. Using $K^{\mathrm{CII}}_{\mathrm{ATP/ADP}} = 0.10$, Fig. 4.1D shows that the phosphorylation time traces at different values of α_{ATP}, are in good agreement with experiments: Changing α_{ATP} from 100% to 25%, the steady state phosphorylation level drops from 80% to 40%.

To quantify the change in phosphorylation speed by varying α_{ATP}, we fit the first 4 hours of the phosphorylation time traces in Fig. 4.1D, $p(t)$, with a 2 state model,

$$p(t) = k_{\mathrm{phos}}/(k_{\mathrm{phos}} + k_{\mathrm{dephos}})\,(1 - \exp(-(k_{\mathrm{phos}} + k_{\mathrm{dephos}})\,t). \tag{4.16}$$

Here, the system switches between the phosphorylated and dephosphorylated state with the rates k_{phos} and k_{dephos}. Fig. 4.1E shows that k_{phos} linearly increases with α_{ATP}, as was found in [7, 26], but with a slope that is less steep. Furthermore, fitting the two state model to our modeling data yields a k_{dephos} that decreases with α_{ATP}, while the fitting to the experimental data yields a k_{dephos} that is virtually independent of α_{ATP}. This decrease in our model is due to the fact that the effective dephosphorylation rate is proportional to the fraction of ADP in the CII binding pocket. We attribute the inconsistency, at least in part, to the lower number of data points in the experimental time traces that are available for fitting to the two state model.

4.5.3. ORDERED PHOSPHORYLATION OF THE S AND T SITES PERSISTS IN STEADY STATE

The phosphotransfer reactions and the binding of KaiA all fulfill detailed balance, and only the irreversible hydrolysis reaction in the CI and CII domains do not. This raises the question whether, when the solution contains only KaiC hexamers and KaiA, individual KaiC monomers continue to go through the ordered cycle U → T → D → S → U. In this case there will be no macroscopic oscillations in the phosphorylation fraction because KaiA is never sequestered by KaiB and the phosphorylation cycles of the individual KaiC hexamers are not synchronized. The concentrations of the U,T,D and S phosphorylated monomers will therefore be in steady state. If we find ordered phosphorylation of the threonine and serine sites, this has to be driven by the hydrolysis of ATP in the CI and/or the CII domain. We want to know whether the phosphorylation cycle is mainly driven by hydrolysis in the CI or the CII domain. To this end, below we consider scenarios where we remove the hydrolysis in the respective domains and study its effect on the phosphorylation dynamics.

To find out if the ordered phosphorylation cycle persist in a system with only KaiA and KaiC, we need to know if there are net fluxes between states in the phosphorylation state space, indicating that detailed balance is broken [138]. To this end, we keep track of the number of phosphorylated threonine, n_{T}(t), and serine residues, n_{S}(t), in each individual hexamer in the ensemble. From this data we can calculate the probability, $P_{n_{\mathrm{T}},n_{\mathrm{S}}}$, that a hexamer is in phosphorylation state $(n_{\mathrm{T}}, n_{\mathrm{S}})$, and the number of times the

hexamer switches from this state to one of its neighboring states, $N^{\mathrm{x}}_{\alpha,\beta}$. Here $x \in \{T, S\}$ indicates whether the (de)phosphorylation event involved a threonine or serine residue, α gives the number of phosphorylated sites of this type before the transition, and β the number after the transition. The net flux between two neighboring states is given by

$$W^{\mathrm{x}}_{\alpha,\beta} = \frac{N^{\mathrm{x}}_{\alpha,\beta} - N^{\mathrm{x}}_{\beta,\alpha}}{\Delta t_{\mathrm{sim}}}, \tag{4.17}$$

where Δt_{sim} is the time interval over which these time traces were measured. As was done in [139], we can define a vector, $\vec{J}_{n_{\mathrm{T}},n_{\mathrm{S}}}$, that points in the direction of the mean net flux through the state $(n_{\mathrm{T}},n_{\mathrm{S}})$

$$\vec{J}_{n_{\mathrm{T}},n_{\mathrm{S}}} = \frac{1}{2}\begin{pmatrix} W^{\mathrm{T}}_{\alpha^-,\alpha} + W^{\mathrm{T}}_{\alpha,\alpha^+} \\ W^{\mathrm{S}}_{\alpha^-,\alpha} + W^{\mathrm{S}}_{\alpha,\alpha^+} \end{pmatrix}. \tag{4.18}$$

Here the pair α^-,α indicates the net flux along the threonine or serine axis from below the coordinate α, and the pair α,α^+ indicates the net flux to above α.

Hydrolysis of ATP in the CII domain provides the ADP required for dephosphorylation. However, when there is no hydrolysis in the CII domain and the bulk only contains ATP, ATP is never converted to ADP in the CII domain (See Fig. 4.1), such that dephosphorylation becomes impossible. All the monomers will be permanently in the D state with an ATP in the binding pocket, blocking the possibility of a phosphorylation cycle. Therefore, in all scenarios discussed in this section, we will use a lower ATP fraction of $\alpha_{\mathrm{ATP}} = 0.5$ such that there is ADP from the bulk available for dephosphorylation.

In Fig. 4.3, we show a heat map of $P_{n_{\mathrm{T}},n_{\mathrm{S}}}$ together with the vectors $\vec{J}_{n_{\mathrm{T}},n_{\mathrm{S}}}$, for a system with KaiA and KaiC, in the presence of ATP hydrolysis in both the CI and CII domains. In panel A we show the behavior of KaiC in solution with KaiA, which clearly shows cyclic net fluxes in phosphorylation state space. Starting in the lower left corner, where the hexamer is unphosphorylated, first the threonine sites will phosphorylate after which the serine sites become phosphorylated. After reaching the upper right corner of the state space, the hexamer will first dephosphorylate the threonine sites and then the serine sites. As explained in the theory section on phosphorylation dynamics, the combination of hydrolysis in the CII domain and differential affinity for KaiA, results in a high ATP fraction in the CII binding pockets when the hexamer is predominantly in the U and T state, and a low ATP fraction when it is in the D and S state. Therefore, KaiC, on average, phosphorylates when it is in the U and T state and dephosphorylates when it is in the D and S state, which is the origin of the cycle in $n_{\mathrm{T}} - n_{\mathrm{S}}$ space.

Cycle is driven by ATP hydrolysis and differential affinity. We then asked what drives the phosphorylation cycle at the level of the individual hexamers. We first set the hydrolysis rates in both domains to zero, thereby removing all irreversible pathways in the model. Clearly, without ATP hydrolysis, as shown in Fig. 4.3B, the fluxes disappear completely. This shows that our model obeys detailed balance when we remove the two irreversible pathways. When we remove hydrolysis in CI, $k^{\mathrm{CI}}_{\mathrm{hyd}} = 0$, panel C, ordered phosphorylation persists, showing that hydrolysis in the CI domain is not essential for generating a cycle at the level of the individual hexamers. If we only remove hydrolysis

in CII and not in CI, there is still a small but clear cycle in state space due to the orchestrated switching between the active and inactive state caused by the phosphorylation state, and its effect on the affinity of the CII domain for KaiA (data not shown). Clearly, ATP hydrolysis in CI or CII is necessary to generate a cycle. Yet, it is not sufficient: If we remove differential affinity of KaiA binding to the CII domain but keep ATP hydrolysis in CI and CII, $\delta g_{\text{bind}}^{\text{CII}\cdot\text{KaiA}}(X_i) = 0$ and $\Delta G_{\text{A,I}}^{\text{CII}\cdot\text{KaiA}} = 0$, all fluxes disappear in phosphorylation state space, as shown in Fig. 4.3D. Hence, both differential affinity and ATP hydrolysis, most notably in the CII domain, are necessary to generate a phosphorylation cycle at the level of the individual hexamers.

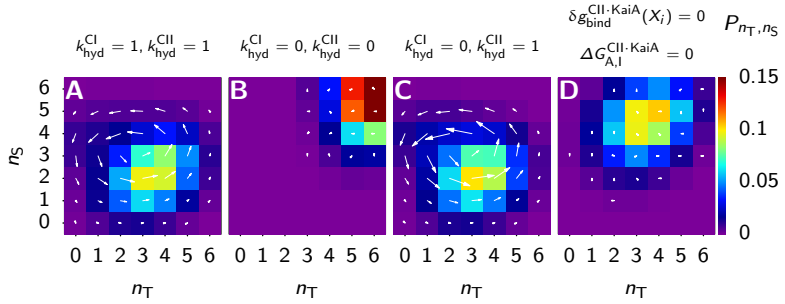

Figure 4.3: Hydrolysis in the CII domain, without KaiA sequestration by the CI domain, is sufficient to generate the ordered phosphorylation of the T and S sites in a solution of KaiC and KaiA. In all panels, the bulk ATP fraction is 50%. (A-D) Heat maps of the probability, P_{n_T,n_S}, for a single hexamer, of having n_T phosphorylated threonine sites and n_S phosphorylated serine sites. Arrows indicate the net flux through a state, where the length is proportional to the magnitude of the flux. (A) KaiC and KaiA, with hydrolysis of ATP in both the CI and CII domains, shows a clear ordered cycle in state space. (B) Without hydrolysis in both the CI and CII domains, all the equations in our model obey detailed balance, and the fluxes in state space disappear. (C) When we remove hydrolysis in the CI domain, the fluxes in phosphorylation state space are little affected compared to panel A. (D) When there is hydrolysis in both domains, but no differential affinity of the CII domain for KaiA, the hydrolysis cycle in the CII domain does not couple to the phosphorylation cycle, and there is no ordered phosphorylation.

4.6. RESULTS ON KAIC ATPASE AND CYCLE DYNAMICS

To test our model of the CI domain and to find the correct parameter values, we compare with the experiments on the evolution of the ATP fraction in the binding pockets of KaiC and the ATPase rate of KaiC [114–116, 127]. As is done in these experiments, we will study the behavior of different combinations of the main actors: KaiA, KaiB and KaiC and the ATP fraction α_{ATP}. First we will study a system containing only KaiC, which allows us to constrain the parameters setting the rate of the hydrolysis cycle in the CI domain. Next, we look at the effect of dephosphorylation on the transient ATP fractions in the CI and CII domains, which provides an informative and testable prediction for how the CII domain regulates the ATP fraction in CI, which is indeed one of the key characteris-

tics of our model. Then we study the effect of KaiA on the steady state ATPase rate and the dynamics of the ATP fraction in phosphorylating KaiC. For the full oscillating system, we will study the phase difference between the phosphorylation level and ATPase rate. Then we study whether the oscillations are robust to changes in the ATP fraction in solution, and whether our model can reproduce the experimental observation that the period of the oscillation is insensitive to these changes. Again, all simulations were performed with 720 KaiC hexamers and 720 KaiA dimers corresponding to the experimental standard condition of a 0.6μM concentration in a volume of 2 cubic micron.

4.6.1. ENSEMBLE OF DEPHOSPHORYLATING KAIC SHOWS A TRANSIENT DECREASE IN ATPASE RATE

To constrain the hydrolysis rate constant and the ADP dissociation rate in the CI domain of unphosphorylated KaiC, we used the experimental observation that in a system with only KaiC that has reached steady state, the ATP fraction in the binding pockets is around 30% [127] and the ATPase rate is 0.6 ATP per KaiC monomer per hour [114]. As we argued in the theory section on KaiA acting as a nucleotide exchange factor, without KaiA, the CII binding pockets are predominantly occupied by ADP, the monomers are in the U state, and the ATPase activity comes mainly from hydrolysis in the CI domain. Now, since the total fraction of ATP in the binding pockets is given by $0.5(\beta^{\mathrm{CI}}_{\mathrm{ATP}}+\beta^{\mathrm{CII}}_{\mathrm{ATP}})=0.3$, where $\beta^{\mathrm{CI}}_{\mathrm{ATP}}$ and $\beta^{\mathrm{CII}}_{\mathrm{ATP}}$ are the ATP fractions in the CI and CII domain, respectively, we estimate that the fraction of ATP in the CI domain, $\beta^{\mathrm{CI}}_{\mathrm{ATP}}=0.6$, because $\beta^{\mathrm{CII}}_{\mathrm{ATP}}\approx 0$. Assuming the measured ATPase rate equals the hydrolysis rate constant times the fraction of ATP in the CI binding pocket, $k^{\mathrm{CI}}_{\mathrm{hyd}}\,\beta^{\mathrm{CI}}_{\mathrm{ATP}}$, we estimate that $k^{\mathrm{CI}}_{\mathrm{hyd}}=1.0\,\mathrm{h}^{-1}$. Since $\beta^{\mathrm{CI}}_{\mathrm{ATP}} = k^{\mathrm{CI{\cdot}ADP}}_{\mathrm{off}}/(k^{\mathrm{CI{\cdot}ADP}}_{\mathrm{off}} + k^{\mathrm{CI}}_{\mathrm{hyd}})$, we deduce that $k^{\mathrm{CI{\cdot}ADP}}_{\mathrm{off}}=1.5\,\mathrm{h}^{-1}$.

To find out if an ensemble of only KaiC hexamers has the observed dynamics of the ATP fraction, we first consider a system in which KaiC is unphosphorylated and the ATP fraction in the binding pockets is 100%. We then study its relaxation to steady state. Fig. 4.1A shows an exponential decay of the ATP fraction in both domains, on a similar timescale and steady state value as was found in [116]. The mean ATP fraction and ATPase rate, given in Table 4.1, are in quantitative agreement with experimental data presented above.

Next we want to find out how the ATP fraction evolves in a system in which KaiC is initially highly phosphorylated, with the concentrations of monomers in the U,T,D and S state evolving as shown in Fig. 4.1A. Please note that in our model, the T state will lower the activation energy for ADP dissociation set by $\delta g^{\mathrm{CI{\cdot}ADP}}_{\mathrm{barrier,A/I}}(X_i)$, while the D and S state will increase the activation energy.

The parameter values for the contributions to the activation energy, $\delta g^{\mathrm{CI{\cdot}ADP}}_{\mathrm{barrier,A/I}}(X_i)$, given in Table 4.2, were chosen to give good agreement of the oscillation dynamics presented below. We predict that the peak in the number of monomers in the S state (Fig. 4.1A), which will decrease the ADP dissociation rate, causes a transient lowering of the overall bound ATP fraction, particularly in the CI domain. Indeed, Fig. 4.1B shows a clear trough in the fraction of ATP in the binding pockets, which is most pronounced for CI. An experiment tracking the ATP fraction in the binding pockets, as performed in [127], but now starting with phosphorylated monomers, would be able to verify this

Mixture			⟨ATPase⟩ [#ADP/KaiC/day]			
KaiA	KaiB	Condition	CI	CII	CI+CII	From [114]
-	-		14.4	0.0	14.4	14.5
+	-		11.5	17.0	28.5	18.1
-	+		14.2	0.0	14.2	8.9
+	+		12.3	8.8	21.1	15.8
+	+	$\alpha_{\mathrm{ATP}} = 50\%$	13.0	8.3	21.3	-
+	-	$k_{\mathrm{hyd}}^{\mathrm{CII\cdot KaiA}} = 0$	10.8	4.2	15.0	-
+	+	$k_{\mathrm{hyd}}^{\mathrm{CII\cdot KaiA}} = 0$	13.9	3.0	16.9	-
+	-	$K_{\mathrm{ATP/ADP}}^{\mathrm{CII}}(\mathrm{D})=10$	11.2	11.4	22.6	-
+	+	$K_{\mathrm{ATP/ADP}}^{\mathrm{CII}}(\mathrm{D})=10$	12.8	7.4	20.2	-

Table 4.1: Measured ATPase activity in ADP molecules produced per KaiC monomer per day (24 hours), under different conditions. First rows show results for parameters given in Tables 4.1 and 4.2. The last four rows show results for alternative models. When $k_{\mathrm{hyd}}^{\mathrm{CII\cdot KaiA}} = 0$, the ATP hydrolysis in the CII domain is blocked when KaiA is bound to CII. When, $K_{\mathrm{ATP/ADP}}^{\mathrm{CII}} = 10$, the CII domain has a higher relative affinity for ADP when the monomer is in the D state. The experimental values for the combined ATPase activity of the CI and CII domain, given in the last column, are taken from [114], and are shown for comparison.

prediction. Lastly, in our model the ATPase rate is proportional to the fraction of ATP in the binding pockets. Our model thus predicts a transient dip in the ATPase rate for dephosphorylating KaiC hexamers.

4.6.2. KAIA HAS OPPOSITE EFFECTS ON THE ATP FRACTION IN THE CI AND CII DOMAINS

Adding KaiA to an ensemble of KaiC will immediately increase the ATP fraction in the CII binding pockets, causing phosphorylation of KaiC as shown in Fig. 4.1B. Due to the phosphorylation of KaiC, the ATP fraction in the CI domain will decrease, because the ADP release in CI depends on the phosphorylation state of CII. These results are illustrated in Fig. 4.1C, which also shows that the total ATP fraction in CI and CII stabilizes around 60%, in good agreement with [127]. However, the steady state ATPase rate of the ensemble is around 29 ADP molecules produced per monomer per day (ADP/KaiC/day), Table 4.1, higher than the observed rate of 18 ADP/KaiC/day [114]. This descrepancy could be due to our assumption that the ATP hydrolysis rate is constant, independent of both the phosphorylation state and whether or not KaiA is bound, and/or the assumption that the nucleotide affinities in the CII domain are independent of the phosphorylation state. It is conceivable that KaiA, when bound to CII, decreases the ATP hydrolysis rate in the CII domain. When we set the hydrolysis rate to zero when KaiA is bound, $k_{\mathrm{hyd}}^{\mathrm{CII\cdot KaiA}}$=0, and adjust the ADP dissociation rate to keep the phosphorylation dynamics unchanged, $k_{\mathrm{off,KaiA}}^{\mathrm{CII\cdot ADP}}$=0.2, we find that the steady state ATPase rate in the system with only KaiA and KaiC drops to 15 ADP/KaiC/day (see Table 4.1), in good agreement with experiment. Another possibility is indeed that the affinity of doubly phosphorylated KaiC for

Parameters relating to the CI domain				
Parameter	Value			Explanation
Nucleotide binding pocket				
$k_{\text{hyd}}^{\text{CI}}$	1.00	h^{-1}		ATP hydrolysis in CI domain
$k_{\text{off}}^{\text{CI·ADP}}$	1.50	h^{-1}		ADP off-rate in CI domain
$\delta g_{\text{act,A}}^{\text{CI·ADP}}(U)$	0.00	kT		Activation energy contributions from the respective monomers, in the active state.
$\delta g_{\text{act,A}}^{\text{CI·ADP}}(T)$	-0.80	kT		
$\delta g_{\text{act,A}}^{\text{CI·ADP}}(D)$	0.40	kT		
$\delta g_{\text{act,A}}^{\text{CI·ADP}}(S)$	0.80	kT		
$\delta g_{\text{act,I}}^{\text{CI·ADP}}(U)$	-0.20	kT		Activation energy contributions from the respective monomer, in the inactive state.
$\delta g_{\text{act,I}}^{\text{CI·ADP}}(T)$	-0.80	kT		
$\delta g_{\text{act,I}}^{\text{CI·ADP}}(D)$	0.40	kT		
$\delta g_{\text{act,I}}^{\text{CI·ADP}}(S)$	0.80	kT		
KaiA and KaiB sequestration dynamics				
$k_{\text{on,A}}^{\text{CI·KaiA}}$	1.00	$\cdot 10^{6}$	μMh^{-1}	KaiA on-rate, active
$k_{\text{off,A}}^{\text{CI·KaiA}}$	1.00	$\cdot 10^{1}$	h^{-1}	KaiA off-rate, active
$k_{\text{on,I}}^{\text{CI·KaiA}}$	1.00	$\cdot 10^{6}$	μMh^{-1}	KaiA on-rate, inactive
$k_{\text{off,I}}^{\text{CI·KaiA}}$	1.00	$\cdot 10^{-1}$	h^{-1}	KaiA off-rate, inactive
$k_{\text{on,A}}^{\text{CI·KaiB}}$	1.00	$\cdot 10^{-1}$	h^{-1}	KaiB on-rate, active
$k_{\text{off,A}}^{\text{CI·KaiB}}$	1.00	$\cdot 10^{1}$	h^{-1}	KaiB off-rate, active
$k_{\text{on,I}}^{\text{CI·KaiB}}$	2.00	$\cdot 10^{0}$	h^{-1}	KaiB on-rate, inactive
$k_{\text{off,I}}^{\text{CI·KaiB}}$	1.00	$\cdot 10^{-2}$	h^{-1}	KaiB off-rate, inactive
$n_{\text{max}}^{\text{CI·KaiA}}$	6			#KaiA sequestered/hexamer
$n_{\text{max}}^{\text{CI·KaiB}}$	6			#KaiB sequestered/hexamer
Conformational state				
k_{0}^{conf}	10		h^{-1}	prefactor conformational switch
$n_{0}^{\text{CI·ADP}}$	5			Offset energy in conformation
$\delta g_{\text{A,I}}^{\text{ATP,ADP}}$	19		kT	A,I ADP dependent energy

Table 4.2: Model parameters relating to the CI domain are introduced in the theory section on the power cycle and their values are motivated in the results section. Energies are given in units of kT, were k is Boltzmann's constant and T the temperature. Note that the rates of binding and unbinding of KaiA are for the CI domain. For rates relating to the CII domain, see Table 4.1.

ATP is lower than assumed in our model. In our current model, doubly phosphorylated KaiC has the same high affinity for ATP as KaiC in the other phosphorylation states (see Fig. 4.2). When only KaiA is present and KaiC is often in the doubly phosphorylated D state, this sets up a futile cycle, in which KaiC continually binds ATP and then hydrolyzes it.

