Characterization and Analysis of Synthetic Genetic Clocks

by

Nicolas Perry

**A dissertation submitted in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy
(Biophysics)
in The University of Michigan
2011**

Doctoral Committee

Professor Alexander J. Ninfa, Co-Chair
Associate Professor Daniel B. Forger, Co-chair
Associate Professor Mark A. Saper
Assistant Professor Nina Lin
Assistant Professor David K. Lubensky

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

CHAPTER 1

INTRODUCTION: SYNTHETIC BIOLOGY, GENETIC CLOCKS AND THE NINFA CLOCK

Synthetic biology is a relatively new field, in which bacteria or other cells are genetically engineered to produce novel cell behaviors. These behaviors may include modified information processing, modified chemical processing, or modified production of various molecules. (Andrianantoandro et al, 2006; Benner and Sismour, 2005; Endy, 2005; Heinemann and Panke, 2006; Kaern et al, 2003). Synthetic biology studies are used for two distinct purposes: for the advancement of practical applications that directly benefit society, and for the development of model experimental systems that may be used to advance our basic understanding of network behaviors and their control. Although the field of synthetic biology is still in its infancy in terms of practical applications, current lines of research include the production of biofuels (Connor and Atsumi, 2010; Ghim et al, 2010; Clomburg and Gonzalez, 2010), drug discovery and production (Keasling, 2008; Weber and Fussenegger, 2009; Neumann and Neumann-Staubitz, 2010), and bioremediation (de Lorenzo 2008; Porcar and Moya, 2010). Numerous studies have used synthetic biology methods for the fabrication and analysis of model genetic networks; examples include the engineering of logic gates (Guet et al. 2002; Mayo et al, 2006; Anderson et al, 2007; Tamsir et al, 2010), two-dimensional patterns (Basu et al, 2004; Basu et al, 2005; Tabor et al, 2009), bistable gene networks (genetic toggle switches)

(Gardner et al, 2000; Tropynina et al, 2002; Kramer and Fussenegger, 2005; Chang et al, 2010) and oscillators (Atkinson et al, 2003; Elowitz and Leibler, 2000; Fung et al, 2005; Stricker et al, 2008; Tigges et al, 2009; Danino et al, 2010).

Construction of synthetic genetic clocks

The construction of genetic oscillators (clocks) has provides a minimal network capable of producing sustained oscillations, in order to provide insight into the inner workings of natural genetic clocks. This subject matter has been a major focus of the synthetic biology field owing to the importance of natural circadian clocks, their widespread occurance in nature, the complexity of natural clocks and difficulty in studying them, and the non-intuitive mechanisms involved in the generation of oscillations. Natural genetic clocks have been described and studied in detail in cyanobacteria (Golden et al, 1997; Golden and Canales, 2003), the fungus Neurospora (Loros, 1998; Loros and Dunlap, 2001), the fruit fly Drosophila (Helfrich-Forster, 1996; Scully and Kay, 2000), and mammals (Takahashi, 1995; King and Takahashi, 2000). Although these natural systems have been studied in detail, the structures of these natural networks are complex, and the mechanisms for the generations of oscillations remain elusive. Thus, studies with simpler, synthetic, networks might be able to elucidate design principles involved in the generation of oscillations.

In addition to providing model systems for study, synthetic clocks may also have practical applications in the future. In our studies, we introduce time-keeping capabilities to an organism (*E. coli*) that does not have a genetic clock. This

introduction of time-keeping capabilities may be useful in the future for controlling complex synthetic gene networks.

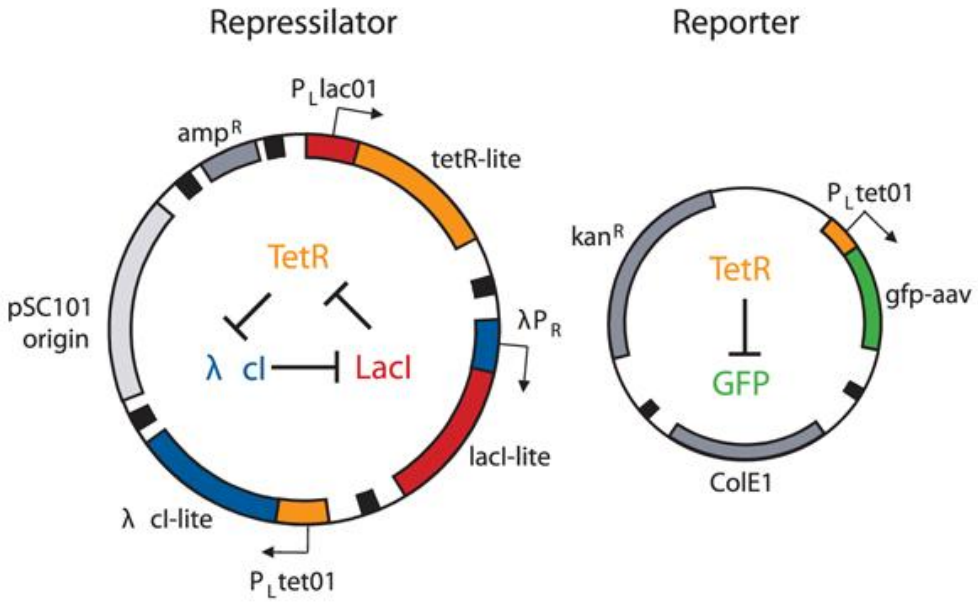


Figure 1 The Repressilator. The Repressilator was the first attempt to construct a synthetic network capable of oscillations. It consisted of three genes, each encoding a repressor that represses the transcription of the next gene in the network. A reporter plasmid was included to observe