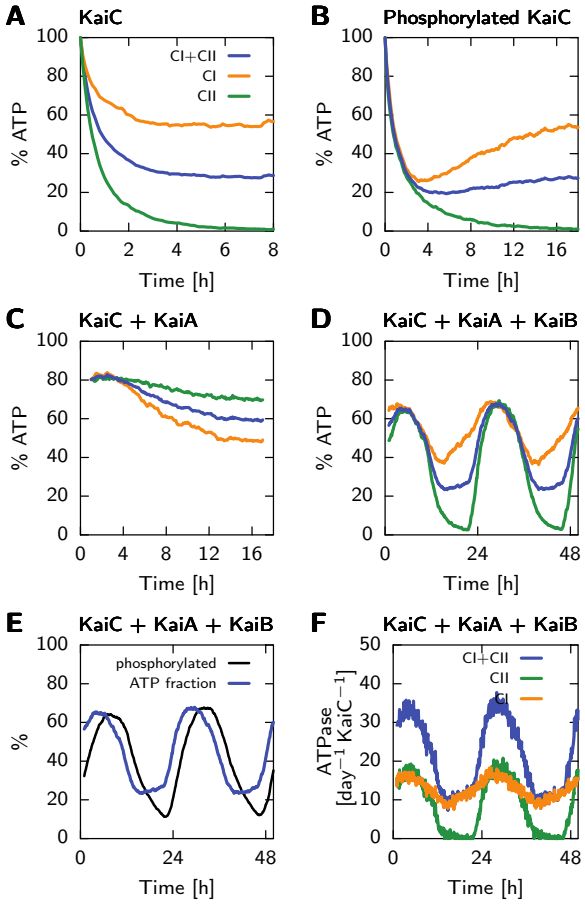

Figure 4.1: Model correctly predicts the ATP fractions in KaiC nucleotide binding pockets and the phase difference between this fraction and the phosphorylation level. Figures A-D show the fraction ATP/(ATP+ADP) in the nucleotide binding pockets of the CI domain (orange), the CII domain (green) and their sum (blue), in a 100% ATP solution, for different scenarios: KaiC initially unphosphorylated, no KaiA and KaiB present (A), KaiC initially phosphorylated, no KaiA and KaiB (B), initially unphosphorylated KaiC + KaiA (C) and KaiC + KaiA + KaiB (D). (A) The ATP levels drop monotonically due to the slow hydrolysis in both the CI and CII domain. (B) ATP fractions in dephosphorylating KaiC shows a clear trough in the ATP fraction of the CI domain due to the peak in the number of monomers in the S state (Fig. 4.1A), which temporarily decreases the ADP-off-rate in the CI domain. (C) For a system with KaiC and KaiA, the ATP fraction is higher in the CII domain and lower in the CI domain. KaiA increases the nucleotide exchange rate in the CII domain, and the resultant high phosphorylation level decreases the ADP dissociation rate in CI, which decreases the ATP fraction in this domain. (D) Full system with KaiA, KaiB and KaiC shows oscillations in the ATP fractions of both the CI and CII domains. (E) The phase difference and amplitude of the ATP fraction (% ATP, blue line) and the phosphorylation level (% phosphorylated monomers, black line) agree well with experimental results in [127]. (F) ATPase levels of the KaiC domains show oscillations proportional to the ATP fractions in the respective domains.

Decreasing the relative affinity of D for ATP versus ADP, such that $K^{\mathrm{CII}}_{\mathrm{ATP/ADP}} = 10$, lowers the steady state ATPase rate to about 22 ADP/KaiC/day. By making minor adjustments to parameters of the model, our model can thus reproduce the ATPase rate of KaiC in the presence of KaiA.

4.6.3. DELAYED KAIA SEQUESTRATION SYNCHRONIZES THE HEXAMERS

Given our analysis in section 4.5, which showed that KaiC goes through a ordered phosphorylation cycle in solution with KaiA, we wanted to know if the sequestration of KaiA, after the slow binding of KaiB, will synchronize the KaiC hexamers. To this end we chose rates for KaiB binding and unbinding as presented in Table 4.2, which correspond to a very low affinity for KaiC in the active state, where KaiB is almost never bound, and a high affinity for the inactive state where KaiB binds and unbinds slowly. KaiA binds rapidly to the inactive CI domain when 6 KaiB monomers are bound to it, and dissociates from this complex very slowly. The last important quantity relating to the CI domain, which determines the stabilization of the inactive state of KaiC by ADP in the CI domain, $\delta g^{\mathrm{ATP,ADP}}_{\mathrm{A,I}}$, is constrained from below by the relative affinities of KaiB and KaiA for the inactive state compared to the active state, as discussed in the theory section on the power cycle. Given the parameters for KaiA and KaiB binding to the CI domain in Table 4.2, we choose $\delta g^{\mathrm{ATP,ADP}}_{\mathrm{A,I}} = 19$ kT.

Figs. 4.1 D,E and F show clear oscillations in the ATP fraction in the CI and CII binding pockets, the phosphorylation fraction and the ATPase rates of the Kai oscillator, respectively. Panel E shows that the phase of the phosphorylation fraction is a few hours ahead of the ATP fraction in the binding pockets, which is in good agreement with experiments [127]. Also the amplitudes of both oscillations are in good agreement. The average ATPase rate of the oscillator is about 21 ADP/KaiC/day, which is slightly higher than the observed rate of 15 ADP/KaiC/day. We hypothesize that this high ATPase activity has to be attributed to the ATP hydrolysis in the CII domain. To test this, we set, as in the previous section, the hydrolysis rate constant to zero, $k^{\mathrm{CII\cdot KaiA}}_{\mathrm{hyd}} = 0.0$, when KaiA is bound to the CII domain. In order for the phosphorylation dynamics to be comparable to our original model, we set $k^{\mathrm{CII\cdot ADP}}_{\mathrm{off,KaiA}} = 0.2$. In this model, the average ATPase activity has dropped to 16 ADP/KaiC/day, almost equal to the experimentally observed value (Table 4.1). Decreasing the affinity of doubly phosphorylated KaiC for ATP, which strongly reduced the ATPase rate of KaiC in the presence of KaiA only (see previous section), has a much smaller effect when both KaiA and KaiB are present, lowering the ATPase rate to 20 ADP/KaiC/day.

4.6.4. CLOCK PERIOD INDEPENDENT OF BULK ATP FRACTION

To find out how robust our model of the Kai oscillator is against changes in the steady-state ATP level, we simulate the oscillator at different ATP fractions, α_{ATP}, as was done experimentally in [26]. Figs. 4.2 A,B and C show that the time traces of concentrations of monomers in the T,D and S state with $\alpha_{\mathrm{ATP}} = 1.0, 0.75$ and 0.50, respectively, are in good quantitative agreement with experiments. Both the amplitude of the concentrations and their relative phases agree. However, contrary to experiments, at $\alpha_{\mathrm{ATP}} = 0.25$ the ensemble does not oscillate anymore. The reason is that not enough hexamers bind

6 KaiB monomers to sequester all the KaiA. Since the absence of KaiA normally synchronizes the phosphorylation state of all hexamers, now the ensemble quickly becomes unsynchronized, and the oscillations disappear. Interestingly, in the range of ATP fractions where the Kai system does oscillate, the period of the oscillations is independent of α_{ATP}, as is shown in panel E, and experimentally observed in [7]. This is remarkable, because we did not design the system to have a period independent of α_{ATP}. In panel F we plot the peak, trough and mean phosphorylation level of the oscillations at different values of α_{ATP}. The increase in peak height with α_{ATP} is in good agreement with experiments in [26]. However, while in our simulations the level of the troughs only marginally rises as α_{ATP} decreases, experiments show a considerable increase. Because in our simulations the appearance of free KaiA at the end of the cycle always occurs at a fixed phosphorylation ratio, it is hard to explain this discrepancy with experiments.

4.6.5. WHAT IS THE PRINCIPLE DRIVER OF THE OSCILLATIONS: HYDROLYSIS IN THE CI OR THE CII DOMAIN?

In our model, there are two reactions that break detailed balance and allow the system to oscillate: Hydrolysis of ATP in the CI and in the CII domain. In the results section on phosphorylation dynamics we showed that hydrolysis of ATP in CII, in combination with differential affinity, is sufficient to generate cycles of phosphorylation at the level of the individual hexamers. However, without hydrolysis of ATP in the CI domain, the macroscopic oscillations inevitably come to a halt, because not enough KaiC can reach the inactive conformational state to allow for the necessary level of periodic KaiA sequestration. Clearly, while ATP hydrolysis in CI is not essential for generating cycles at the level of individual hexamers, it is necessary for generating macroscopic oscillations. The question that remains is whether hydrolysis of ATP in CII is likewise necessary for creating coherent, macroscopic oscillations. In the model presented so far, hydrolysis of ATP in CII is needed to allow for dephosphorylation: During the dephosphorylation phase, when a KaiC protein has made the transition from the D to the S state, it has ATP in the binding pocket, which needs to be hydrolyzed to generate ADP, thereby enabling the transition from S to U. However, this ATP hydrolysis reaction does not seem of fundamental importance for breaking detailed balance and creating macroscopic oscillations. Could this system generate oscillations with only turnover of ATP in the CI domain? To address this question, we here investigate a slightly modified version of our original model that has no ATP hydrolysis in the CII domain. This modified version does not represent the real Kai system, but rather serves as a thought experiment to clarify the different thermodynamic roles of ATP hydrolysis in the two domains.

We change our existing model so that KaiC can dephosphorylate without hydrolysis in the CII domain while still having KaiA stimulated phosphorylation in order to synchronize the hexamers. To this end, we set the ATPase rate in the CII domain to zero, while still allowing phosphates to be transferred in both directions between ATP (or ADP) and the serine and threonine residues. For dephosphorylation to occur, there must then be some mechanism other than ATP hydrolysis to introduce ADP's into the CII binding pocket to receive the phosphates. We thus make the nucleotide exchange rate high in the inactive state. In the active state, however, the nucleotide exchange rate should still be low unless KaiA is bound to CII, to preserve the mechanism of KaiA stimulated phospho-

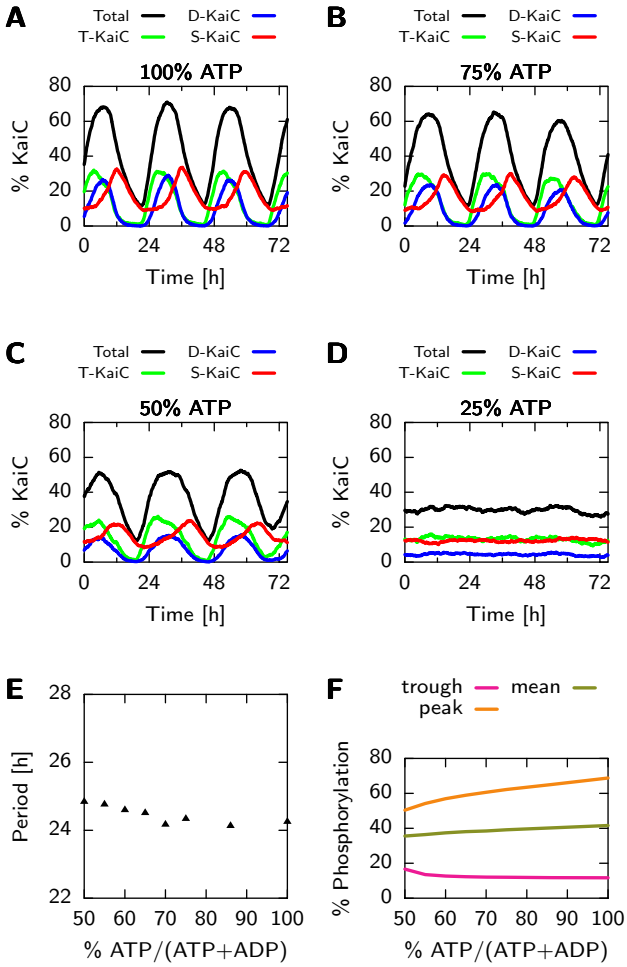

Figure 4.2: The Kai system oscillates over a wide range of ATP fractions, while the period remains unchanged. (A-D) Time traces of monomers in the T state (green), D state (blue), S state (red) and total phosphorylation level (black), for different ATP fractions in the bulk. Amplitudes and phases in good agreement with experiment (Compare with [26], Fig. S1), except that our system does not oscillate at 25% ATP fractions. (E) The average peak-to-peak time of the phosphorylation fraction for a 1000 hour time trace at different ATP fractions. The period is remarkably unaffected by the ATP fraction, even though it has a big influence on the phosphorylation speed, as we showed in Fig. 4.1, D. (F) Mean peak and trough phosphorylation levels at different ATP fractions. The increase in peak-hight with α_{ATP} is in in good agreement with experiment. However, in our simulations, the level of the trough is almost independent of α_{ATP}, while in experiments it increases with α_{ATP} (Compare with Fig. S6 in [7] and Fig. 1 in [26])

rylation. To maintain the correct relative stability of the two KaiC conformations, ADP in the CII domain now also stabilizes the inactive state, such that ADP has a high affinity for

the CII domain when the hexamer is in the inactive state, $\tilde{K}^{\mathrm{CII}}_{\mathrm{ATP/ADP}}$=10.0, and the original low relative affinity for ADP when in the active state $K^{\mathrm{CII}}_{\mathrm{ATP/ADP}}$=0.10. Finally, because ADP from the bulk is required for dephosphorylation, we set α_{ATP}=0.5. Changed parameters are listed in Table 4.3 below.

Parameter		Active	Inactive
$k^{\mathrm{CII}}_{\mathrm{hyd}}$	(h^{-1})	0	0
$K^{\mathrm{CII}}_{\mathrm{ATP/ADP}}$		0.1	10
$k^{\mathrm{CII\cdot ATP}}_{\mathrm{off,0}}$	(h^{-1})	0.6	6.0
$k^{\mathrm{CII\cdot ATP}}_{\mathrm{off,KaiA}}$	(h^{-1})	6.0	6.0
$g^{\mathrm{ATP/ADP}}_{\mathrm{A,I}}$	(kT)	30	30

Table 4.3: Parameters used in the alternative model without hydrolysis in the CII domain that are different from the values in our original model in Tables 4.1 and 4.2. Parameters values are listed for both the active and inactive conformations if they have been changed for either conformation.

Our modified model shows robust macroscopic oscillations, as shown in Fig. 4.3A. The total ATP consumption has dropped to 11.6 ATP/KaiC/day, all due to the CI domain's ATPase activity. This shows that the Kai oscillator can in principal generate macroscopic oscillations with only hydrolysis in the CI domain, and that hydrolysis in CII is not essential. However, as shown in Fig. 4.3B, the rate of dephosphorylation now strongly depends on the ATP fraction in the bulk, because α_{ATP} affects the probability that, upon the D→S transition and subsequent ATP release, CII will bind ADP instead of ATP, which is necessary for the next S→U transition. This dependence of the dephosphorylation rate on the bulk ATP fraction in this modified model is contrary to what is observed in experiments. Even when we allow for hydrolysis in the CII domain in addition to nucleotide exchange in the inactive state, Fig. 4.3C, the dephosphorylation speed still strongly depends on α_{ATP}. As argued in [26] and in chapter 5, such a dependence of the dephosphorylation rate on α_{ATP} would hamper input compensation, and the period of the oscillations would not be constant any more at different ATP fractions. A low nucleotide exchange rate in the inactive state (as included in our main model described in the preceding sections) seems therefore critically important for the real Kai oscillator.

4.7. DISCUSSION

We set out to develop a thermodynamically correct model of the post-translational Kai oscillator that is consistent with the large body of quantitative experimental data available. In particular, the recent experimental observation that KaiC regenerates ATP during dephosphorylation [116, 117] made us rethink the thermodynamics behind the phosphorylation cycle: If phosphorylation and dephosphorylation require no net turnover of ATP, what drives the thermodynamic cycle of the oscillator? We built our new model on our earlier model of the Kai system [15, 19], where each individual KaiC hexamer goes through a cycle of phosphorylation and dephosphorylation. The hexamers are synchronized through the mechanism of differential affinity, in which the affinity of KaiA for KaiC

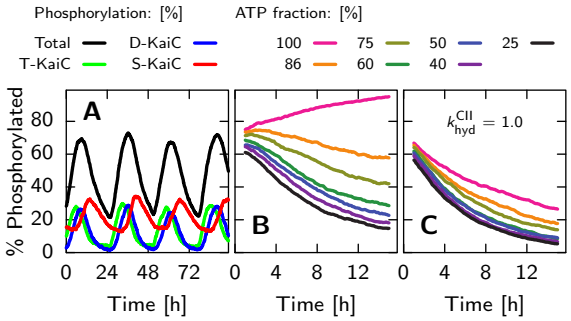

Figure 4.3: An alternative model, without hydrolysis in the CII domain, can generate robust macroscopic oscillations. (A) Time traces of monomers in the T state (green), D state (blue), S state (red) and total phosphorylation level (black), at 50% ATP level in the bulk, for the alternative model with hydrolysis only in the CI domain. (B) Dephosphorylation of KaiC for different bulk fractions of ATP, α_{ATP}. Due to the high nucleotide exchange rate of the CII domain in the inactive conformation, the dephosphorylation speed becomes sensitive to α_{ATP}, contrary to what is observed experimentally. (C) Even when we add the hydrolysis of ATP in the CII domain, $k_{hyd}^{CII} = 1$, the speed still depends on α_{ATP}.

changes with the phosphorylation state and the complex of inactive KaiC with KaiB has the strongest affinity for KaiA; sequestration of KaiA by this complex allows the hexamers lagging behind and which are still in the dephosphorylation phase of the cycle to remove KaiA from the front runners, which have already completed their cycle yet need KaiA to be phosphorylated again. Here, we have extended these ideas with a more detailed monomer model, taking into account the two phosphorylation sites per monomer and the nucleotide binding pockets in the CI and CII domain. This allowed us to include the effects of the ordered phosphorylation cycle on the monomer level, found in [22, 26].

4.7.1. SUMMARY OF THE KEY RESULTS OF THE MODEL

Here we give an overview of the most important conclusions we can draw from our model.

By enhancing nucleotide exchange, KaiA both stimulates phosphorylation and blocks dephosphorylation, preventing futile cycles. In our new model, next to the threonine and serine phosphorylation sites, we explicitly track whether there is ATP or ADP present in the nucleotide binding pockets of the CI and CII domains of KaiC. Dephosphorylation proceeds exclusively by phosphotransfer between the phosphorylation site and ADP in the CII domain; direct release of inorganic phosphate from serine or threonine residues to the bulk can not occur. KaiA acts as a nucleotide exchange factor [127], enhancing the release rates of nucleotides from the CII domain. In combination with the hydrolysis of ATP in the CII domain, KaiA effectively increases the fraction of ATP in the binding pocket, thereby driving the hexamer towards the highly phosphorylated state. Because dephosphorylation with ATP in the binding pocket is impossible, KaiA, when bound, also effectively blocks dephosphorylation. In this way, the serine and thre-

onine sites can be phosphorylated with a rate similar to the rate of dephosphorylation, because the phosphorylation rate does not need to compensate for the spontaneous release of phosphate groups. The capacity of KaiA to block dephosphorylation has two important consequences. First, if KaiA were not to block dephosphorylation, then the latter would inevitably occur, which must then be followed by another round of phosphorylation. Such futile cycles would make the oscillator less efficient. Furthermore, if KaiA were not to impede dephosphorylation, then *net* phosphorylation could only occur if the phosphorylation rate is larger than the dephosphorylation rate. Simulations reveal, however, that in such a model the concentration of threonine phosphorylated KaiC would rise too fast compared to the phosphorylation assay shown in Fig. 4.1B. The timing of the peak of T state monomers would then be wrong.

Differential affinity of the CII domain for KaiA stimulates the ordered phosphorylation of the S and T sites. Next to regulating the ATP level in CII binding pockets, KaiA also steers the order of phosphorylation of the threonine and serine sites, which arises in our model as a logical consequence of differential affinity. In our model, KaiA has a higher affinity for KaiC that is in the T state, with threonine phosphorylated, than in the S state, with serine phosphorylated. Detailed balance then implies that the binding of KaiA raises the energy level of the S state compared to that of the T state (see the energy levels of Fig. 4.2). In this way, the binding of KaiA to CII drives the ordered phosphorylation cycle, where first the threonine site is phosphorylated and then the serine site. We tested this with simulations, and found that KaiC hexamers in a solution with KaiA and no KaiB, go through the ordered phosphorylation cycle of the T and S sites. Hydrolysis of ATP bound to the CII domain is sufficient to drive this cycle. This prediction could conceivably be tested by performing mass-spectrometry experiments [22, 136], tracking the phosphorylation states starting from different initial conditions with synchronized hexamers, and testing whether the measured rates of phosphorylation and dephosphorylation obey detailed balance.

ATP hydrolysis in CI drives conformational switching and macroscopic oscillations. Even though hydrolysis in the CII domain is sufficient to give rise to the phosphorylation cycle of the *individual* hexamers, we conclude that this is not the principal driver of the *macroscopic* oscillations, and in particular of the periodic sequestration of KaiA. If the phosphorylation of KaiC were to directly stabilize the inactive state and consequently the binding of KaiB, then detailed balance would dictate that, conversely, the latter also stabilize the phosphorylated state. Adding KaiB to a dephosphorylation assay would then decrease the dephosphorylation speed, contrary to what is found in experiments [22], which show no change in dephosphorylation dynamics. In our model, the phosphorylation cycle only sets the timing of the conformational switch by regulating the activation energy for ADP dissociation in the CI domain. Hydrolysis of ATP in the CI domain will continually generate ADP in the binding pocket of CI, but only when enough serine sites on CII have been phosphorylated, does the ADP release rate drop sufficiently so that the ADP level in CI will rise. It is this rise in ADP level that stabilizes the inactive state of KaiC. Hydrolysis in the CI domain thus drives the conformational switch and provides the large change in affinity between the active and inactive states, necessary for KaiB binding and KaiA sequestration. The latter, in turn, underlies the synchronization of the phosphorylation cycles of the individual hexamers, which is essential

for generating the macroscopic oscillations in phosphorylation level.

Positive feedback is not essential; time delay and negative feedback are sufficient. Unlike in the model by Van Zon *et al.*, here the phosphorylation dynamics of the hexamer is independent for each monomer, and a hexamer does not need to be fully phosphorylated before flipping to the inactive state. Phosphorylation of the threonine and serine sites in each KaiC monomer has antagonistic effects on the ADP dissociation rate from the CI domain, and consequently on the switch of the conformational state. Due to this antagonism, the conformational switch depends on the difference of T to S phosphorylated residues, and not on the absolute phosphorylation level. Therefore, in our model, a hexamer does not need to go through a full phosphorylation cycle each period, as was the case for the model by Van Zon and coworkers. Furthermore, there is no direct cooperativity between monomers; their states all add linearly to the activation energy for ADP dissociation in CI and the free energies of the conformational states. In particular, the D and S state have a similar effect on the ADP dissociation rate in CI, and hence on the conformation of the hexamer and on its ability to sequester KaiA. This means that the synchronization mechanism of the original monomer model by Rust et. al. [22] does not apply. In their model, KaiA prevents the occupation of the S state by enhancing the transition from the S to the D state, while only the S state sequesters KaiA. This mutual inhibition between KaiA and the S state creates a positive feedback loop for KaiA sequestration that is essential to the oscillations in that model. In contrast, in our model the D and S states both stimulate KaiA sequestration, so it does not exhibit this positive feedback mechanism. Because our model acts at the level of hexamers rather than monomers, it does not need the positive feedback: the delay between the conformational switch and the subsequent binding and sequestration of KaiA is sufficiently long that, together with the negative feedback of KaiA sequestration on phosphorylation, it can generate oscillations.

The model is robust to variations in ATP fraction. Our model correctly reproduces the phosphorylation and dephosphorylation time traces of the T,D and S state monomers. Furthermore, as in experiments [7], dephosphorylation is independent of the bulk ATP fraction, whereas phosphorylation does strongly depend on it. The idea that dephosphorylation is dominated by the phosphotransfer pathway [116, 117] is confirmed by good agreement with the experimentally observed ATP production in the CII domain and the delay in the appearance of inorganic phosphate in the bulk [116]. To compare the ATP consumption of KaiC in our model with experiments, we investigated the ATP fraction in the nucleotide binding pockets and the ATPase activity of KaiC, for different mixtures of Kai proteins. Both the transient dynamics of the ATP fraction in the binding pockets and the steady state ATPase rate in mixtures with KaiC only or KaiC and KaiA are in qualitative agreement with experiments. In a system with macroscopic oscillations, both the phase difference between the ATP fraction in the binding pockets and the phosphorylation fraction and the amplitude of the ATP fraction are in excellent agreement with experiments. To check the robustness of the oscillator against variations in the bulk ATP fraction, we checked whether oscillations persist at lower ATP fractions and if the phosphorylation period is independent of the fraction. We found that the clock period is indeed constant over a wide range of ATP fractions.

4.7.2. OPEN QUESTIONS

Although our model is able to reproduce most of the available experimental data, there are a few observations that it cannot replicate in its current form.

First, in our model, KaiB barely interacts with unphosphorylated KaiC, in contrast with the observation that KaiB lowers the ATPase activity of a solution with only KaiC [114], and that KaiB can bind unphosphorylated WT KaiC [137] at micromolar concentrations. In our model, it is essential that unphosphorylated KaiC is predominantly in the active conformation: Monomers in the U state increase the dissociation rate from ADP in the CI domain, thereby stabilizing the active state which leads to the dissociation of the sequestered KaiA and KaiB at the end of the cycle. The consequence is that unphosphorylated KaiC is predominantly in the active conformation, which has a very low affinity for KaiB. Making the inactive state of unphosphorylated KaiC more stable would remedy this shortcoming of the model, because in the inactive state KaiC can bind KaiB. Consistent with this idea, very recent experiments suggest that the inactive conformational state is indeed more stable: About half of the unphosphorylated KaiC hexamers in a system without KaiA or KaiB, are in the inactive conformational state [140]. However, increasing the affinity of KaiB for unphosphorylated KaiC also increases, in the current model, the capacity of unphosphorylated KaiC to sequester KaiA, which impedes the release of KaiA at the end of the cycle. For future research, it will be interesting to see whether by amending the model these experimental observations can be reproduced.

Secondly, the oscillations in our model come to a standstill when the bulk ATP fraction drops below 40%, while in experiments they continue to exist until the ATP fraction drops below 25% [7, 26]. In our model, when the ATP fraction drops below a critical value of around 40%, not enough hexamers have 6 KaiB monomers bound to sequester all KaiA from solution at the required phase of the oscillation. This impedes the synchronization of the individual hexamers, necessary for coherent macroscopic oscillations. One way to resolve this might be to increase the affinity of unphosphorylated KaiC for KaiB, but as mentioned above, this impairs the release of KaiA at the end of the cycle. Alternatively, or in addition, making the window of KaiA sequestration more deterministic e.g. by making dephosphorylation of the respective monomers within a hexamer more concerted or by more tightly coupling KaiB-KaiC binding to the KaiC phosphorylation state, is expected to extend the range of ATP concentration over which the model exhibits oscillations. We have not investigated the effects of the concerted phosphorylation of hexamers, because it deviates too much from our current model where monomers phosphorylate independently. It is also possible that including monomer exchange might improve synchronization of the KaiC hexamers and thus allow oscillations to persist to lower ATP fractions, but such an effect likewise cannot readily be included in the current model.

Lastly, the ATP consumption in our model is slightly higher than observed. We hypothesized that this can be attributed to the ATP hydrolysis in the CII domain. To provide support for this idea, we looked at an alternative model where the binding of KaiA suppresses ATP hydrolysis in CII. This model did show ATPase rates very similar to experimentally observed values.

4.7.3. PREDICTIONS AND EXPERIMENTAL VERIFICATION

Here we explore the possibilities for experimentally verifying the predictions from our model. Our model is based on two important ingredients: 1) The relative stability of the two conformations is determined by the ATP fraction in the binding pockets of the CI domain and 2) this fraction is set by the number of phosphorylated serine and threonine sites in the CII domain of the hexamer.

The dependence of the conformation on the nucleotide binding state of CI can be tested by measuring the ATP fraction in the binding pockets of KaiC [114, 127] while at the same time probing the conformational state as in the study of [140], where the authors track the fractions of active and inactive KaiC hexamers over time. Our analysis predicts a positive correlation between the fraction of ADP bound to CI and the fraction of inactive KaiC, including KaiC mutants with a different hydrolysis rate constant in CI [114] and in an assay where KaiC is in the presence of KaiA and KaiB, and oscillates over time.

The dependence of the ATP fraction in the binding pockets of the CI domain on the phosphorylation state of the CII domain can be tested by measuring the concentrations of monomers in the U,T,D and S phosphorylated state [22, 136] and again the ATP fraction in the nucleotide binding pockets [114, 127]. We predict that serine-phosphorylated KaiC slows down the dissociation of ADP from CI, decreasing its ATP fraction in the binding pocket, and threonine-phosphorylated KaiC should antagonize this effect. Starting with different ratios and levels of monomers phosphorylated at their threonine and serine sites, either using KaiC phosphomimics or aliquots from an oscillating system, the ADP fraction in the CI binding pockets should show a positive correlation with the *difference* between the number of phosphorylated serine sites and threonine sites. Adding KaiB should enhance the effect, because it cooperatively stabilizes the inactive state with ADP via the MWC mechanism. Our model can also explain the observation described in [141], where they found that adding ATP to unphosphorylated KaiC leads to a transient dip in the ATPase rate: The transient phosphorylation of KaiC temporarily lowers the CI-ADP dissociation rate. Clearly, it would be of interest to repeat these experiments starting with KaiC in different phosphorylation states, and in the presence and absence of KaiB. Our model predicts that, starting from fully phosphorylated KaiC, during dephosphorylation the ATP fraction in the binding pockets will exhibit a dip (see Fig. 4.1).

Related to this, and more specifically, our model predicts that the ADP dissociation rate in the CI domain is set by the relative number of phosphorylated serine and threonine sites in the CII domain, and not by their absolute levels. This implies that not all the residues have to be phosphorylated before a hexamer can switch to the inactive conformation and complete its cycle. Indeed as we show in a forthcoming publication, at lower ATP fractions of the buffer, hexamers go through a smaller phosphorylation cycle. This could in principle be tested experimentally if it were possible to track individual hexamers as they go through their phosphorylation cycle.

While these experiments test our predictions on the connection between the CI and CII domain, our analysis also predicts an interesting consequence of the idea that phosphotransfer is the major pathway for dephosphorylation [116, 117]. This could be tested by revisiting the experiments on dephosphorylation of radioactively labeled KaiC [116], but now in solution with non-hydrolyzable ATP. Since the ATP can not be hydrolyzed,

there will be no ADP in the CII binding pockets, and the phosphate groups on the S and T sites in KaiC can not be transferred to ADP in the CII domain. This should significantly slow down the dephosphorylation speed if indeed phosphotransfer is the dominant pathway.

Our model of the interaction of KaiA with the CI and CII domains of KaiC also allows us to make predictions for how long KaiA is bound to one of the domains during an oscillation. In our model, the dissociation rate of KaiA from the CII domain is much higher than the frequency of the oscillation. This is essential for differential affinity, where KaiA continually binds different KaiC hexamers during the phosphorylation phase to promote the phosphorylation of hexamers that are lagging behind. Our model therefore predicts a sharp peak in the distribution of times for which the CII domain of a certain KaiC hexamer is bound to KaiA during an oscillation period, $\Delta t_{\text{CII·KaiA}}$. Indeed, Fig. 4.1A shows a clear single peak in this distribution. When we lower the dissociation rate of KaiA from the CII domain, this distribution broadens, Fig. 4.1B, indicating differential affinity is hampered. One might think that when the dissociation rate is decreased even further, the distribution in $\Delta t_{\text{CII·KaiA}}$ becomes bimodal, because once KaiA is bound to the CII domain of a certain hexamer, it will continue to stay bound to this hexamer during the whole phosphorylation phase of an oscillation cycle; because KaiA is limiting, this means that other KaiC hexamers will not, via their CII domain, bind KaiA during that cycle. In this case one fraction of hexamers does not or only very briefly binds KaiA via the CII domain while the other fraction is bound to KaiA for most of the time during the phosphorylation phase. However, in our model oscillations stop when we decrease the dissociation rate of KaiA to such a low level to allow for a bimodal distribution. Nevertheless, bi-modality could arise in experiments, which would indicate that the role of KaiA during the phosphorylation phase is very different from what we predict in our model. For KaiA bound to the CI domain, our simulations show exponentially distributed bound times, Fig. 4.1B. This distribution is bimodal, as only 70% of the hexamers sequester KaiA, while the other 30% do not make it to the inactive state and bind six KaiB during an oscillation. Techniques to follow protein complex formation at the single molecule level have been developed [142], suggesting that future experiments might be able to reveal how long KaiA is bound to KaiC during an oscillation cycle.

4.8. METHODS

We model the post-translational Kai oscillator using kinetic Monte Carlo [143, 144]. Each monomer has two nucleotide binding states for both the CI and CII domain and 4 phosphorylation states. Additionally, a hexamer can be in an active or inactive state. This results in $(2\cdot2\cdot4)^6\cdot2 = 2^{25}$ different states a single hexamer can be in. Furthermore, there are the binding reactions with the KaiC hexamer: 2 for KaiA with the CII domain, 6 for KaiB monomers to the CI domain and another 6 reactions for KaiA binding to the CI domain of KaiC with 6 KaiB monomers. The total number of reactions therefore exceeds a billion. Clearly, the reaction combinatorics makes a straightforward ODE or Gillespie simulation unfeasible. We thus set out to design a dedicated kinetic Monte Carlo algorithm to simulate the Kai system, where we do not have to write down all possible reactions explicitly.

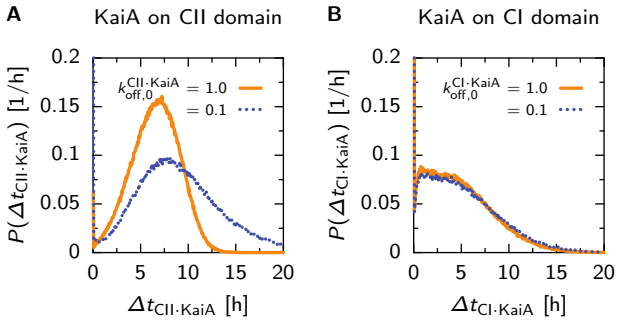

Figure 4.1: Probability density of the time, per period of the phosphorylation cycle, the CII domain of a KaiC hexamer is bound to KaiA, $\Delta t_{\text{CII·KaiA}}$ (A), or the CI domain has at least one KaiA dimer bound, $\Delta t_{\text{CI·KaiA}}$ (B). We compare situations with our standard value for the KaiA dissociation rate, $k_{\text{off,0}}^{\text{CII·KaiA}} = 1.0$ (orange solid line) with a much lower value $k_{\text{off,0}}^{\text{CII·KaiA}} = 0.1$ (blue dotted line). (A) For the CII domain, both values of the KaiA dissociation rate show unimodal distributions of the time KaiA is bound. This indicates that, during a period, KaiA binds to all hexamers in the ensemble equally likely. Of all hexamers, only 1% do not bind KaiA to CII at any time during the period. (B) For the CI domain, the time KaiA is sequestered by KaiC is roughly exponentially distributed, and is the same for both KaiA dissociation rates from the CII domain. Here, 30% of the hexamers do not sequester KaiA at all during a period, indicating many hexamers do not sequester KaiA during a cycle.

To this end, we designed an algorithm that keeps track of N^{hex} KaiC hexamers, which in turn consist of six explicitly simulated monomers, and $N_{\text{tot}}^{\text{KaiA}}$ KaiA dimers. Since we have defined individual hexamers and monomers, we can calculate the propensity that a reaction occurs in a particular hexamer, and, in turn, the propensity that a reaction occurs in a monomer that is part of this hexamer. This allows us to create a layered version of the original algorithm by Gillespie [143], where we first determine the next reaction time and in which hexamer this reaction will take place. Then, in the next step, we determine which reaction or monomer of this hexamer fires. If a monomer fires, we choose which reaction of this monomer happens. This layered approach allows us to separate state changes in a hexamer that only modify the reaction propensities of that specific hexamer from state changes that influence the KaiA concentration in solution, which affects all hexamers. This greatly reduces the reaction combinatorics and the computational cost of the algorithm.

Specifically, the state of the whole system, $\mathbf{s}^{\text{tot}}$, consists of the state of the hexamers with index h, $\mathbf{s}_{\text{h}}^{\text{hex}}$, and the number of KaiA dimers in solution, $n_{\text{sol}}^{\text{KaiA}}$: $\mathbf{s}^{\text{tot}} = \{\{\mathbf{s}_{\text{h}}^{\text{hex}}\}_{h=1}^{N^{\text{hex}}}, n_{\text{sol}}^{\text{KaiA}}\}$. The hexamer state contains the state vectors of its six monomers with index m, $\mathbf{s}_{\text{h,m}}^{\text{mon}}$, the conformational state, C, the number of KaiA bound to CI, $n^{\text{CI·KaiA}}$, the number of KaiB bound to CI, $n^{\text{CI·KaiB}}$, and the number of KaiA bound to the CII domain, $n^{\text{CII·KaiA}}$: $\mathbf{s}_{\text{h}}^{\text{hex}} = \{\{\mathbf{s}_{\text{h,m}}^{\text{mon}}\}_{m=1}^{6}, C, n^{\text{CI·KaiA}}, n^{\text{CI·KaiB}}, n^{\text{CII·KaiA}}\}_h$. The state vector of a monomer consists of the threonine and serine phosphorylation site and the nucleotide binding pockets of the CI and CII domain: $\mathbf{s}_{\text{h,m}}^{\text{mon}} = \{S, T, n_{\text{nucl}}^{\text{CI}}, n_{\text{nucl}}^{\text{CII}}\}_{h,m}$.

Given the system state, we can calculate the firing propensity of reaction μ that changes the state vector of monomer m, which is part of hexamer h, $q^{\mu}_{\mathrm{h,m}}(\mathbf{s}^{\mathrm{hex}}_{\mathrm{h}})$. Note that this reaction propensity can depend on the state of the whole hexamer, and therefore is a function of the hexamer state vector, and not only of the state vector of monomer m. The propensity for firing a reaction with index ν, which changes the state variables of hexamer h, denoted $q^{\nu}_{\mathrm{h}}(\mathbf{s}^{\mathrm{hex}}_{\mathrm{h}}, n^{\mathrm{KaiA}}_{\mathrm{sol}})$, depends on the state vector of hexamer h only, and the number of KaiA in solution. Given these reaction propensities, we can calculate the accumulated propensities, denoted by $\tilde{q}$, of firing a single monomer m in hexamer h, a single hexamer h and the total propensity as

$$\tilde{q}^{\mathrm{mon}}_{\mathrm{h,m}}(\mathbf{s}^{\mathrm{hex}}_{\mathrm{h}}) = \sum_{\mu} q^{\mu}_{\mathrm{h,m}}, \tag{4.19}$$

$$\tilde{q}^{\mathrm{hex}}_{\mathrm{h}}(\mathbf{s}^{\mathrm{hex}}_{\mathrm{h}}, n^{\mathrm{KaiA}}_{\mathrm{sol}}) = \sum_{m=1}^{6} \tilde{q}^{\mathrm{mon}}_{\mathrm{h,m}} + \sum_{\nu} q^{\nu}_{\mathrm{h}}, \tag{4.20}$$

$$\tilde{q}^{\mathrm{tot}}(\{\mathbf{s}^{\mathrm{hex}}\}_{h=1}^{N^{\mathrm{hex}}}, n^{\mathrm{KaiA}}_{\mathrm{sol}}) = \sum_{h=1}^{N^{\mathrm{hex}}} \tilde{q}^{\mathrm{hex}}_{\mathrm{h}}, \tag{4.21}$$

respectively.

Since the firing of a reaction is a Markov process, the probability of hexamer h firing in the infinitesimal time interval $[t+\tau, t+\tau+d\tau]$ is

$$P(\tau, h|\mathbf{s}^{\mathrm{tot}}, t)d\tau = \tilde{q}^{\mathrm{hex}}_{\mathrm{h}} \exp\left(-\tilde{q}^{\mathrm{tot}}\tau\right) d\tau. \tag{4.22}$$

Now, given two random numbers, ρ_1, ρ_2, drawn from a uniform distribution with domain $[0,1)$, we calculate the next event time, τ, and the hexamer to fire, h, as

$$\tau = \frac{1}{\tilde{q}^{\mathrm{tot}}} \ln\left(\frac{1}{\rho_1}\right), \tag{4.23}$$

$$h = \text{the smallest integer satisfying } \sum_{h'=1}^{h} \tilde{q}^{\mathrm{hex}}_{\mathrm{h}} > \rho_2 \tilde{q}^{\mathrm{tot}}, \tag{4.24}$$

respectively.

Having defined the important propensities, our dedicated kinetic Monte Carlo algorithm becomes

0. Initialize the time $t = t_0$ and the system's state $\mathbf{s}^{\mathrm{tot}} = \mathbf{s}^{\mathrm{tot}}_0$. Calculate the propensities $q^{\mu}_{\mathrm{h,m}}$, q^{ν}_{h}, $\tilde{q}^{\mathrm{mon}}_{\mathrm{h,m}}$, $\tilde{q}^{\mathrm{hex}}_{\mathrm{h}}$ and $\tilde{q}^{\mathrm{tot}}$.
1. Calculate the time interval to the next reaction, τ, using Eq. 4.23.
   a. Choose which hexamer, h, to fire, with $P(h|\tau) = \tilde{q}^{\mathrm{hex}}_{\mathrm{h}} / \tilde{q}^{\mathrm{tot}}$.
   b. Choose which reaction, ν, with $P(\nu|\tau, h) = q^{\nu}_{\mathrm{h}} / \tilde{q}^{\mathrm{hex}}_{\mathrm{h}}$ or monomer, m, with $P(m|\tau, h) = q^{\mathrm{mon}}_{\mathrm{h,m}} / \tilde{q}^{\mathrm{hex}}_{\mathrm{h}}$, to fire.
   c. If a monomer was chosen, choose which reaction to fire with $P(\mu|\tau, h, m) = q^{\mu}_{\mathrm{h,m}} / \tilde{q}^{\mathrm{mon}}_{\mathrm{h,m}}$.

2. Fire reaction and update the state vector of hexamer h, and $n_{\mathrm{sol}}^{\mathrm{KaiA}}$ in case of a bimolecular reaction. Recalculate all reaction propensities q_{h}^{ν} and $q_{\mathrm{h,m}}^{\mu}$ for each monomer m. In case a bimolecular reaction was fired, change $n_{\mathrm{sol}}^{\mathrm{KaiA}}$ accordingly, and update the bimolecular reaction propensities in all hexamers.

3. Recalculate $\tilde{q}_{\mathrm{h,m}}^{\mathrm{mon}}$ and $\tilde{q}_{\mathrm{h}}^{\mathrm{hex}}$ for the fired hexamer h. In case of a bimolecular reaction, also recalculate $\tilde{q}_{\mathrm{h}'}^{\mathrm{hex}}$ for all hexamers h'. Update $\tilde{q}^{\mathrm{tot}}$.

4. Record $(t, \mathbf{s}^{\mathrm{tot}}(t))$ as desired, return to step 1.

5

PERIOD ROBUSTNESS AND ENTRAINABILITY UNDER CHANGING NUCLEOTIDE CONCENTRATIONS IN THE IN-VITRO KAI CIRCADIAN CLOCK

Circadian clocks must be entrainable to keep their oscillations in phase with the day-night rhythm. On the other hand, they must also exhibit input compensation: the period must remain about one day in different constant environments. The post-translational oscillator of the Kai system can be entrained by transient or oscillatory changes in the ATP fraction, yet is insensitive to constant changes in this fraction. We study in three different models of this system how these two seemingly conflicting criteria are met: the Van Zon model (Van Zon et al., PNAS, 2007), the Rust model (Phong et al., PNAS, 2013), and our new model presented in chapter 4. We find that the new model exhibits the best trade-off between input compensation and entrainability: on the footing of equal phase-response curves, it exhibits the strongest input compensation. Performing stochastic simulations at the level of individual hexamers allows us to identify a new mechanism, which is employed by the new model to achieve input compensation: At lower ATP fraction, the individual hexamers make a shorter cycle in the phosphorylation state space, which compensates for the lower pace at which they traverse the cycle.

5.1. INTRODUCTION

Circadian clocks help organisms to coordinate their metabolism and behavior with the daily changes in the environment [145]. These clocks are prevalent in a wide range of organisms from bacteria to humans, but all have three important features in common: First, a circadian clock is a *self-sustained oscillator*, meaning that oscillations persist even in the absence of any external cue, with a rhythm of about 24 hours. Second, the clock is *entrainable* so that its oscillations can be kept in phase with the day-night rhythm. To this end, a circadian clock must be able to respond to daily cues such as rhythmic changes in light and temperature. Third, the clock has some form of *input compensation*, such that the period is constant, even when the temperature or light intensities change for longer times. Importantly, these last two requirements seem to be at odds with each other. A clock that is easily entrained, because, for instance, the rates of the biochemical reaction that determine the period strongly depend on temperature, would seem to have a period that also depends on the temperature. This clock would be a bad predictor of time. As was shown in [146, 147], cyanobacteria with clock mutants having intrinsic periods ranging from 22 to 30 hours, grow significantly slower in a 12:12 light-dark cycle than wild-type cells. Therefore, input compensation is critically important to the function of a circadian clock. On the other hand, if a clock would be completely insensitive to any external change, it would trivially have a period that is insensitive to changes in the environment. However, such a clock would not be entrainable, and as a result of biochemical noise, it would inevitably run out of phase with the day-night rhythm. Clearly, entrainability is essential to keep the clock in phase with the day, which means that the clock needs to be sensitive to changes in the environment. What the properties of an oscillator should be to fulfill both conditions has been studied in different systems, both experimentally and theoretically [6, 52, 148–151].

We use the post-translational oscillator of the Kai circadian clock, found in the freshwater cyanobacterium *Synechococcus elongatus* PCC 7942, as a model system. It is well known that this circadian clock is sensitive to transient changes in its environment, and hence entrainable, yet robust to permanent changes in its input, thus showing input compensation [5, 53, 73, 152, 153]. It was shown in an seminal experiment in 2005, that the core oscillator can be reconstituted in-vitro, and consists of the three proteins KaiA, KaiB and KaiC, in solution with ATP [12]. KaiC is a phosphotransferase [116, 117], which, depending on its conformation, switches between phases of phosphorylation and dephosphorylation [15, 123]. KaiA is a nucleotide exchange factor, that facilitates exchange of ADP to ATP in the nucleotide binding pockets of KaiC [127], which enhances phosphorylation of KaiC (see chapter 4). KaiB counteracts the effect of KaiA, by binding to KaiC, and sequestering KaiA from solution [109, 111]. Remarkably, even in an in-vitro essay, the phase of the oscillator can be reset while keeping a robust circadian period, by changing the redox state [154, 155], temperature [12, 136, 156] or the ratio of ATP to ADP in the buffer [7, 26].

In this work, we study entrainability and input compensation under changes in the bulk ATP fraction, by comparing three different models of the post-translational Kai oscillator. We have chosen to investigate the influence of the ATP fraction, because experiments suggest that this is the most important mechanism for entrainment of the oscillator [30], and becuase its effect has been studied in great detail [7, 26, 157]. Specifically,

we will study the effect of the ATP fraction on the hexamer model by Van Zon *et. al.* [15], the extended monomer model, based on the original model by Rust *et. al.* [22, 26], and the new model introduced in this thesis (see chapter 4). In general, as shown in experiments, the effect of the ATP fraction on the oscillator is that the phosphorylation rate is roughly proportional to the ATP fraction, and the overall rate of dephosphorylation is unaffected [7]. We will study the entrainability by applying a 6 hour pulse, during which the ATP fraction is reduced from 100% to 40%, at different phases of the oscillation, and compare the maximally induced phase shift in each model. We analyze period stability in each model by running the models at different constant ATP fractions and observe how the period and other important quantities of the oscillations change. Lastly, simulations of our new model allow us to track each individual hexamer, and measure the timing between states as it proceeds through its cycle. In this way, to our knowledge for the first time, we can predict how individual KaiC hexamers respond to external cues.

Below we briefly describe the three models, and the mechanisms they employ to achieve input compensation and entrainability. We discuss three mechanisms of input compensation, two of which have been partly identified elsewhere, and one of which is novel. The first, present in all models, is related to the delay between the moment the KaiC front runners (hexamers that are more phosphorylated than the average) reach the top of the cycle, which is the point they no longer need KaiA to progress along the cycle, and the later time at which the front runners reach a state in which they sequester KaiA [26]. As the ATP fraction in the bulk is reduced, the rate of phosphorylation decreases, which means that it takes longer to reach the top of the cycle, thus extending the phosphorylation phase. However, the lower rate of phosphorylation also means that during the delay less hexamers make it to the top of the cycle. This decreases the number of KaiC molecules that can participate in sequestering KaiA, thus shortening the dephosphorylation phase, counteracting the longer phosphorylation phase. The second mechanism, present only in the Rust model, is related to the positive feedback in that model which results from the mutual inhibition between the sequestration of KaiA by serine-phosphorylated KaiC (the S state, see chapter 4), and KaiA stimulating the transition from serine-phosphorylated to doubly phosphorylated KaiC (the D state). The third mechanism, present in our new model, and first identified here, concerns the path individual hexamers take through phosphorylation state space: At lower ATP fraction, the individual hexamers move through a smaller phosphorylation cycle, which compensates for the lower rate at which they traverse this cycle. The generic idea of input compensation via a trade-off between the size of the cycle in state space and the speed at which it is traversed was presented by Hatakeyama and Kaneko [149]. Here we present a specific manifestation of this mechanism at the level of individual hexamers. This mechanism is absent in the Van Zon and Rust models, where the phosphorylation cycles of the hexamers or monomers through state space are independent of the ATP fraction.

As period robustness can trivially be achieved by making the oscillator completely insensitive to the ATP fraction, we next investigate the phase response curve for each model, which describes the induced phase shift of the clock upon a pulse of ADP [145]. We find that the models of Van Zon and Rust are, depending on their sensitivity to the ATP fraction, either strongly entrainable yet have a period that depends on the ATP fraction or have a very stable period but are not entrainable. In contrast, our new model

exhibits both input compensation and entrainability.

5.2. THEORY

Here we give a description of the three models of the post-translational Kai circadian clock studied here, and how we included the sensitivity of the phosphorylation rates to the ATP fractions in the bulk. Furthermore, we explain which mechanisms for period stability are present in each model, to compensate for the dependence of the phosphorylation rates on the ATP fraction. Two of the three mechanisms that we describe below have been partly presented elsewhere [26], yet we will discuss here how they are implemented specifically in the respective models. For completeness, we give here also a qualitative description of the third, novel, mechanism, which is employed only by the new model; this mechanism is discussed in much more detail in the Results section.

5.2.1. VAN ZON MODEL

The Van Zon model describes the phosphorylation cycle at the level of KaiC hexamers, and does not explicitly keep track of the KaiC monomers. A simplified scheme of the model is shown in Fig. 5.1A: A hexamer can be in the active conformational state, denoted by C_i, or in the inactive state, denoted by $\tilde{C}_i$, where i denotes its phosphorylation level. In the presence of free KaiA, active KaiC (enclosed by the green box in Fig. 5.1A) is phosphorylated with a rate which dependents on the KaiA concentration. When the hexamer has reached the fully phosphorylated state, C_6, it flips to the inactive conformation, $C_6 \rightarrow \tilde{C}_6$, where it immediately binds KaiB. The delay between full phosphorylation and KaiA sequestration, essential for synchronized oscillations, is set by the two dephosphorylation steps, $\tilde{C}_6 \rightarrow \tilde{C}_5 \rightarrow \tilde{C}_4$ (blue box). The complexes $\tilde{C}_4 - \tilde{C}_1$, have a very high affinity for KaiA, allowing them to sequester all free KaiA from the solution (red box). This sequestration of KaiA forces the front runners that have reached the bottom of the cycle (C_0) and are ready to be phosphorylated again, to wait, because KaiA is needed for phosphorylation. Sequestration of KaiA thus allows the laggards (the hexamers that are falling behind in the phosphorylation cycle) to catch up with the front runners, leading to the synchronization of the oscillations of the individual hexamers. Only when most KaiC has reached C_0, is KaiA released in solution, and can a new phosphorylation cycle start again.

We use a coarse grained description to model the effect of the ATP fraction, α_{ATP}=[ATP]/([ATP]+[ADP]), where [ATP] and [ADP] are the ATP and ADP concentrations in the bulk, respectively, on the phosphorylation rates. Just like in the Rust model [7], we assume that, when KaiA is bound, the probability of having ATP instead of ADP bound, is

$$\beta_{\mathrm{ATP}} = \frac{\alpha_{\mathrm{ATP}}}{\alpha_{\mathrm{ATP}} + K_{\mathrm{ATP/ADP}}(1-\alpha_{\mathrm{ATP}})}, \tag{5.1}$$

where $K_{\mathrm{ATP/ADP}}$ is the relative dissociation constant for binding ATP over ADP. The effective phosphorylation rates become $k_{\mathrm{phos}} = \beta_{\mathrm{ATP}}\, k^0_{\mathrm{phos}}$, where k^0_{phos} is the phosphorylation rate at 100% ATP. Dephosphorylation rates are independent of α_{ATP}.

The Van Zon model employs one mechanism for period stability, which is partly identified in [26]. The mechanism is a direct consequence of the temporal delay be-

tween the moment a hexamer reaches the top of the cycle, i.e. the state $\tilde{C}_6$ in which it no longer needs KaiA to progress along the cycle, and the time at which it reaches $\tilde{C}_4$ and starts sequestering KaiA. Thus, in particular, there is a lag between the moment when enough KaiC hexamers to fully sequester KaiA have passed $\tilde{C}_6$, and so are committed to the path towards sequestration, and the moment when full sequestration is actually reached. The number of additional hexamers that reach $\tilde{C}_6$ during this delay is given by the duration of the delay multiplied by the phosphorylation speed. Importantly, while the duration of the delay is independent of the ATP fraction (since dephosphorylation from $\tilde{C}_6$ till $\tilde{C}_4$ is independent of the ATP fraction), the rate of phosphorylation decreases as the ATP fraction decreases. Consequently, the lower the ATP fraction, the smaller the number of hexamers that can reach the state $\tilde{C}_6$ during the delay. The smaller number of $\tilde{C}_6$, in turn, leads to a shorter time interval in which all KaiA stays sequestered. This shortens the dephosphorylation phase, which counteracts the longer phosphorylation phase, stabilizing the period.

Fig. 5.1B shows time traces for the KaiC phosphorylation level $p(t) = \sum_{i=1}^{6} i(C_i + \tilde{C}_i)/(6\,\mathrm{KaiC_{tot}})$ (dotted lines), and the fraction of inactive hexamers (solid lines), at α_{ATP}= 100% and 50%. The number of hexamers that can sequester KaiA increases with the amplitude of the inactive fraction, such that this amplitude sets the duration of the dephosphorylation phase. Indeed, as a clear signature of the stability mechanism, at 50% ATP fraction, both the phosphorylation level and the fraction of inactive KaiC rise slower while having a lower amplitude compared to the oscillations at 100% ATP fraction.

5.2.2. RUST MODEL

Contrary to our new model and the Van Zon model, the Rust model describes the oscillations at the level of single monomers [7, 22, 26]. As shown in Fig. 5.1C, each monomer goes through the ordered phosphorylation cycle U → T → D → S → U. Only when monomers have reached the phosphorylation states S and D, they can bind KaiB with a low rate, and form D·B and S·B, respectively. Importantly, only the S·B state sequesters KaiA, while KaiA impedes the occupation of the S·B state by enhancing the transition from S·B back to D·B. This mutual inhibition between KaiA and the S·B state creates a positive feedback loop for KaiA sequestration that is essential for oscillations. Initially, when D·B transforms into S·B, KaiA stimulates the reverse reaction. During this period of a quasi-equilibrium between the D·B and S·B states, the concentration of their sum rises, [D·B] + [S·B], up to the point that [S·B] reaches a level where it sequesters all KaiA. At this moment, the positive feedback is broken, and the system rapidly switches to the dephosphorylation phase in which KaiA is fully sequestered for a long time. The positive feedback thus creates a sharp transition between the phase in which KaiA is free to simulate phosphorylation, and the phase in which all KaiA is sequestered.

To include the effect of the bulk ATP fraction, we use the same coarse grained description on the phosphorylation rates as in the original work: $k_{\mathrm{phos}} = \beta_{\mathrm{ATP}}\, k^0_{\mathrm{phos}}$, where β_{ATP} is defined in Eq. 5.1. As was shown in [26], the ATPase activity in the CI domain, which sets the rate of KaiB binding in this model, does not depend on the bulk ATP fraction.

The Rust model implements two methods for period stability, where the first is similar to that identified in the Van Zon model and described by Phong *et. al.* [26]. The slow

KaiB binding step creates a temporal delay between, on the one hand, the phosphorylation states D and S, which can only be reached in the presence of free KaiA, and, on the other hand, the state S·B, which sequesters KaiA. This delay allows monomers to reach the D and S state through phosphorylation before all KaiA is sequestered. The number of monomers that can reach the KaiB bound states increases with the phosphorylation speed, set by α_{ATP}, and will determine the duration of the period in which all KaiA is sequestered.

The previous mechanism assumes that the delay between reaching the D or S phosphorylated states and sequestering KaiA is constant, but the input compensation can be enhanced further if this delay gets shorter when ADP is added. More precisely, in order to sequester all KaiA, the concentration of S·B monomers has to fulfill $n^{\mathrm{S}}_{\mathrm{seq}}[(S\cdot B)_{\mathrm{min}}] \geq [\mathrm{KaiA_{tot}}]$, where $n^{\mathrm{S}}_{\mathrm{seq}}$ is the number of KaiA monomers sequestered by a single S·B KaiC monomer and $[(S\cdot B)_{\mathrm{min}}]$ is the minimal concentration of S·B monomers to sequester all KaiA, $\mathrm{KaiA_{tot}}$. Because the transitions between the states S·B and D·B are faster compared to the transitions from these states to the U and T, the S·B $\rightleftharpoons$ D·B transitions are in quasi equilibrium. This means that [D·B] is related to [S·B] via the steady state relation, $[\mathrm{D\cdot B}]/[\mathrm{S\cdot B}] \approx k_{\mathrm{SD}}(\alpha_{\mathrm{ATP}})/k_{\mathrm{DS}} = f(\alpha_{\mathrm{ATP}})$, which depends on the effective (de)phosphorylation rates, $k_{\mathrm{SD}}(\alpha_{\mathrm{ATP}})$ and k_{SD}, respectively, and, importantly, on the bulk ATP fraction. Consequently, the amount of D·B required to have enough S·B to sequester all KaiA, is $\mathrm{D\cdot B} \approx \mathrm{S\cdot B}\, f(\alpha_{\mathrm{ATP}})$, which is related to α_{ATP}. Thus, at a lower bulk ATP fraction, the concentration of KaiC bound to KaiB necessary to sequester all KaiA, [D·B]+[S·B], will be lower, which compensates for the slower formation of these complexes during the phosphorylation phase. Indeed, as shown in Fig. 5.1D, the concentration of KaiC-bound KaiB increases much slower at a 50% ATP fraction compared to 100%, but the concentration of KaiC-bound KaiB at the moment that all KaiA is sequestered, is also lower. Because fewer D phosphorylated monomers are required to sequester all KaiA, the subsequent sequestration time is smaller, shortening the period even more.

5.2.3. NEW MODEL

Our new model, explained in more detail in chapter 4, is again a hexamer model. This model, shown in Fig. 5.1E, explicitly describes the state of individual monomers, and in particular their serine and threonine phosphorylation sites. Each monomer in a hexamer is phosphorylated in a well defined order: First the threonine site is phosphorylated and then the serine site. Phosphorylation of the two sites has an antagonistic effect on the conformational state of the hexamer: The U and T states stabilize the active conformation and the D and S states stabilize the inactive conformation. Due to this antagonism, the relative stability of the conformations do not depend on the absolute number of monomers in a certain state, but rater on the difference between the number of phosphorylated threonine and serine sites [18]. Roughly, when more serine sites are phosphorylated than threonine sites, the hexamer will switch conformation. After flipping to the inactive state, the hexamer binds KaiB, but it can only sequester KaiA after 6 KaiB monomers are bound. This delay allows hexamers lagging behind to continue phosphorylation and reach the inactive state, which is an essential property of our model to generate robust oscillations.

The new model explicitly simulates the binding and unbinding of nucleotides, and the hydrolysis of ATP, in the CII domain, where the ATP is used to phosphorylate the threonine and serine sites. The ATP fraction in the CII binding pocket is dependent on the bulk ATP fraction, because the binding of nucleotides is directly proportional to α_{ATP}. Furthermore, the ATP fraction in the binding pocket dependents on the hydrolysis rate of ATP, and the relative affinity for ATP and ADP. Importantly, both the effective phosphorylation *and* dephosphorylation rates depend on the ATP fraction of the binding pockets, because both events occur via phosphotransfer with the nucleotide. This means that a change in α_{ATP} has a much bigger effect on the (de)phosphorylation dynamics compared to models where dephosphorylation proceeds through a α_{ATP} independent phosphatase reaction.

Motivated in chapter 4, we choose the relative affinity of nucleotides for the binding pockets of the CII domain, $K^{\mathrm{CII}}_{\mathrm{ATP/ADP}} = 0.1$, much lower than what is used in the Rust model [7]. We have, however, also investigated the behavior of the other two models with a relative affinity that is similar to our model. In the results section below, we will compare the other models with the high and low relative affinity, with our new model.

The oscillator employs two mechanisms of period stability. First, due to the slow KaiB binding, there is a delay between flipping to the inactive conformation after phosphorylation, and KaiA sequestration. This creates a pool of hexamers in the inactive conformation that increases with the speed of phosphorylation, set by the ATP fraction. Again, at lower α_{ATP}, fewer hexamers make it to the inactive state, such that the dephosphorylation phase is shorter, which counteracts the longer phosphorylation phase.

The second mechanism for period stability is related to the path individual hexamers traverse through phosphorylation state space before switching to the inactive conformation. A hexamer switches to the inactive state when the number of phosphorylated serine sites, n_{S}, exceeds the number of phosphorylated threonine sites, n_{T}. Due to the ordered phosphorylation of each monomer, the threonine sites are phosphorylated before the serine sites, such that a hexamer makes a wide arch in phosphorylation state space, $(n_{\mathrm{T}}, n_{\mathrm{S}})$, before the diagonal, and the flipping criterion $n_{\mathrm{S}} > n_{\mathrm{T}}$, is reached. As the effective phosphorylation rates decrease for lower α_{ATP}, the size of the arch decreases, as discussed in more detail below. This shorter path in state space counteracts the effect of a slower progression along the path (due to the slower phosphorylation), and creates another mechanism for input compensation, at the level of inidividual hexamers.

Note that the positive feedback loop on KaiA sequestration in the Rust model, due to the mutual repression between KaiA and the S·B state of KaiC, is not present in our model. In this feedback loop in the Rust model, KaiA stimulates the transition from the S to the D state, thereby preventing its own sequestration (because only the S·B state significantly sequesters KaiA). However, because in our model both the S and the D state stabilize the ADP bound state in the binding pocket of the CI domain, which stabilizes the inactive conformation and the subsequent KaiB binding (leading to KaiA sequestration), KaiA does not prevent its own sequestration by stimulating the S to D transition. Therefore, the mechanism of input compensation in the Rust model resulting from the positive feedback loop, does not apply to our model.

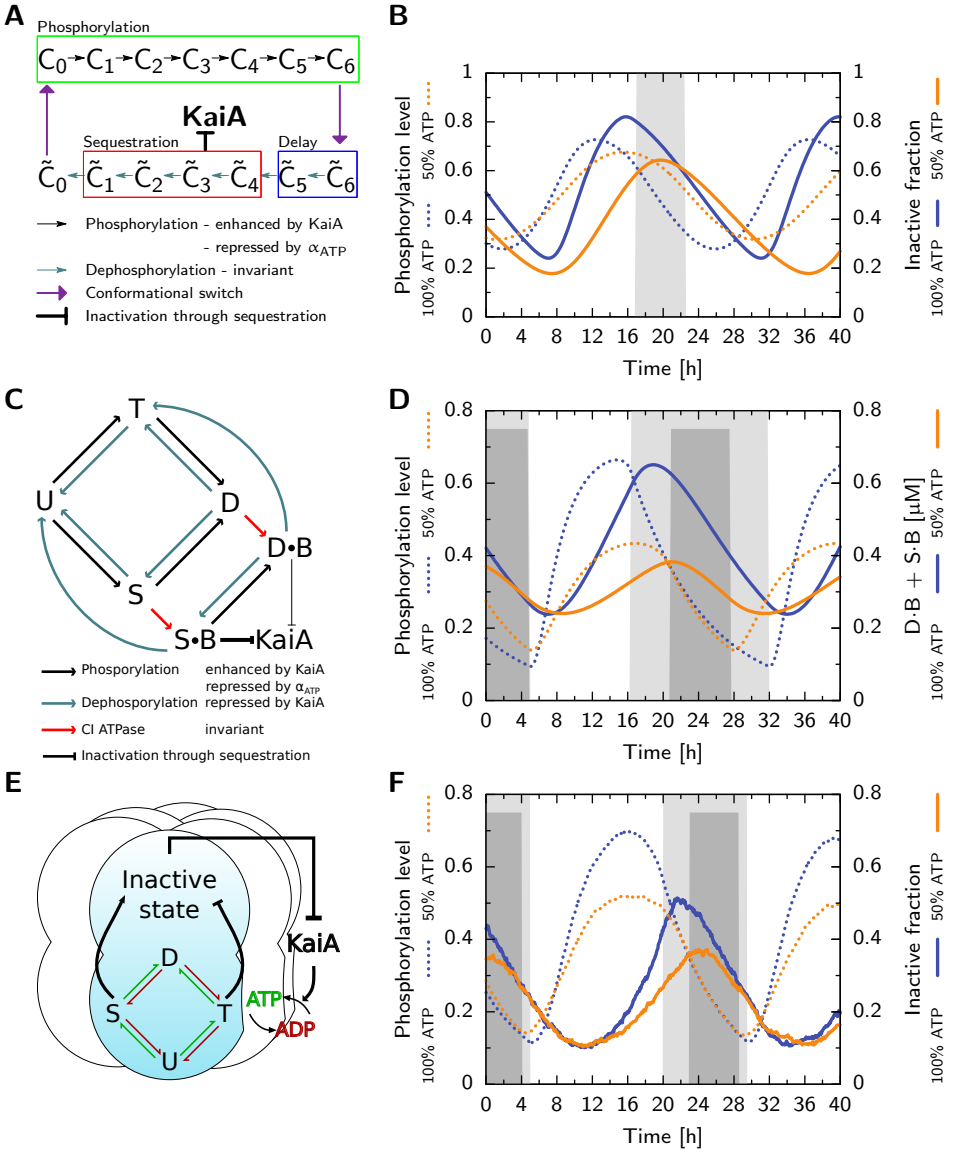

Figure 5.1: The different models of the Kai system employ different mechanisms of period stabilization. (A,B): Van Zon model; (C,D): Rust model (from [26]); (E,F): new model. Phosphorylation level (dotted lines,B,D,F) and fraction of KaiC in inactive state (solid lines,B,F) or in S.B + D.B (D, solid lines) at α_{ATP}=100% (blue) and 50% (orange). Shaded regions indicate the phase where all KaiA is sequestered, at α_{ATP}=100% (gray) and 50% (dark gray).

(continued next page)

Figure 5.1: (*continuing from previous page*) (A) Van Zon model: Hexamers phosphorylate in the active phase (green box), with a rate that depends on α_{ATP} and the amount of free KaiA, switch to the inactive state where they dephosphorylate and, after a delay (blue box), sequester KaiA (red box). (B) Both the phosphorylation level and the fraction of inactive hexamers rise slower at α_{ATP}=50%, but due to the delay between phosphorylation and KaiA sequestration, the dephosphorylation phase starts at a lower maximum of the inactive fraction, which shortens the time window where all KaiA is sequestered. (C) Rust model: Monomer is predominantly phosphorylated in the order U,T,D and S. Phosphorylation depends on the ATP fraction and KaiA (black arrows), dephosphorylation only on KaiA (gray arrows). The delay is set by the slow KaiB binding step (red arrows), and mutual inhibition between KaiA and the sequestration state S·B. (D) Because at α_{ATP}=50% the rate from S to D is smaller, the ensemble starts the dephosphorylation phase at a lower concentration of KaiB-bound KaiC. This shortens the dephosphorylation phase, which compensates for slower phosphorylation. (E) New model: In each hexamer, monomers go through the ordered phosphorylation cycle, where T stabilizes the active state, and S the inactive state. Phosphorylation can only occur with ATP (green arrows) and dephosphorylation only with ADP (red arrows) in the CII binding pocket. (F) The inactive fraction (solid lines) increases slower at α_{ATP}=50%, but reaches a lower maximum, leading to a shorter dephosphorylation phase. However, the effect on the duration of the KaiA sequestration phase is not as strong as in the Rust model. Because in the new model the switch to the inactive state depends not on the absolute number of serine and threonine phosphorylated residues, but rather on their difference, the decrease in the phosphorylation level (dotted lines) has a smaller impact on the inactive fraction (solid lines), as compared to the impact on the S.B+D.B fraction in the Rust model (panel D).

5.3. RESULTS

5.3.1. DEPENDENCE OF OSCILLATIONS ON THE ATP FRACTION REVEALS INPUT COMPENSATION ON THE ENSEMBLE LEVEL

To find out how effective the mechanisms for period stability are in the three models, we run simulations of the models at constant bulk ATP fractions from 100% to 50%. For the Van Zon and Rust model, we consider two values of the relative dissociation constant for ATP versus ADP: 1) With equal affinity for ATP and ADP, $K_{\mathrm{ATP/ADP}} = 1.0$ (solid lines in Fig. 5.1), as used and motivated in the original model by Rust [7], and 2) With a lower affinity for ADP, $K_{\mathrm{ATP/ADP}} = 0.19$ (dashed lines), such that the effect on the ATP fraction in nucleotide binding pocket, β_{ATP} (Eq. 5.1), as α_{ATP} decreases from 100% to 50%, is similar to the drop in our new model of about 15%.

Fig. 5.1A shows how the period varies with decreasing α_{ATP}. Remarkably, the three models have a different response to lowering α_{ATP}: Whereas in our new model the period is almost constant, in the Van Zon model it increases by 20%, while in the Rust model it decreases by 20%, as α_{ATP} decreases from 100% to 50%. This is reflected in the change in the amplitude of the phosphorylation levels, panel B, which decreases the strongest in the Rust model and the least in the Van Zon model, with decreasing α_{ATP}. The results in panels A and B are consistent: given the period stabilization mechanism due to the delay, we expect that in all models fewer hexamers or monomers make it through the cycle as α_{ATP} decreases, such that the amplitude of the oscillation decreases, shortening the length of the dephosphorylation phase. This view is further supported in panels C and D:

Panel C shows the length of the phosphorylation phase, $\Delta t_{\text{phosphorylation}}$, defined as the time between a trough and the next peak in phosphorylation level. As expected, this time increases in all models as α_{ATP} decreases: due to the lowering of the phosphorylation rates, it takes more time to reach the required phosphorylation state to sequester enough KaiA. The length of the subsequent dephosphorylation phase, $\Delta t_{\text{dephosphorylation}}$, defined as the time between a peak and the next trough in the phosphorylation level, decreases in the new model and the Rust model, as a result of the stability mechanism. However, in the Van Zon model, the stability mechanism does not work as $\Delta t_{\text{dephosphorylation}}$ increases with decreasing α_{ATP}. Because KaiC is unable to sequester all KaiA in the system during the dephosphorylation phase, the phosphorylation of active hexamers continues during this phase, which decreases the *net* dephosphorylation rate and extends its duration.

We point out that as the phosphorylation rates decrease, the number of hexamers and monomers that traverse a full cycle each period, and sequester KaiA, decreases. To show the effect of α_{ATP} on this fraction, we plot the total flux of monomers or hexamers that move through the cycle per period, described in more detail in section 5.5, for each model in Fig. 5.1E. Panel E shows a decrease in flux of around 15% for the Van Zon and our new model at $\alpha_{\text{ATP}} = 50\%$. The Rust model has a much larger decrease in flux of around 60%, which confirms the idea that input compensation is achieved in this model by letting fewer monomers participate in the cycle and sequester KaiA per period. Panel F gives the time interval per period when all KaiA is sequestered by KaiC, $\Delta t_{\text{KaiA sequestered}}$, as defined in section 5.5, which shows that in all models the time of full sequestration indeed shortens. Here, remarkably, the decreases upon lowering α_{ATP} is the strongest for the Van Zon model, even though the decrease in the fraction of hexamers that go through a cycle and sequester KaiA (panel E) is much less. This is probably related to the fact that in the Van Zon model a hexamer can only sequester 2 KaiA dimers, while in our model 6 dimers per hexamer and in the Rust model 2.5 dimers per monomer are sequestered. Therefore, the number of inactive hexamers required to sequester all KaiA, is much higher in the Van Zon model, compared to the other two models. Hence, a small change in the amplitude in the concentration of inactive hexamers, has a big effect on $\Delta t_{\text{KaiA sequestered}}$. Note that below α_{ATP} of 65%, KaiA is never fully sequestered in the Van Zon model, while oscillations persist.

Comparing results between different relative affinities, $K_{\text{ATP/ADP}}$, it is clear that both the Van Zon and the Rust model are much less affected by the bulk ATP fraction, when $K_{\text{ATP/ADP}}$ is lower. For lower $K_{\text{ATP/ADP}}$, the probability that KaiC is bound to ATP instead of ADP, is much less affected by changes in α_{ATP}. This makes the phosphorylation rates, and hence the oscillations, less sensitive to α_{ATP}. However, a lower sensitivity might hamper the entrainability of the clocks. Therefore, in the next section, we study how each model of the Kai oscillator responds to a transient lowering of the ATP fraction.

5.3.2. VAN ZON AND RUST MODELS ONLY SHOW STRONG ENTRAINABILITY AT EQUAL NUCLEOTIDE AFFINITIES

We expose the oscillator, running at a 100% ATP fraction, to a transient, six hour pulse of a 40% ATP fraction, starting at different times from the last trough in the phosphorylation fraction. To calculate the phase shift induced, we compare the time trace of the

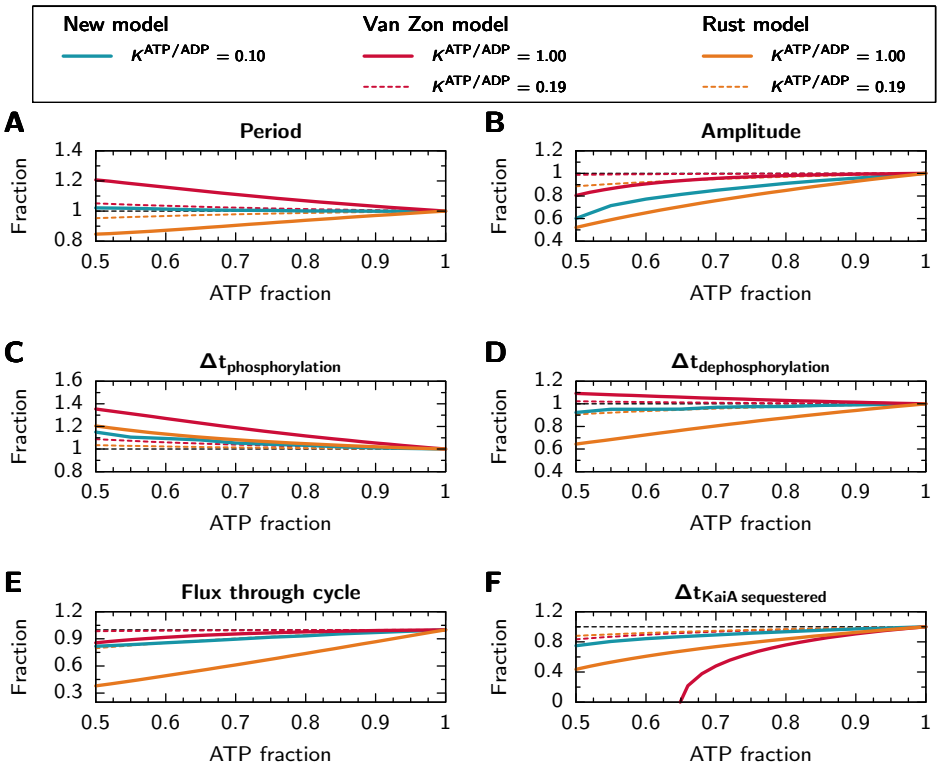

Figure 5.1: Dependence of the new (blue solid lines), Van Zon (red) and the Rust (orange) models on the bulk ATP fraction, α_{ATP}. All curves are normalized by their values at 100% ATP fraction, indicated by the horizontal dashed line. For the Van Zon and the Rust models, we study versions with equal sensitivity for ATP and ADP, $K_{ATP/ADP}$=1.0 (solid lines), and a smaller sensitivity for ADP, $K_{ATP/ADP}$=0.19 (dashed lines). (A) The ATP fraction has an opposite effect on the period in the Van Zon and our new model compared to the Rust model, (B) but the amplitude increases with increasing α_{ATP} in all models. (C) The time between a through and the next peak in the phosphorylation level, $\Delta t_{phosphorylation}$, decreases with α_{ATP} in all models, because the phosphorylation rates increase. (D) In both our new model and the Rust model, the time between a peak and the next trough, $\Delta t_{dephosphorylation}$, increases with α_{ATP}, which compensates for the faster phosphorylation rates. In the Van Zon model however, the dephosphorylation time interval decreases with α_{ATP} such that the compensation for a stable period does not occur. (E) As the ATP fraction increases, the number of hexamers or monomers that go through a full cycle per period increases, such that the amount of KaiA that can be sequestered increases. (F) Therefore, the time interval per period that all KaiA in the system is sequestered by KaiC, becomes longer with increasing α_{ATP}.

phosphorylation fraction two troughs after the onset of the pulse, with a control where no pulse is given.

Fig. 5.2, panels A, B and C, show phase response curves for our new model, the Van Zon model and the Rust model, respectively, showing the induced phase advances.

Comparing our new model, with $K_{\mathrm{ATP/ADP}} = 0.1$, with the other two models with equal nucleotide affinities, $K_{\mathrm{ATP/ADP}} = 1.0$ (solid lines), we see that the phase response curves are comparable: Both the maximally induced phase advance and delays are around 4 hours in all models, although the Rust model is capable of a particularly large phase advance of almost 10 hours. All models have a dead zone, where the pulse does not induce a phase change, starting about 12 hours from the trough, which is thought to be essential for entrainment [6]. However, when we reduce in the Van Zon and Rust models the relative affinity to $K_{\mathrm{ATP/ADP}} = 0.19$ (dashed lines), the amplitude of the phase response curves become much smaller: In the Van Zon model the maximal phase advance and delay are reduced to around one hour. In the Rust model, the maximal phase delay is also reduced to about one hour, which is significantly smaller than that in the new model; the maximal phase advance is also strongly reduced, although it is reduced to a value that is only slightly lower than that in the new model. Comparing Fig. 5.2A-C with Fig. 5.1A shows that the new model gives the best trade-off between entrainability and input compensation: On the footing of equal phase response curves, Fig. 5.2A-C, the new model has the smallest change in the period upon changing α_{ATP}, Fig. 5.1A.

We want to know how the phase advance and delay is achieved in each model. To this end, in panels D-I, we look at time traces of the phosphorylation level (solid black lines) and the fraction of sequestered KaiA (solid purple lines), when a pulse of ADP is given (shaded regions), where we set $K_{\mathrm{ATP/ADP}} = 1.0$ for the Van Zon and Rust models. For comparison, the dashed lines show the control where no pulse is given. We choose the starting times of the pulses such that a maximum phase advance (panels D-F) or phase delay (G-I) is induced, as indicated by the arrows in the phase response curves, panels A-C.

When the pulse is given at the moment that leads to a maximum phase delay, then, in all three models, the ADP pulse only slows down the increase of the phosphorylation level. When the ATP level returns to 100%, the phosphorylation level reaches the same peak height as compared to a situation where no pulse is given. Therefore, the subsequent dephosphorylation phase, when all KaiA is sequestered, has the same length as when no pulse is given. The ADP pulse thus predominantly slows down phosphorylation.

When the pulse is given at the moment that leads to a maximum phase advance, then, in all three models, the ADP pulse causes the immediate start of the dephosphorylation phase. Therefore, the peak in the phosphorylation level is lower compared to the unperturbed case, and the number of hexamers or monomers that is able to sequester KaiA is also smaller. The subsequent dephosphorylation phase shortens, causing the phase advance. Only in the Rust model, the ADP pulse also initiates the immediate sequestration of KaiA, because, due to the sudden drop in phosphorylation rate from the S to the D state, KaiA is incapable anymore to prevent its own sequestration.

5.3.3. AS THE ATP FRACTION DECREASES, INDIVIDUAL HEXAMERS GO THROUGH A SMALLER CYCLE IN PHOSPHORYLATION STATE SPACE

Up to now, the effect of ATP on the phosphorylation level has only been studied at the level of the mean phosphorylation level, both experimentally and theoretically. We want to know, for our new model, the effect of the ATP fraction on the phosphorylation dy-

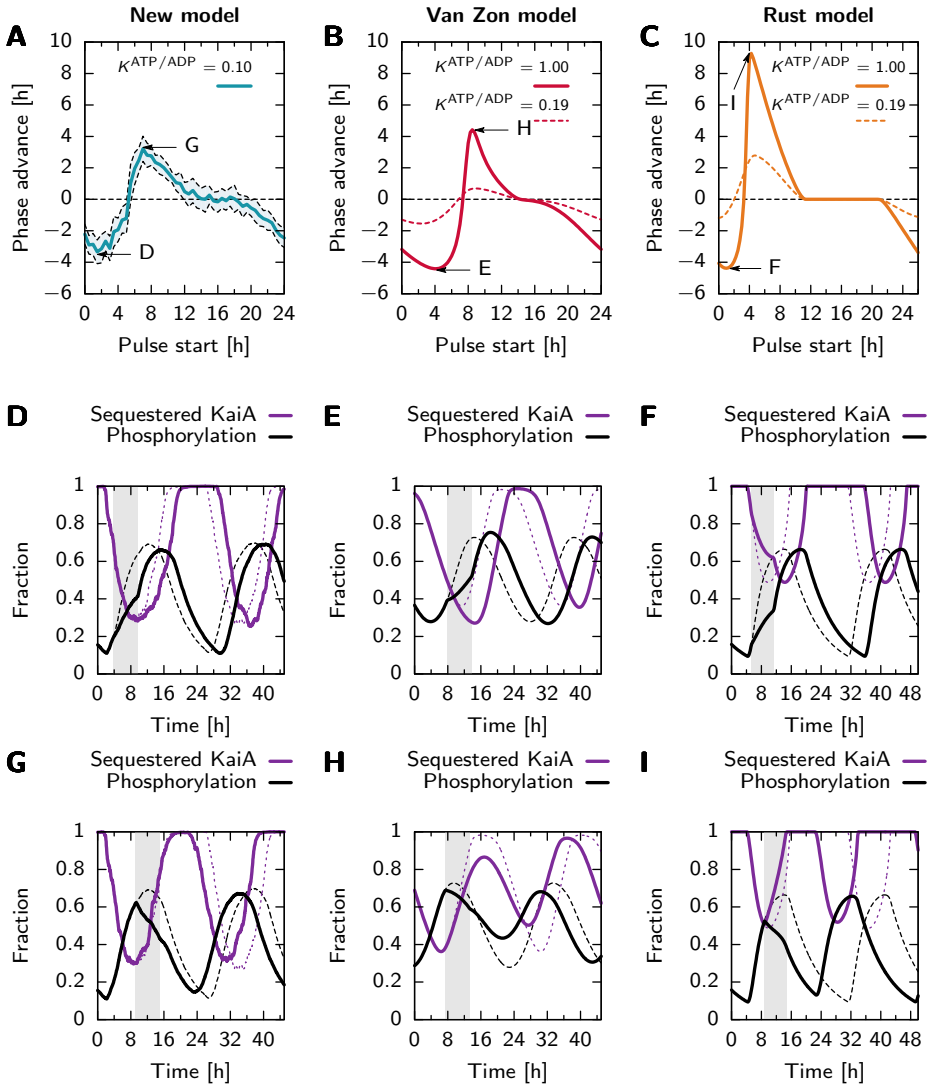

Figure 5.2: Sensitivity of the phase of the phosphorylation level, in our new model (left column), Van Zon model (middle column) and the Rust model (right column), to a 6 hour pulse of lowering of the ATP fraction to 40%. Panels A, B and C show the phase advance in hours, due to the pulse, starting at the time indicated on the x-axis, measured from the minimum in the phosphorylation level. The shaded region in panel A shows the standard deviation of 10 independent runs of the new stochastic model.

(continued next page)

Figure 5.2: (*continuing from previous page*) For the Van Zon and Rust models, we show phase response curves for a scenario with an equal relative affinity for ATP and ADP, $K^{\mathrm{ATP/ADP}} = 1.0$ (solid line), and a version with a lower affinity for ADP, $K^{\mathrm{ATP/ADP}} = 0.19$ (dashed line). Arrows indicate extrema in the phase response curves, for which we show the corresponding time traces in the panels with the adjacent label. Panels D,E and F show the effect of a pulse (shaded region) when the phase delay is the largest, on the phosphorylation level (solid black line) and the fraction sequestered KaiA (solid purple line). Dashed lines show the development in the case no pulse is given. Panels G,H and I show the effect of the pulse when the phase advance is the largest. Results for the new model were obtained using kinetic Monte Carlo, and ODE's for the Van Zon and Rust models.

namics of individual hexamers. To this end, for each hexamer, we track in time the number of phosphorylated threonine sites, n_{T}, and serine sites, n_{S} , as it moves through its phosphorylation cycle.

In Fig. 5.3, we show the distribution of hexamers in phosphorylation state space, $P_{\mathrm{n_T,n_S}}$, and the fluxes between states, at different phases of the oscillation. Panel A shows, as indicated by the arrows, the time intervals at which we take data during the oscillation, and panels B-F show the state of the ensemble during these intervals, at α_{ATP}=100%. Panels G-L show the same, but now for α_{ATP}=50%. Comparing the most populated phosphorylation states in panels D and J, taken at the peak of the phosphorylation level, it is clear that more threonine and serine sites are phosphorylated at 100% ATP level than at 50%. Thus, higher phosphorylation rates result in wider path through phosphorylation state space for individual hexamers. This is a clear signature of another mechanism of period robustness: at higher phosphorylation rates, corresponding to higher α_{ATP}, the rate at which a system goes through state space increases, yet the length of one cycle also increases [149].

5.3.4. MICROSCOPIC CYCLES ILLUMINATE INPUT COMPENSATION IN INDIVIDUAL HEXAMERS

Given our observation that individual hexamers make smaller cycles in phosphorylation state space when α_{ATP} decreases, we wanted to know how this affects the timing for switching between the active and inactive conformation. This timing is an important factor in the period of the oscillation, because it determines when and how long KaiA is sequestered. To this end, we measure the time between two important events during a full cycle of a hexamer, illustrated in Fig. 5.4A: 1) The start of the phosphorylation cycle, when a hexamer is in the active state and, for the first time, either a threonine site or serine site is phosphorylated; 2) and half-way of the cycle, when the hexamer is in the inactive state and is bound to six KaiB monomers. For all hexamers, we track the first passage time from the start to reaching half-way, $\Delta t_{\mathrm{active}}$, and from half-way to the start of a new cycle, $\Delta t_{\mathrm{inactive}}$. The time of a full cycle is defined as $\Delta t_{\mathrm{cycle}} = \Delta t_{\mathrm{active}} + \Delta t_{\mathrm{inactive}}$.

Note that the time in the active phase, $\Delta t_{\mathrm{active}}$, not only includes the phosphorylation and the switching to the inactive state, but also includes the binding of KaiB monomers. We include the binding of KaiB in this state definition, because KaiC recrosses the dividing surface that separates the active from the inactive state many times, before it is

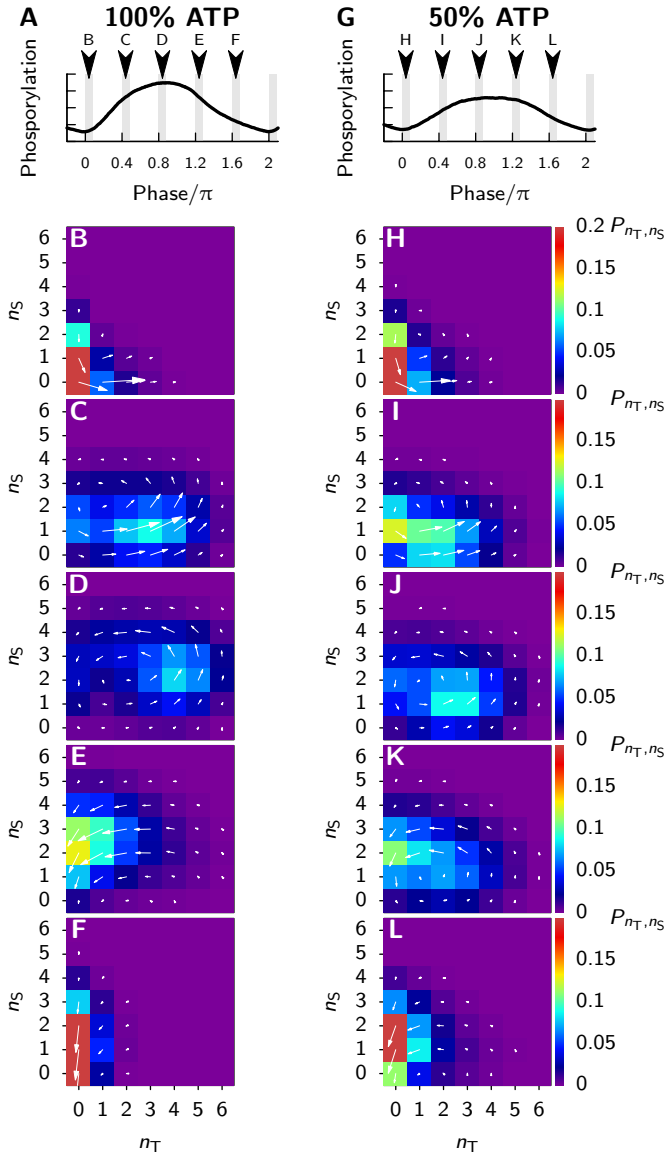

Figure 5.3: In the new model, at a lower bulk ATP fraction, individual hexamers make a smaller cycle through phosphorylation state space. We track the number of phosphorylated threonine sites, n_T, and serine sites, n_S, in each individual hexamer for one simulated hour at different phases of the oscillation. In panels A and G, the shaded regions indicate at which phase of the oscillation hexamers were probed and in which panel the result is shown.

(continued next page)

Figure 5.3: (*continuing from previous page*) Panels B-F (α_{ATP}=100%, left column) and H-L (α_{ATP}=50%, right column) show histograms of the probability of finding a hexamer in a certain phosphorylation state, indicated by the color bar to the right of the right column. Arrows are proportional to the flux through the state the arrow originates from. Comparing states near the peak of the phosphorylation level, panels D and J, hexamers at α_{ATP}=50% go through a smaller cycle compared to the situation at 100% ATP. Specifically, the majority of hexamers go through n_{T}=4 or 5 at α_{ATP}=100%, while at 50% hexamers only reach n_{T}=2 or 3. Furthermore, the ensemble is less synchronized near the trough of the phosphorylation levels, comparing panels F and L. Results shown are averaged over 400 consecutive oscillation cycles.

finally committed to the inactive state. Including the binding of KaiB into our criterion does not affect the results, because the rate of KaiB binding is independent of α_{ATP}, and any change in $\Delta t_{\mathrm{active}}$ due to a different α_{ATP} is therefore related to changes in phosphorylation rates.

In Fig. 5.4B we show histograms of individual cycle times, $\Delta t_{\mathrm{cycle}}$, at 100% ATP (solid blue line) and 50% (dashed orange line). The distribution has maxima at multiples of the period of the oscillation, indicated by dashed vertical lines. Clearly, the cycle times of individual hexamers coincide with the period of the oscillation. Peaks at multiples of $\Delta t_{\mathrm{cycle}}$ correspond to hexamers that do not bind 6 KaiB monomers during the first period, and hence have to wait for another round, or more, to make the full cycle. Since the histogram for 50% ATP has a fatter tail, hexamers are more likely to wait multiple periods before completing the cycle, showing that indeed fewer hexamers participate in an oscillation at lower ATP fractions. Fig. 5.4C shows the distributions of $\Delta t_{\mathrm{active}}$, at 100% and 50% ATP. The distribution again has multiple peaks, mirroring those in the distribution of cycle times (panel B). This indicates that at lower α_{ATP}, synchronization becomes impaired because fewer hexamers make it to the top of the cycle, where they have 6 KaiB bound and are committed to the inactive state.

The inset of Fig. 5.4C zooms in on the first peak of the distribution, emphasizing that even though the phosphorylation rates are different, the modes of the first passage time distribution are remarkably similar; the difference is only 1-2 hours. This paradox can be resolved by noting that the switch from the active to inactive state is determined by the difference between the number of phosphorylated serine and threonine sites, $n_{\mathrm{S}} - n_{\mathrm{T}}$, respectively. In our model, the phosphotransfer rates for the threonine site are much faster than for the serine site (see Table 4.1 and [22]). Therefore, compared to the slow phosphorylation of the serine sites, n_{T} will quickly reach its steady-state level during the phosphorylation phase. The steady-state level of n_{T} will thus set the number of serine sites that need to be phosphorylated before the hexamer can switch to the inactive state. The steady-state level of n_{T} decreases as the ATP fraction of the buffer is reduced, because the rate of phosphorylation decreases and the rate of dephosphorylation increases (in the presence of KaiA) with lower α_{ATP}, respectively. Consequently, as α_{ATP} is decreased, less serine sites need to be phosphorylated for the hexamer to switch conformation, which compensates for the lower rate of phosphorylation. This reasoning implies that the levels n_{S} and n_{T}, at which a hexamer switches to the inactive state, also decreases. This can be seen in panel E and F, which show the distribution of phosphory-

lation states (n_S,n_T), at the moment a hexamer flips from active to inactive, at α_{ATP}=100% and 50%, respectively. Clearly, n_S and n_T tend to be lower at the moments of switching, when α_{ATP}=50%. This also suggests that at lower α_{ATP}, fewer monomers are double phosphorylated. Fig. 5.4G shows histograms of the phosphorylation states of the hexamers at the moment when all KaiA is sequestered in the system for the first time during a period. At α_{ATP} = 50%, 40% of the hexamers have one or more monomers in the D state, while at α_{ATP}=50%, this fraction is reduced to 10%.

Lastly, the distribution of inactive times, $P(\Delta t_{inactive})$ in Fig. 5.4D, can be explained. The distribution exhibits a shoulder at α_{ATP}=100%, which is due to the fact that: 1) The number of hexamers that are in the inactive state is higher, reflected by the higher first peak in Fig. 5.4B, resulting in a longer time where all KaiA is sequestered. Therefore, the time hexamers have to wait before another round of phosphorylation starts increases, which is included in $\Delta t_{inactive}$; 2) Hexamers start their inactive phase at a higher phosphorylation level, comparing panels E and F in Fig. 5.4, which results in a longer dephosphorylation phase.

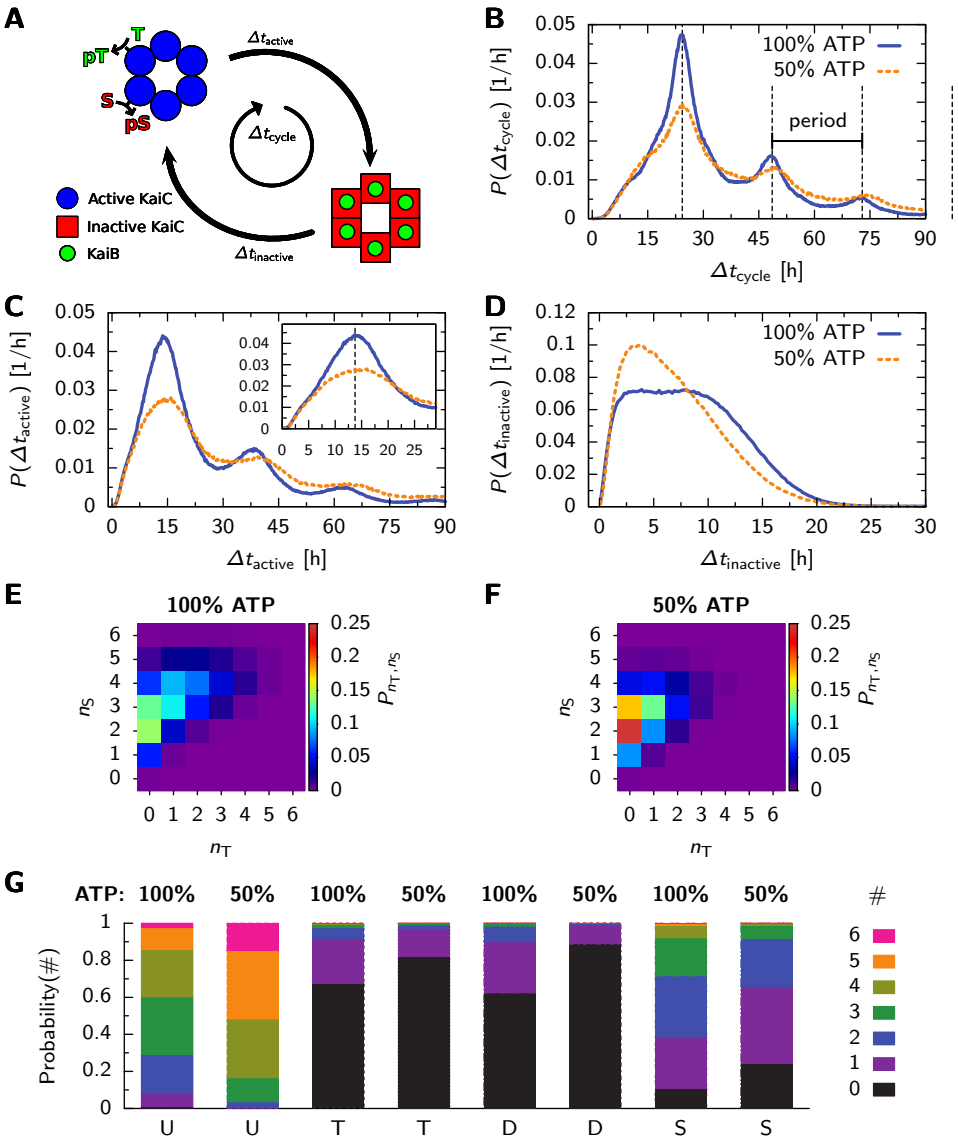

Figure 5.4: Microscopic dynamics of the new model reveals input compensation at the level of individual hexamers. (A) For each hexamer we measure, 1) Δt_{active}: The time between the first phosphorylation event, and having six KaiB bound to the hexamer for the first time. 2) $\Delta t_{\text{inactive}}$: The time between having 6 KaiB bound and, after dephosphorylation and switching back to the active state, the first phosphorylation event. The time for completing a full cycle is: $\Delta t_{\text{cycle}} = \Delta t_{\text{active}} + \Delta t_{\text{inactive}}$.

(continued next page)

Figure 5.4: (*continuing from previous page*) B, C and D show histograms of these time intervals, comparing situations with 100% ATP (blue solid lines) and 50% ATP in the bulk (orange dashed lines). (B) Distribution of times for completing a cycle. Peaks are at multiples of the oscillator's period of 24.3 hrs, as indicated by the vertical dotted lines (period at 100% ATP). Peaks at $\Delta t_{\text{cycle}} > 24.3$hrs show hexamers that could not complete a full cycle during one period of the oscillation. (C) Distribution of times for phosphorylation and KaiB binding. Event though phosphorylation rates are lower at α_{ATP}=50% as compared to 100%, the modes of their distributions are remarkably similar, as emphasized by the inset which zooms in on the first peak. (D) Distribution of times for dephosphorylation and waiting for a new round of phosphorylation. The bigger shoulder at α_{ATP}=100% is a manifestation of the longer time during which all KaiA is sequestered at this ATP level. At the end of their cycle, hexamers have to wait longer before KaiA returns to solution and phosphorylation starts again. (E,F) Histograms of the number of phosphorylated threonine sites, n_{T}, and serine sites, n_{S}, in individual hexamers, at the moment when they switch to the inactive state, at α_{ATP}=100% (E) and 50% (F). At higher α_{ATP}, hexamers switch to the inactive state at higher phosphorylation levels. (G) Histograms of the number (different colors, defined on the right side) of U,T,D and S phosphorylated monomers inside a hexamer, at the moment when all KaiA is sequestered. At α_{ATP}=100% more monomers are doubly phosphorylated.

5.4. DISCUSSION

All circadian clocks have to fulfill two seemingly conflicting requirements in order to be a good predictor of time: A robust circadian period under a wide range of external conditions, and entrainability such that it always moves in phase with the day-night cycle. As shown in recent experiments, the daily change in ATP fraction in the cyanobacterium *Synechococcus elongatus*, is an important cue for entrainment of its circadian Kai oscillator. We compared two canonical models of the post-translational oscillator, the hexamer model by Van Zon *et. al.* and the monomer model by Rust *et. al.*, with our new model, and studied how well they fulfill the robustness and entrainability criteria.

We find, in agreement with experiments [26], that the period in our new model is almost unaffected by the bulk ATP fraction, and that its hard to determine from the mean quantities related to the phosphorylation level what sets this period stability. Apart from the amplitude in the phosphorylation level, other quantities such as the length of the phosphorylation and dephosphorylation phase, do not change much with the bulk ATP fraction in our new model. The other two models show clear signatures of input compensation. The Van Zon model has too little input compensation, however, as shortening the dephosphorylation phase by sequestering less KaiA is unable to compensate slower phosphorylation. On the other hand, the Rust model has too much input compensation, since the period decreases with lower ATP. Lowering the phosphorylation rates has a too strong effect on the positive feedback loop regulating KaiA sequestration, shortening the period too much. The fact that our new model is so stable, is perhaps not so surprising, as the sensitivity to the ADP level was a factor 10 lower in our model compared to the other two.

We then checked the entrainability of the respective models, and found that the entrainability of our model is comparable to that of the other models, even though the relative affinity of ADP versus ATP is lower in our model. When we make the relative ADP/ATP sensitivity in the Van Zon and Rust models similar to that of our new model, the amplitudes of the phase response curves in these other models become very small, leading to poor entrainability. This showed that our new model, contrary to previous models, is capable of maintaining a robust circadian period, while at the same time being strongly entrainable.

To elucidate how our model achieves this combination of period robustness and good entrainability, we studied the phosphorylation cycle of the threonine and serine sites in individual hexamers, at 100% and 50% bulk ATP fractions. This showed that at 50% ATP, when effective phosphorylation rates are lower, individual hexamers go through a smaller phosphorylation cycle. Our analysis also revealed that the distribution of times for hexamers to complete a full cycle, peaks at multiples of the period, and that peaks at times higher than the period become more pronounced at lower ATP fraction, as fewer hexamers make it trough the full cycle each period. Remarkably, the time required to complete the first part of the cycle, between when phosphorylation starts and when 6 KaiB monomers are bound, seems to be little affected by the bulk ATP fraction. This leads to the question of how, at different effective phosphorylation rates, the timing of the conformational switch can be almost unaffected.

To address this question, we made histograms of the phosphorylation states of the hexamers when they switch to the inactive state. At lower phosphorylation rates, hex-

amers switch to the inactive state at a lower phosphorylation level, compensating the longer time they need to phosphorylate the sites. This is in marked contrast with previous models by Van Zon *et. al.* and Rust *et. al.*, where the time required for individual units (be it hexamers or monomers) to complete a full cycle, can only increase with decreasing ATP fractions. In these models, period stability can only be achieved through the interaction *between* the units via KaiA sequestration [149], that is, via the delay between the moment where KaiC no longer needs KaiA to progress along the cycle and the point where KaiC sequesters KaiA: At lower α_{ATP}, less KaiC makes it to the top of the cycle during the delay, making the dephosphorylation phase shorter, which can then counteract the longer phosphorylation phase [26].

The new model presented here exploits this mechanism too, yet also employs another one, which acts at the level of the individual hexamers. This mechanism utilizes the ordered phosphorylation of the threonine and serine site in the monomers, in combination with their antagonistic effect on the conformation of the hexamer. Previously, Lin *et. al.* argued that this antagonism creates an ultra-sensitive switch which provides robustness against a varying KaiA concentration [18]. However, his antagonism also plays a key role in our mechanism of input compensation. The antagonism entails that the state at which the hexamer switches conformation does not depend on the absolute number of phosphorylated threonine and serine residues, but rather only on their relative amounts. When the ATP fraction is low, their are typically fewer serine and threonine sites phosphorylated (see Fig. 4.1 and [26]), but since the switch to the inactive conformation does not depend on the absolute number of phosphorylated sites, but on their relative amounts, hexamers switch conformation at a lower number of phosphorylated sites. The net result is that, although the cycle in state space is smaller at lower ATP fraction, the moment a hexamer switches is robust against changes in the ATP fraction. And, since the phosphorylation rates *a*re sensitive to changes in the fraction, the hexamer still undergoes a phase shift, essential for entrainability. We thus propose a new function for the ordered phosphorylation cycle of the KaiC monomers: It allows the oscillator to combine period robustness with high entrainability.

5.5. METHODS

We use the Rust model described the SI of [26] and the Van Zon model in the SI of [15], both described with ordinary differential equations (ODE's), propagated using the NDSolve function of Mathematica 8 (Wolfram Research). Our new model, introduced in chapter 4, is propagated using the dedicated Monte Carlo algorithm described in the same chapter.

We want to compare the Van Zon and Rust models, using a relative affinity for ATP versus ADP, such that the change of the ATP fraction in the nucleotide binding pockets given that KaiA is bound, due to changes in α_{ATP}, is similar to the change in our new model. In our new model, the steady state fraction of ATP in the CII binding pocket, $\beta_{\mathrm{ATP}}^{\mathrm{CII}}$, given that KaiA is always bound to CII, is given by (section 4.3.2)

$$\beta_{\mathrm{ATP}}^{\mathrm{CII}} = \frac{\alpha_{\mathrm{ATP}}\, k_{\mathrm{off,KaiA}}^{\mathrm{CII\cdot ADP}}}{\alpha_{\mathrm{ATP}}\, k_{\mathrm{off,KaiA}}^{\mathrm{CII\cdot ADP}} + k_{\mathrm{hyd}}^{\mathrm{CII}} + (1-\alpha_{\mathrm{ATP}})\, k_{\mathrm{off,KaiA}}^{\mathrm{CII\cdot ADP}} K_{\mathrm{ATP/ADP}}^{\mathrm{CII}}}. \tag{5.2}$$

Here, $k_{\text{off,KaiA}}^{\text{CII·ADP}}$ and $k_{\text{hyd}}^{\text{CII}}$ are the dissociation rate of ADP when KaiA is bound and the hydrolysis rate of ATP, in the CII domain of KaiC, respectively. Using the parameters presented in Table 4.1, we find that $\beta_{\text{ATP}}^{\text{CII}}$ decreases with 15% as we lower α_{ATP} from 100% to 50% in Eq. 5.2. The ATP fraction in β_{ATP}, Eq. 5.1, has a similar scaling when we set $K_{\text{ATP/ADP}} = 0.19$.

5.5.1. CALCULATING MEAN QUANTITIES, FIG. 5.1 AND FIG. 5.2

Results shown in Fig. 5.1 were taken after 10 oscillations, such that the system has reached steady state oscillations. For the new model, presented quantities are averages over 400 consecutive oscillations. The period is defined as the mean of the peak-to-peak time in the phosphorylation level.

To calculate the flux trough a cycle, Fig. 5.1E, we require a reaction that has to take place at least once, in order for a hexamer or monomer to complete a full cycle. Here, a full cycle is defined as a series of states the hexamer or monomer has to go through in order to be able to sequester KaiA. For the Rust model, we calculate the flux between the U and T phosphorylation states, integrated over a period P:

$$Q_{\text{Rust}} = \frac{1}{P}\int_0^P \left(k_{\text{UT}}(t)\,U(t) - k_{\text{TU}}(t)\,T(t)\right) dt, \tag{5.3}$$

where k_{UT} and k_{TU} are the time dependent (depend on the free KaiA concentration) rates for phosphorylation and dephosphorylation of the T state, and $U(t)$ and $T(t)$ are the concentrations of U and T phosphorylated monomers. For the Van Zon model, we calculate the flux between switching to the inactive state

$$Q_{\text{VanZon}} = \frac{1}{P}\int_0^P \left(k_{\text{fw}}\,\text{C}_6(t) - k_{\text{bw}}\,\tilde{\text{C}}_6(t)\right) dt, \tag{5.4}$$

where k_{fw} and k_{bw} are the rate constants for switching to the inactive or active state, respectively. In our new model we measure the flux by counting the number of hexamers that have six KaiB monomers bound at some point during the period, averaged over 400 oscillations.

The time interval in which all KaiA is sequestered, Fig. 5.1F, is defined as the time interval when more than 99% of all KaiA dimers are sequestered by KaiC.

To generate the phase response curves (PRC) shown in Fig. 5.2A-C, we applied a 6 hour step-wise decrease in the ATP fraction. For the ODE models, we derived the induced phase change by comparing the time of second through in the phosphorylation level after the onset of the pulse, with the same trough of the control where no pulse is given. For our new model, which contains stochasticity, we fit a sinusoidal function to three oscillations in the time trace of the phosphorylation level, one oscillation after the pulse is given. We compare this with a fit to the phosphorylation level in a control simulation, where no pulse is given, and calculate the phase shift between the fits. To suppress the intrinsic number fluctuations present in the Monte Carlo simulations, we used a simulation volume of 6 cubic microns, which is three times as large as the original volume. Finally, to get the PRC in panel A, we averaged over 10 different trajectories for each pulse start. The shaded region in Fig. 5.2A indicates the standard deviation in the phase shift of these 10 runs.

5.5.2. MICROSCOPIC DYNAMICS OF THE HEXAMER, FIG. 5.3 AND FIG. 5.4

To get the snapshots of the distribution of hexamers in phosphorylation state space shown in Fig. 5.3, we first need to register the time of the troughs in the phosphorylation level, t_i^{tr}, for a time trace containing over 400 oscillations. This allows us to calculate the trough-to-trough time, $T_i^{tr} = t_{i+1}^{tr} - t_i^{tr}$, for each oscillation i. Then we define time intervals of one hour for the whole trajectory, $(t_i^{tr} + \gamma\, T_i^{tr}, t_i^{tr} + \gamma\, T_i^{tr} + 1.0)$, where $\gamma \in [0,1)$ sets the phase of the period at which the time intervals start. During these intervals we keep track of the number of phosphorylated threonine sites, n_T, and serine sites, n_S, in each hexamer. This allows us to calculate the occupancy of each phosphorylation state, P_{n_T,n_S}, and the fluxes between these states, during the one hour time window. These quantities are described in more detail in section 4.5.3.

BIBLIOGRAPHY

[1] Johnson CH, Egli M, Stewart PL (2008) Structural insights into a circadian oscillator. *Science* 322:697–701.

[2] Elowitz MB, Levine AJ, Siggia ED (2002) Stochastic gene expression in a single cell. *Science* 297:1183–1186.

[3] Mori T, Binder B, Johnson CH (1996) Circadian gating of cell division in cyanobacteria growing with average doubling times of less than 24 hours. *Proc Natl Acad Sci* 93:10183–10188.

[4] Walker N, Nghe P, Tans SJ (2016) Generation and filtering of gene expression noise by the bacterial cell cycle. *BMC Biol* 14:11.

[5] Kondo T, et al. (1993) Circadian rhythms in prokaryotes : Luciferase as a reporter of circadian gene expression in cyanobacteria. *Proc Natl Acad Sci* 90:5672–5676.

[6] Pfeuty B, Thommen Q, Lefranc M (2011) Robust entrainment of circadian oscillators requires specific phase response curves. *Biophys J* 100:2557–2565.

[7] Rust MJ, Golden SS, O'Shea EK (2011) Light-driven changes in energy metabolism directly entrain the cyanobacterial circadian oscillator. *Science* 331:220–223.

[8] DeMairan J (1729) *Histoire de l'Academie Royale des Sciences* (Academie Royale des Sciences, Paris), pp 35–36.

[9] Konopka RJ, Benzer S (1971) Clock mutants of Drosophila melanogaster. *Proc Natl Acad Sci* 68:2112–2116.

[10] Ishiura M, et al. (1998) Expression of a gene cluster kaiABC as a circadian feedback process in cyanobacteria. *Science* 281:1519–1523.

[11] Nishiwaki T, et al. (2004) Role of KaiC phosphorylation in the circadian clock system of Synechococcus elongatus PCC 7942. *Proc Natl Acad Sci* 101:13927–13932.

[12] Nakajima M, et al. (2005) Reconstitution of circadian oscillation of cyanobacterial KaiC phosphorylation in vitro. *Science* 308:414–415.

[13] Bialek W (2015) Perspectives at the interface of physics and biology. *arXiv* 1512.08954.

[14] Emberly E, Wingreen NS (2006) Hourglass model for a protein-based circadian oscillator. *Phys Rev Lett* 96:038303.

[15] van Zon JS, Lubensky DK, Altena PRH, ten Wolde PR (2007) An allosteric model of circadian KaiC phosphorylation. *Proc Natl Acad Sci* 104:7420–7425.

[16] Eguchi K, Yoda M, Terada TP, Sasai M (2008) Mechanism of robust circadian oscillation of KaiC phosphorylation in vitro. *Biophys J* 95:1773–1784.

[17] Yoda M, Eguchi K, Terada TP, Sasai M (2007) Monomer-shuffling and allosteric transition in KaiC circadian oscillation. *PLoS One* 5:1–8.

[18] Lin J, Chew J, Chockanathan U, Rust MJ (2014) Mixtures of opposing phosphorylations within hexamers precisely time feedback in the cyanobacterial circadian clock. *Proc Natl Acad Sci* 111:E3937–E3945.

[19] Zwicker D, Lubensky DK, ten Wolde PR (2010) Robust circadian clocks from coupled protein- modification and transcription-translation cycles. *Proc Natl Acad Sci* 107:22540–22545.

[20] Teng SW, Mukherji S, Moffitt JR, de Buyl S, O'Shea EK (2013) Robust circadian oscillations in growing cyanobacteria require transcriptional feedback. *Science* 340:737–740.

[21] Paijmans J, Bosman M, ten Wolde PR, Lubensky DK (2016) Discrete gene replication events drive coupling between the cell cycle and circadian clocks. *Proc Natl Acad Sci* 113:4063–4068.

[22] Rust MJ, Markson JS, Lane WS, Fisher DS, O'Shea EK (2007) Ordered phosphorylation governs oscillation of a three-protein circadian clock. *Science* 318:809–812.

[23] Tsai TYC, et al. (2008) Robust, tunable biological oscillations from interlinked positive and negative feedback loops. *Science* 321:126–129.

[24] Li C, Chen X, Wang P, Wang W (2009) Circadian KaiC phosphorylation : A multi-Layer network. *PLoS Comput Biol* 5:e1000568.

[25] Nagai T, Terada TP, Sasai M (2010) Synchronization of circadian oscillation of phosphorylation level of KaiC in vitro. *Biophys J* 98:2469–2477.

[26] Phong C, Markson JS, Wilhoite CM, Rust MJ (2013) Robust and tunable circadian rhythms from differentially sensitive catalytic domains. *Proc Natl Acad Sci* 110:1124–1129.

[27] Chang Yg, et al. (2015) A protein fold switch joins the circadian oscillator to clock output in cyanobacteria. *Science* 349:324–328.

[28] Elowitz MB, Leibler S (2000) A synthetic oscillatory network of transcriptional regulators. *Nature* 403:335–338.

[29] Stricker J, et al. (2008) A fast, robust and tunable synthetic gene oscillator. *Nature* 456:516–520.

[30] Pattanayak GK, Lambert G, Bernat K, Rust MJ (2015) Controlling the cyanobacterial clock by synthetically rewiring metabolism. *Cell Rep* 13:2362–2367.

[31] Salter MG, Franklin KA, Whitelam GC (2003) Gating of the rapid shade-avoidance response by the circadian clock in plants. *Nature* 426:680–683.

[32] Johnson C (2010) Circadian clocks and cell division: What's the pacemaker? *Cell Cycle* 9:3864–3873.

[33] Sotak M, Sumova A, Pacha J (2014) Cross-talk between the circadian clock and the cell cycle in cancer. *Ann Med* 46:221–232.

[34] Matsuo T, et al. (2003) Control mechanism of the circadian clock for timing of cell division in vivo. *Science* 302:255–259N.

[35] Nagoshi E, et al. (2004) Circadian gene expression in individual fibroblasts: Cell-autonomous and self-sustained oscillators pass time to daughter cells. *Cell* 119:693–705.

[36] Dong G, et al. (2010) Elevated ATPase activity of KaiC applies a circadian checkpoint on cell division in Synechococcus elongatus. *Cell* 140:529–539.

[37] Yang Q, Pando BF, Dong G, Golden SS, van Oudenaarden A (2010) Circadian gating of the cell cycle revealed in single cyanobacterial cells. *Science* 327:1522–1526.

[38] Gerard C, Goldbeter A (2012) Entrainment of the mammalian cell cycle by the circadian clock: Modeling two coupled cellular rhythms. *PLoS Comp Biol* 8:e1002516.

[39] Feillet C, et al. (2014) Phase locking and multiple oscillating attractors for the coupled mammalian clock and cell cycle. *Proc Natl Acad Sci* 111:9828–9833.

[40] Bieler J, et al. (2014) Robust synchronization of coupled circadian and cell cycle oscillators in single mammalian cells. *Mol Syst Biol* 10:739.

[41] Bosman M (2012) Master's thesis (Universiteit van Amsterdam).

[42] Volfson D, et al. (2006) Origins of extrinsic variability in eukaryotic gene expression. *Nature* 439:861–864.

[43] Irish VF, Gelbart WM (1987) The decapentaplegic gene is required for dorsal-ventral patterning of the Drosophila embryo. *Genes Dev* 1:868–879.

[44] Trcek T, Larson DR, Moldon A, Query CC, Singer RH (2011) Single-molecule mRNA decay measurements reveal promoter-regulated mRNA stability in yeast. *Cell* 147:1484–1497.

[45] Zopf CJ, Quinn K, Zeidman J, Maheshri N (2013) Cell-Cycle dependence of transcription dominates noise in gene expression. *PLoS Comp Biol* 9:e1003161.

[46] Narula J, et al. (2015) Chromosomal arrangement of phosphorelay genes couples sporulation and DNA replication. *Cell* 162:328–337.

[47] Hensel Z, Marquez-Lago TT (2015) Cell-cycle-synchronized, oscillatory expression of a negatively autoregulated gene in E. coli. *arXiv* 1506.08596.

[48] Pikovsky A, Rosenblum M, Kurths J (2003) *Synchronisation: A universal concept in nonlinear sciences* (Cambridge University Press, Cambridge).

[49] Mori T, Johnson CH (2001) Independence of circadian timing from cell division in cyanobacteria. *J. Bacteriol.* 183:2439–2444.

[50] Kageyama H, et al. (2006) Cyanobacterial circadian pacemaker: Kai protein complex dynamics in the KaiC phosphorylation cycle in vitro. *Mol Cell* 23:161–171.

[51] Clodong S, et al. (2007) Functioning and robustness of a bacterial circadian clock. *Mol Syst Biol* 3:90.

[52] Brettschneider C, et al. (2010) A sequestration feedback determines dynamics and temperature entrainment of the KaiABC circadian clock. *Mol Syst Biol* 6:1–10.

[53] Qin X, Byrne M, Xu Y, Mori T, Johnson CH (2010) Coupling of a core post-translational pacemaker to a slave transcription/translation feedback loop in a circadian system. *PLoS Biol* 8:e1000394.

[54] Takai N, et al. (2006) A KaiC-associating SasA-RpaA two-component regulatory system as a major circadian timing mediator in cyanobacteria. *Proc Natl Acad Sci* 103:12109–12114.

[55] Taniguchi Y, Takai N, Katayama M, Kondo T, Oyama T (2010) Three major output pathways from the KaiABC-based oscillator cooperate to generate robust circadian kaiBC expression in cyanobacteria. *Proc Natl Acad Sci* 107:3263–3268.

[56] Gutu A, O'Shea EK (2013) Two antagonistic clock-regulated histidine kinases time the activation of circadian gene expression. *Mol Cell* 50:288–294.

[57] Markson JS, Piechura JR, Puszynska AM, O'Shea EK (2013) Circadian control of global gene expression by the cyanobacterial master regulator RpaA. *Cell* 155:1396–1408.

[58] Jain IH, Vijayan V, O'Shea EK (2012) Spatial ordering of chromosomes enhances the fidelity of chromosome partitioning in cyanobacteria. *Proc Natl Acad Sci* 109:13638–13643.

[59] Lengeler JW, Drews G, Schlegel HG (1999) *Biology of the Prokaryotes* (Georg Thieme Verlag), p 481.

[60] Nishiwaki T, et al. (2007) A sequential program of dual phosphorylation of KaiC as a basis for circadian rhythm in cyanobacteria. *EMBO J* 26:4029–4037.

[61] Goldbeter A, Koshland DE (1981) An amplified sensitivity arising from covalent modification in biological systems. *Proc Natl Acad Sci* 78:6840–6844.

[62] Binder BJ, Chisholm SW (1990) Relationship between DNA cycle and growth rate in Synechococcus sp. strain PCC 6301. *J Bact* 172:2313–2319.

[63] Griese M, Lange C, Soppa J (2011) Ploidy in cyanobacteria. *FEMS Micr Lett* 323:124–131.

[64] Chen AH, Afonso B, Silver Pa, Savage DF (2012) Spatial and temporal organization of chromosome duplication and segregation in the cyanobacterium Synechococcus elongatus PCC 7942. *PLoS ONE* 7:e47837.

[65] Watanabe S, et al. (2012) Light-dependent and asynchronous replication of cyanobacterial multi-copy chromosomes. *Mol Microbiol* 83:856–865.

[66] Kollmann M, Løvdok L, Bartholomé K, Timmer J, Sourjik V (2005) Design principles of a bacterial signalling network. *Nature* 438:504–507.

[67] Shultzaberger RK, Boyd JS, Katsuki T, Golden SS, Greenspan RJ (2014) Single mutations in sasA enable a simpler ΔcikA gene network architecture with equivalent circadian properties. *Proc Natl Acad Sci* 111:E5069–E5075.

[68] Masri S, Cervantes M, Sassone-Corsi P (2013) The circadian clock and cell cycle: Interconnected biological circuits. *Curr Opin Cell Biol* 25:730–734.

[69] Ito H, et al. (2009) Cyanobacterial daily life with Kai-based circadian and diurnal genome-wide transcriptional control in Synechococcus elongatus. *Proc Natl Acad Sci* 106:14168–14173.

[70] Tomita J, Nakajima M, Kondo T, Iwasaki H (2005) No transcription-translation feedback in circadian rhythm of KaiC phosphorylation. *Science* 307:251–254.

[71] Delaune EA, Francois P, Shih NP, Amacher SL (2012) Single-cell-resolution imaging of the impact of Notch signaling and mitosis on segmentation clock dynamics. *Dev Cell* 23:995–1005.

[72] Bird AJ, Turner-Cavet JS, Lakey JH, Robinson NJ (1998) A carboxyl-terminal Cys2/His2-type zinc-finger motif in DNA primase influences DNA content in Synechococcus PCC 7942. *J Biol Chem* 273:21246–52.

[73] Kitayama Y, Nishiwaki T, Terauchi K, Kondo T (2008) Dual KaiC-based oscillations constitute the circadian system of cyanobacteria. *Genes Dev* 22:1513–1521.

[74] Collins JJ, Gardner TS, Cantor CR (2000) Construction of a genetic toggle switch in Escherichia coli. *Nature* 403:339–342.

[75] Becskei a, Serrano L (2000) Engineering stability in gene networks by autoregulation. *Nature* 405:590–593.

[76] Ernesto A, Basu S, Karig DK, Weiss R (2006) Synthetic biology: New engineering rules for an emerging discipline. *Mol Syst Biol* 2:1–14.

[77] Tigges M, Marquez-Lago TT, Stelling J, Fussenegger M (2009) A tunable synthetic mammalian oscillator. *Nature* 457:309–312.

[78] Nandagopal N, Elowitz MB (2011) Synthetic Biology: Integrated Gene Circuits. *Science* 333:1244–1248.

[79] Sowa SW, Gelderman G, Contreras LM (2015) Advances in synthetic dynamic circuits design: Using novel synthetic parts to engineer new generations of gene oscillations. *Curr Opin Biotechnol* 36:161–167.

[80] Elowitz MB, Hsing W, Leibler S (2002) Combinatorial synthesis of genetic networks. *Science* 296:1466–1470.

[81] Rosenfeld N, Young JW, Alon U, Swain PS, Elowitz MB (2005) Gene regulation at the single-cell level. *Science* 307:1962–1965.

[82] Cai L, Friedman N, Xie XS (2006) Stochastic protein expression in individual cells at the single molecule level. *Nature* 440:358–362.

[83] Chabot JR, Pedraza JM, Luitel P, van Oudenaarden A (2007) Stochastic gene expression out-of-steady-state in the cyanobacterial circadian clock. *Nature* 450:1249–1252.

[84] Novak B, Tyson JJ (2008) Design principles of biochemical oscillators. *Nat Rev Mol Cell Biol* 9:981–991.

[85] Mather W, Bennett MR, Hasty J, Tsimring LS (2009) Delay-induced degrade-and-fire oscillations in small genetic circuits. *Phys Rev Lett* 068105:1–4.

[86] Woods ML, Leon M, Perez-Carrasco R, Barnes CP (2016) A statistical approach reveals designs for the most robust stochastic gene oscillators. *ACS Synth Biol* 5:459–470.

[87] Garcia-Ojalvo J, Elowitz MB, Strogatz SH (2004) Modeling a synthetic multicellular clock: repressilators coupled by quorum sensing. *Proc Natl Acad Sci* 101:10955–10960.

[88] Mondragón-Palomino O, Danino T, Selimkhanov J, Tsimring L, Hasty J (2011) Entrainment of a population of synthetic genetic oscillators. *Science* 333:1315–1319.

[89] Prindle A, et al. (2014) Rapid and tunable post-translational coupling of genetic circuits. *Nature* 508:387–391.

[90] Kang B, Li YY, Chang X, Liu L, Li YX (2008) Modeling the effects of cell cycle M-phase transcriptional inhibition on circadian oscillation. *PLoS Comput Biol* 4:e1000019.

[91] Gonze D (2013) Modeling the effect of cell division on genetic oscillators. *J Theor Biol* 325:22–33.

[92] Dies M, Galera-Laporta L, Garcia-Ojalvo J (2015) Mutual regulation causes co-entrainment between a synthetic oscillator and the bacterial cell cycle. *Integr Biol*.

[93] Cooper S, Helmstetter CE (1968) Chromosome replication and the division cycle of Escherichia coli Br. *J Mol Biol* 31:519–540.

[94] Donachie W (1968) Relationship between cell size and time of initiation of DNA replication. *Nature* 219:1077–1079.

[95] Wallden M, Fange D, Gregorsson Lundius E, Baltekin Ö, Elf J (2016) The synchronization of replication and division cycles in individual E. coli cells. *Cell* 166:729–739.

[96] Adiciptaningrum A, Osella M, Moolman MC, Cosentino Lagomarsino M, Tans SJ (2015) Stochasticity and homeostasis in the E. coli replication and division cycle. *Sci Rep* 5:18261.

[97] Koppes LJH, Meyer M, Oonk HB, de Jong MA, Nanninga N (1980) Correlation between size and age at different events in the cell division cycle of Escherichia coli. *J Bacteriol* 143:1241–1252.

[98] Michelsen O, Teixeira de MJ, Jensen PR, Hansen FG (2003) Precise determinations of C and D periods by flow cytometry in Escherichia coli K-12 and B/r. *Microbiology* 149:1001–1010.

[99] Bogan JA, et al. (2001) P1 and NR1 plasmid replication during the cell cycle of Escherichia coli. *Plasmid* 45:200–208.

[100] Mori T, Johnson CH (2001) Independence of circadian timing from cell division in cyanobacteria independence of circadian timing from cell division in cyanobacteria. *J Bacteriol* 183:2439–2444.

[101] Hurley JM, Loros JJ, Dunlap JC (2016) Circadian oscillators: Around the transcription–translation feedback loop and on to output. *Trends Biochem Sci* 41:834–846.

[102] Xu Y, Mori T, Johnson CH (2000) Circadian clock-protein expression in cyanobacteria: Rhythms and phase setting. *EMBO J* 19:3349–3357.

[103] Nakahira Y, et al. (2004) Global gene repression by KaiC as a master process of prokaryotic circadian system. *Proc Natl Acad Sci* 101:881–885.

[104] Mehra A, et al. (2006) Circadian rhythmicity by autocatalysis. *PLoS Comput Biol* 2:1–8.

[105] Clodong S, et al. (2007) Functioning and robustness of a bacterial circadian clock. *Mol Syst Biol* 3:1–9.

[106] Mori T, et al. (2007) Elucidating the ticking of an in vitro circadian clockwork. *PLoS Biol* 5:0841–0853.

[107] Miyoshi, Fumihiko and Nakayama, Yoichi and Kaizu, Kazunari and Iwasaki, Hideo and Tomita M (2007) A mathematical model for the Kai-protein–based chemical oscillator and cock gene expression rhythms in cyanobacteria. *J Biol Rhythms* 22:69–80.

[108] Markson JS, O'Shea EK (2009) The molecular clockwork of a protein-based circadian oscillator. *FEBS Lett* 583:3938–3947.

[109] Williams S, Vakonakis I (2002) Structure and function from the circadian clock protein KaiA of Synechococcus elongatus: A potential clock input mechanism. *Proc Natl Acad Sci* 99:15357–15362.

[110] Iwasaki H, Nishiwaki T, Kitayama Y, Nakajima M, Kondo T (2002) KaiA-stimulated KaiC phosphorylation in circadian timing loops in cyanobacteria. *Proc Natl Acad Sci* 99:15788–15793.

[111] Xu Y, Mori T, Johnson CH (2003) Cyanobacterial circadian clockwork: roles of KaiA, KaiB and the kaiBC promoter in regulating KaiC. *EMBO J* 22:2117–2126.

[112] Kitayama Y, Iwasaki H, Nishiwaki T, Kondo T (2003) KaiB functions as an attenuator of KaiC phosphorylation in the cyanobacterial circadian clock system. *EMBO J* 22:2127–2134.

[113] Pattanayek R, et al. (2008) Structural model of the circadian clock KaiB-KaiC complex and mechanism for modulation of KaiC phosphorylation. *EMBO J* 27:1767–1778.

[114] Terauchi K, et al. (2007) ATPase activity of KaiC determines the basic timing for circadian clock of cyanobacteria. *Proc Natl Acad Sci* 104:16377–16381.

[115] Murakami R, et al. (2008) ATPase activity and its temperature compensation of the cyanobacterial clock protein KaiC. *Genes Cells* 13:387–95.

[116] Nishiwaki T, Kondo T (2012) Circadian autodephosphorylation of cyanobacterial clock protein KaiC occurs via formation of ATP as intermediate. *J Biol Chem* 287:18030–18035.

[117] Egli M, et al. (2012) Dephosphorylation of the core clock protein KaiC in the cyanobacterial KaiABC circadian oscillator proceeds via an ATP synthase mechanism. *Biochemistry* 51:1547–1558.

[118] Mori T, et al. (2002) Circadian clock protein KaiC forms ATP-dependent hexameric rings and binds DNA. *Proc Natl Acad Sci* 99:17203–17208.

[119] Hayashi F, Suzuki H, Iwase R (2003) ATP induced hexameric ring structure of the cyanobacterial circadian clock protein KaiC. *Genes to Cells* 8:287–296.

[120] Hayashi F, et al. (2004) Roles of two ATPase-motif-containing domains in cyanobacterial circadian clock protein KaiC. *J Biol Chem* 279:52331–52337.

[121] Nishiwaki T, Iwasaki H, Ishiura M, Kondo T (2000) Nucleotide binding and autophosphorylation of the clock protein KaiC as a circadian timing process of cyanobacteria. *Proc Natl Acad Sci* 97:495–499.

[122] Chang Yg, Tseng R, Kuo NW, Liwang A (2012) Rhythmic ring-ring stacking drives the circadian oscillator clockwise. *Proc Natl Acad Sci* 109:16847–16851.

[123] Kim YI, Dong G, Carruthers CW, Golden SS, LiWang A (2008) The day/night switch in KaiC, a central oscillator component of the circadian clock of cyanobacteria. *Proc Natl Acad Sci* 105:12825–12830.

[124] Chang Yg, Kuo Nw, Tseng R, Liwang A (2011) Flexibility of the C-terminal , or CII , ring of KaiC governs the rhythm of the circadian clock of cyanobacteria. *Proc Natl Acad Sci* 108:14431–14436.

[125] Egli M, et al. (2013) Loop-loop interactions regulate KaiA-stimulated KaiC phosphorylation in the cyanobacterial KaiABC circadian clock. *Biochemistry* 52:1208–1220.

[126] Snijder J, et al. (2014) Insight into cyanobacterial circadian timing from structural details of the KaiB-KaiC interaction. *Proc Natl Acad Sci* 111:1379–1384.

[127] Nishiwaki-Ohkawa T, Kitayama Y, Ochiai E, Kondo T (2014) Exchange of ADP with ATP in the CII ATPase domain promotes autophosphorylation of cyanobacterial clock protein KaiC. *Proc Natl Acad Sci* 111:4455–4460.

[128] Monod J, Wyman J, Changeux JP (1965) On the nature of allosteric transitions: A plausible model. *J Mol Biol* 12:88–118.

[129] Kageyama H, Kondo T, Iwasaki H (2003) Circadian formation of clock protein complexes by KaiA, KaiB, KaiC, and SasA in cyanobacteria. *J Biol Chem* 278:2388–2395.

[130] Nakajima M, Ito H, Kondo T (2010) In vitro regulation of circadian phosphorylation rhythm of cyanobacterial clock protein KaiC by KaiA and KaiB. *FEBS Lett* 584:898–902.

[131] Pattanayek R, et al. (2006) Analysis of KaiA-KaiC protein interactions in the cyanobacterial circadian clock using hybrid structural methods. *EMBO J* 25:2017–2028.

[132] Kitayama Y, Nishiwaki-Ohkawa T, Sugisawa Y, Kondo T (2013) KaiC intersubunit communication facilitates robustness of circadian rhythms in cyanobacteria. *Nat Commun* 4:2897.

[133] Hayashi F, et al. (2004) Stoichiometric interactions between cyanobacterial clock proteins KaiA and KaiC. *Biochem Biophys Res Commun* 316:195–202.

[134] Ma L, Ranganathan R (2012) Quantifying the rhythm of KaiB-C interaction for in vitro cyanobacterial circadian clock. *PLoS One* 7:e42581.

[135] Boltzmann L (1872) Weitere Studien über das Wärmegleichgewicht unter Gasmolekülen. *Sitzungsberichte Akad Wiss* 66:275–370.

[136] Murayama Y, et al. (2011) Tracking and visualizing the circadian ticking of the cyanobacterial clock protein KaiC in solution. *EMBO J* 30:68–78.

[137] Villarreal Sa, et al. (2013) CryoEM and molecular dynamics of the circadian KaiB-KaiC complex indicates that KaiB monomers interact with KaiC and block ATP binding clefts. *J Mol Biol* 425:3311–3324.

[138] van Kampen NG (2007) *Stochastic Processes in Physics and Chemistry* (Elsevier, North-Holland), Third edition, p 463.

[139] Battle C, et al. (2016) Broken detailed balance at mesoscopic scales in active biological systems. *Science* 352:604–607.

[140] Oyama K, Azai C, Nakamura K, Tanaka S, Terauchi K (2016) Conversion between two conformational states of KaiC is induced by ATP hydrolysis as a trigger for cyanobacterial circadian oscillation. *Sci Rep* 6:1–11.

[141] Abe J, et al. (2015) Atomic-scale origins of slowness in the cyanobacterial circadian clock. *Science* 349:312–316.

[142] Viani MB, et al. (2000) Probing protein-protein interactions in real time. *Nat Struct Biol* 7:644–647.

[143] Gillespie DT (1977) Exact stochastic simulation of coupled chemical reactions. *J Phys Chem* 81:2340–2361.

[144] Gillespie DT (2007) Stochastic simulation of chemical kinetics. *Annu Rev Phys Chem* 58:35–55.

[145] Winfree AT (1980) *The geometry of Biological Time* (Springer, New-York), 2nd edition.

[146] Ouyang Y, Andersson CR, Kondo T, Golden SS, Johnson CH (1998) Resonating circadian clocks enhance fitness in cyanobacteria. *Proc Natl Acad Sci* 95:8660–4.

[147] Woelfle MA, Ouyang Y, Phanvijhitsiri K, Johnson CH (2004) The adaptive value of circadian clocks: An experimental assessment in cyanobacteria. *Curr Biol* 14:1481–1486.

[148] Hatakeyama TS, Kaneko K (2012) Generic temperature compensation of biological clocks by autonomous regulation of catalyst concentration. *Proc Natl Acad Sci* 109:8109–8114.

[149] Hatakeyama TS, Kaneko K (2015) Reciprocity between robustness of period and plasticity of phase in biological clocks. *Phys Rev Lett* 115:1–5.

[150] Hasegawa Y, Arita M (2014) Optimal implementations for reliable circadian clocks. *Phys Rev Lett* 113:1–5.

[151] Kidd PB, Young MW, Siggia ED (2015) Temperature compensation and temperature sensation in the circadian clock. *Proc Natl Acad Sci* 112:E6284–E6292.

[152] Schmitz O, Katayama M, Williams SB, Kondo T, Golden SS (2000) CikA, a bacteriophytochrome that resets the cyanobacterial circadian clock. *Science* 289:765–768.

[153] Hosokawa N, Kushige H, Iwasaki H (2013) Attenuation of the posttranslational oscillator via transcription-translation feedback enhances circadian-phase shifts in Synechococcus. *Proc Natl Acad Sci* 110:14486–14491.

[154] Wood TL, et al. (2010) The KaiA protein of the cyanobacterial circadian oscillator is modulated by a redox-active cofactor. *Proc Natl Acad Sci* 107:5804–5809.

[155] Kim YI, Vinyard DJ, Ananyev GM, Dismukes GC, Golden SS (2012) Oxidized quinones signal onset of darkness directly to the cyanobacterial circadian oscillator. *Proc Natl Acad Sci* 109:17765–17769.

[156] Yoshida T, Murayama Y, Ito H, Kageyama H, Kondo T (2009) Nonparametric entrainment of the in vitro circadian phosphorylation rhythm of cyanobacterial KaiC by temperature cycle. *Proc Natl Acad Sci* 106:1648–1653.

[157] Pattanayak GK, Phong C, Rust MJ (2014) Rhythms in energy storage control the ability of the cyanobacterial circadian clock to reset. *Curr Biol* 24:1934–1938.

SUMMARY

The earth's rotation around its axis presents us with a challenge, for every day we have to live by the rhythm set by the rising and setting of the sun. Instead of just reacting to the changes in light and temperature, many organisms, like us, have evolved to foretell dusk and dawn and anticipate the impending changes. For instance, the leaves of young sunflower plants track the sun from east to west during the day and, at night, points them towards the east again well before sunrise. Remarkably, when these plants are placed under continuous light, the 24 hour movement of their leaves persists. This shows that the sunflower has an internal 24 hour rhythm, or 'circadian clock' (Latin: *circa* for "around" and *dies* "day"), that is independent of its environment.

I study the clock of *Synechococcus elongatus*, a freshwater cyanobacterium that depends on sunlight for its growth. I use this model organism because it has the simplest known clock in biology, which consists of only three different proteins (the building blocks of the cell), called KaiA, KaiB and KaiC. The central component, KaiC, forms a hexamer structure that contains twelve special sites to which a small phosphate molecule can bind. KaiA can add phosphate molecules to the sites of KaiC and KaiB negates KaiA, such that the bound phosphates fall off again. The antagonistic effects of KaiA and KaiB on KaiC together set up a 24 hour rhythm in the number of phosphate molecules that are bound to KaiC. KaiC thus functions as a register, from which the cyanobacterium can read of the time of the day. However, the speed at which the phosphate molecules are added and removed from KaiC varies over time and strongly depends on the temperature of the surrounding. Furthermore, a cyanobacterium can divide up to four times every 24 hours, whereby all the components that make up the clock are randomly divided between the two daughter cells. In light of these circumstances, it is extraordinarily that the clock is unaffected by the cell division of the cyanobacterium or the temperature of its environment; it always ticks with a constant 24 hour period.

In this thesis I set out to reveal the properties of the biological clock that are essential for having a robust period in the chaotic environment of the cell. To find these properties, I used the large body of experimental data that are available on the circadian clock of *S. elongatus* to design mathematical models of the clock and the cell cycle, which I then simulate using a computer. Computational modeling allows for a systematic study of all the things that can affect the clock's stability, such as the division time of the cell and its environment. Modeling also allows us to dissect each individual component of the circadian clockwork and investigate how it makes the clock more robust.

Starting in chapters 2 and 3, I investigate how circadian clocks and synthetic oscillators (a designed biological clock built into a cell) can have a robust period inside a growing and dividing cell. I find that it is not the division of the cell but rather the copying of its DNA, the blueprint of all the cell's building blocks, that can be detrimental to this robust period. As the cell grows, it has to double all of its components, including its DNA and clock proteins. Each time the DNA is copied, the production of new clock

components also doubles, which causes a periodic perturbation of the clock by the cell cycle. Just like when you push someone on a swing, by pushing at the right moment, you can change the period and the extent of the swing. Similarly, the rhythmic doubling of the production of clock proteins by the growing cell, can change the clocks period to, for example, twenty hours. I first used a simplified version of the circadian clock to show that a too simple designed clock will be completely enslaved by the cell cycle. Our simplified clock highlights why the clock of *S. elongatus* is far more complicated: A clock that can not maintain a constant period is a bad predicator of time. I find two design features that contribute to the clock's stability: First, the rhythm of adding and removing of phosphate molecules from the clock protein KaiC is almost unaffected by the production of new KaiC proteins, which stabilizes the clock against the production doubling. Furthermore, *S. elongatus* has multiple copies of its chromosome which are copied at random times during each cell cycle. In this case, when a chromosome is copied, the number does not increase from one to two but, for instance, from four to five, whereby the effect of a single copy event on the protein production rate is relatively smaller.

Subsequently, in chapter 4, I seek to understand what drives the ticking of the clock. In cells, the hydrolysis of the molecule ATP to ADP is the driving force behind many cellular processes. In the clock of *S. elongatus*, ATP is converted in both the CI and CII domain that together make up KaiC. ATP is used in the CII domain for the rhythmic addition and removal of the phosphate molecules while the function of the CI domain is less obvious. It is unclear which of the two domains is the primary driver of the clock. To find out, I used the large body of experimental data to make the most detailed model of the circadian clock of *S. elongatus* to date. I argue that the CI domain is the engine that makes the clock 'go round', and that the CII domain acts as a 'timer' regulating the hydrolysis in CI.

Finally, in chapter 5, I explore how the clock can fulfill two seemingly contradictory requirements: Its period should be unaffected by the daily changes in the environment, while the phase of the clock needs to be *entrained* by these daily changes in order for the clock to give the correct time of the day. The amount of ATP in the cell changes during the day with the amount of light the cyanobacterium receives. It has been recently shown that the rhythmic fluctuation in the ATP concentration is the primary input for entrainment, or *Zeitgeber*, of the clock. Using our new model, I discovered a new mechanism by which the clock can both have a robust 24 hour period and still be entrainable. Inside a cell of *S. elongatus* there are typically thousands of individual KaiC hexamers that all go through their individual cycles. Contrary to earlier models, the cycle that a KaiC complex has to go through is not fixed but depends on the ATP concentration inside the cell. When the ATP concentration is high, an individual KaiC will move faster through its cycle but this cycle will also be longer, such that the time required to traverse the whole cycle is constant. I show that this mechanism stabilizes the period of the clock against fluctuations in the ATP concentration, while still being sensitive to these fluctuations for entrainability.

To recapitulate, I used the well-studied circadian clock of *S. elongatus*, to reveal the mechanisms that are essential for having a robust clock inside a cell and found how a clock can be both robust and entrainable. In doing so, I have increased insight into the minute precision of circadian clocks despite the chaotic environment of the a living cell.

SAMENVATTING

Het dag- en nachtritme heeft een grote invloed op ons bestaan. Veel organismen, zoals mensen, reageren niet alleen op zonsopgang en ondergang maar kunnen deze ook voorspellen. De bladereren van jonge zonnebloemen bijvoorbeeld, volgen overdag de zon van oost naar west en worden 's nachts ruim voor zonsopkomst weer naar het oosten gericht. Zelfs wanneer zonnebloemen onafgebroken belicht worden blijft het 24 uurs ritme in de beweging van de bladeren doorgaan. Dit toont aan dat de zonnebloem een interm ritme van 24 uur heeft, of een circadiaanse klok (Latijn: *circa* voor "ongeveer" en *dies* "dag"), onafhankelijk van zijn omgeving.

Ik bestudeer de klok van *Synechococcus elongatus*, een zoetwater cyanobacterie die van zonlicht afhankelijk is voor zijn groei. De klok van dit model organisme is de simpelste bekende biologische klok en bestaat uit slechts drie eiwitten: KaiA, KaiB en KaiC. Het belangrijkste eiwit, KaiC, vormt een hexameer structuur waarin twaalf plekken zijn waaraan een fosfaat molecuul kan binden. KaiA voegt fosfaat moleculen toe aan deze plekken en door KaiB vallen de gebonden fosfaat moleculen er weer af. Dit antagonistisch effect van KaiA en KaiB vormt een 24 uurs ritme in het aantal gebonden fosfaat moleculen aan KaiC. KaiC werkt dus als een register, waarvan de cyanobacterie de tijd kan aflezen. Nu varieert de tijd om een fosfaat molecuul toe te voegen of the verwijderen van KaiC sterk per keer en hangt het ook af van de omgevingstemperatuur. Daarnaast kan een cyanobacterie zich vier keer per dag delen, waarbij alle componenten van de klok over de twee dochtercellen verdeeld worden. Het is bijzonder dat de klok in deze chaotische omstandigheden ten alle tijden zijn 24 uurs ritme, of periode, behoudt.

In dit proefschrift zal ik de eigenschappen van de biologische klok onthullen die essentieel zijn voor een robuuste periode in de chaos van de levende cel. Om deze eigenschappen te vinden gebruik ik de schat aan experimentele data over de klok van *S. elongatus* om een wiskundig model te ontwerpen van de klok en de celcyclus. Modelleren stelt ons in staat systematische de invloed van de delingstijd van de cel en veranderingen in de omgeving op de klok te bestuderen. Daarnaast kan ik met het model elk individueel component van de klok ontleden en kijken wat het effect is op de robuustheid van de klok.

In hoofdstuk 2 en 3 onderzoek ik hoe circadiaanse klokken en synthetische oscillatoren (een door mensen ontworpen biologische klok ingebouwd in een cel) een robuuste periode hebben in een groeiende en delende cel. Ik laat zien dat niet de deling van de cel maar het kopiëren van het DNA, de blauwdruk van de cel, een nadelig effect op het ritme van de klok heeft. Terwijl de cel groeit verdubbeld hij al zijn componenten, zoals het DNA en de klokeiwitten. Op het moment dat het DNA gekopieerd wordt verdubbeld ook de productiesnelheid van klokeiwitten. Dit zorgt voor een periodieke pertubatie van de klok door de celcyclus. Vergelijk het met als je iemand duwt op een schommel. Wanneer je op het juiste moment duwt kun je de periode en de slingerwijdte bepalen. Net als bij de schommel zorgt de verdubelling van de productie ervoor dat het ritme van de klok

verandert in, bijvoorbeeld, twintig uur. Eerst bestudeer ik een versimpelde versie van de circadiaanse klok om aan te tonen dat de periode van deze klok volledig bepaald wordt door de celcyclus. Daarmee laat onze versimpelde klok zien waarom deze in *S. elongatus* veel complexer is: Zonder klok met constante periode weet de cel niet meer hoe laat het is. Ik vind twee eigenschappen van de klok die bijdragen aan zijn stabilieit: Ten eerste is het ritmisch toevoegen en verwijderen van fosfaat moleculen onafhankelijk van veranderingen in de productie van klokeiwitten. Ten tweede heeft *S. elongatus* meerdere kopieën van zijn chromosoom, en deze worden gedurende de celcyclus op willekeurige momenten gekopieerd. In dit geval, als een bepaald chromosoom gekopieerd wordt, verandert het aatal niet van een naar twee, maar bijvoorbeeld van vier naar vijf. Het effect van de het kopiëren van een enkel chromosoom is dan relatief kleiner.

Vervolgens, in hoofdstuk 4, ga ik op zoek naar de drijvende kracht achter de klok. In cellen is de hydrolyse van het ATP molecuul in ADP de energiebron van actieve processen zoals transport en groei. In de klok van *S. elongatus* wordt ATP gebruikt in zowel het CI en CII domein van KaiC. ATP wordt in het CII domein gebruikt voor het ritmisch toevoegen en verwijderen van de fosfaat moleculen. De functie in het CI domein is minder duidelijk. Het is onduidelijk welk domein de primaire aandrijving van de klok is. Om hier achter te komen gebruiken we de schat aan experimentele data en ontwerpen het meest gedetaileerde model ooit van de klok van *S. elongatus*. Aan de hand van dit model beredeneer ik waarom het CI domein de 'motor' is die de klok aandrijft en dat het CII domein de 'klok' is die de hydrolyse in het CI domein reguleert.

In hoofdstuk 5 bestudeer ik hoe de klok van *S. elongatus* twee tegenstrijdige eisen in zich verenigt: De periode van de klok mag niet beïnvloed worden door veranderingen in de omgeving, terwijl de klok wel gelijk gezet moet worden met zonsopkomst. Dit betekend dat de fase van de klok wel gevoelig moet zijn voor de omgeving. De hoeveelheid ATP in de cel verandert gedurende de dag evenredig met het invallend licht. Recente experimenten hebben laten zien dat de rhytmische verandering in de hoeveelheid ATP de belangrijkste informatie, of *Zeitgeber*, voor de klok is om de juiste fase te houden. Aan de hand van mijn model heb ik een nieuw mechanisme ontdekt waardoor de klok een robuuste periode heeft terwijl de fase juist goed de omgeving volgt. In een cel van *S. elongatus* zijn typisch duizenden klokeiwitten aanwezig die allemaal door hun individuele cycli gaan van het aantal gebonden fosfaat moleculen. In tegenstelling tot eerdere modellen, hoeft een KaiC complex in mijn model niet een vaste cyclus te doorlopen, maar is de cyclus afhankelijk van de ATP concentratie. Als de ATP concentratie hoog is, doorloopt een KaiC complex typisch sneller zijn cyclus maar wordt deze ook langer. De benodigde tijd om de cyclus te doorlopen blijft hierdoor ongeveer constant. Ik laat zien dat door dit mechanisme de periode van de klok constant blijft bij fluctuaties in de ATP concentratie terwijl de fase nog steeds gevoelig is voor deze veranderingen.

Samengevat heb ik de klok van model organisme *S. elongatus* gebruikt om te laten zien welke mechanismen essentieel zijn voor een robuuste klok, in een groeiende en delende cel. Daarnaast laat ik zien hoe de klok altijd de juiste fase heeft. Dit onderzoek vergroot ons inzicht in hoe de klok kan functioneren in de chaotische omgeving in de levende cel.

ACKNOWLEDGMENTS

It is now almost six years ago, when I first walked into Pieter Rein's office to apply for an internship at AMOLF. Ever since that day, I have found AMOLF to be an outstanding environment to develop ideas with a lot of freedom and to make friends. I feel tremendous gratitude to the people I have met throughout these years.

First, I like to thank Pieter Rein for motivating me to develop and sharpen my ideas. Especially during the last months of my PhD, while frantically writing this thesis, your enthusiasm and insight in your student's psychology were most appreciated. Also, I admire how far you want to go to thoroughly understand my and your own ideas. The many times we discussed in your office until late, when the rest of the institute had become dark and silent, and first your wife and later your daughter would call, desperately asking for you to come home. I feel very grateful for being given the opportunity to spend six years in your group, and to learn from you how to become a researcher.

David, I thank you for being such a critical thinker and your devotion to our projects. A manuscript that has been granted your approval can be send of to the hardest reviewer with utmost confidence. I also enjoyed the three months I got to spend in your research group in Ann Arbor, which allowed me to explore a bit of the US and motivates me to perhaps do a post-doc there.

Martijn, we started almost the same day and, by now, have become the '*mastodonten*' of AMOLF. Staying along my side throughout these years, and now as my paranymph, you have become a very good friend. I thank you for commenting on my work, revising my texts, your humor during the long evenings in the office and all the good moments during the many PhD sketches, trips and movies we experienced together.

Adithya, I thank you for your positive attitude and your capacity to cheer me up — always, for the many cool trips we made on our bikes in Norway and Czech and by car through Scotland and along the Syrian border. I admire your capacity to, contrary to myself, stay relaxed under all circumstances.

Noreen, you impressed me with your integrity and how sincere you approach the things that fascinate you such as your PhD and climbing.

Stephen, I thank you for all the nerdy games evenings we enjoyed together and the traditional American dinners you made.

Tomek, although it has been a while since you finished your PhD at AMOLF, I am very glad we stayed friends. While you live in Vienna now with Annegret and Kasimir, I still get a taste of your fascination for (historic) trivia, politics and Balkan beats through your eloquent emails.

Johannes, self-proclaimed only true liberal at AMOLF, I thank you for your advise in science and life matters during our countless coffee breaks and lunches at restaurant Pythagoras. Hilarious are the many times you position yourself in the middle of our office to give a treatise of your view on daily matters. I appreciate you are the only other person that still takes the etiquette of A.J. Groskamp-Ten Have serious.

Guus, you motivated me to start running and introduced me to a lot of good music and books; thank you. I am very glad we are friends.

I like to thank all my colleagues for the great time I had during my PhD. In particular, I praise Galja, for all your heartwarming greetings, Yuval, for your incapacity to make boring conversation, Mario, for sharing a love for 80's synthesizer extravagance, Olga, for making me realize it's actually crap, Rayan, for the surreal experience of a yoga class, Agata, for your razor sharp commentary, VIKTORIA, for carrying out a message that matches the magnificence of your name, Michele, for feeding me with old Chinese newspapers, Giulia, for introducing me to Italian cinema and food, Bart -Fuxy- Vos, for sharing a fascination for improvised German, Luc, for being a good running mate, Katharina, for challenging me with fencing, Simone, for being the pinnacle of human kindness, Ireth, for spreading Costa Rican cheerfulness, Anders, for being exceptional both in baking and with humor, Fatemeh, for your enlightening views on life, Mr and Mrs Addenbrooke, for letting me experience an English wedding, Mathijs, for never walking, and Nils, for being my scientific 'Doktoronkel'.

I thank my (former) group members Michele, Adithya, Yao, Giulia, Harmen, Maarten, Anouk, Annemarieke, Andrew, Rasa, Chris, Tom, Nils, Martijn, Laurens, Wiet and Filipe for all the helpful discussions and having to listen to my incoherent presentations during the group meetings.

I also like to thank the support staff for making AMOLF such a good place. In particular Jan and Lars for their Linux support, Marga for making coffee for the Wednesday group meeting and Jan and Mark of the workshop in helping out with PV and open-day activities.

Lieve Moeder, Vader, Marrigje en Willemijn, ik wil jullie bedanken dat jullie er altijd voor mij zijn. En Marrigje, heel erg bedankt voor het ontwerp van de omslag van mijn proefschrift.

LIST OF PUBLICATIONS

RELATED TO THIS THESIS

Chapter 2
Paijmans J, Bosman M, ten Wolde PR, Lubensky DK (2016) *Discrete gene replication events drive coupling between the cell cycle and circadian clocks.* Proc Natl Acad Sci USA, 113(15):4063–4068
http://doi.org/10.1073/pnas.1507291113

Chapter 3
Paijmans J, Lubensky DK, ten Wolde PR - *Robustness of synthetic oscillators in growing and dividing cells.*
Submitted to Physical Review E

Chapter 4
Paijmans J, Lubensky DK, ten Wolde PR - *A Thermodynamically consistent model of the post-translational Kai circadian clock.*
Submitted to PLoS Computational Biology

Chapter 5
Paijmans J, Lubensky DK, ten Wolde PR - *Period robustness and entrainability under changing nucleotide concentrations in the in-vitro Kai circadian clock.*
Manuscript in preparation

OTHER PUBLICATIONS

Kaizu K, de Ronde W, **Paijmans J**, Takahashi K, Tostevin F, ten Wolde PR (2014) *The berg-purcell limit revisited.* Biophys J, 106(4), 976–985.
http://doi.org/10.1016/j.bpj.2013.12.030

Paijmans J, ten Wolde PR (2014) *Lower bound on the precision of transcriptional regulation and why facilitated diffusion can reduce noise in gene expression.* Phys Rev E, 90(3), 032708.
http://doi.org/10.1103/PhysRevE.90.032708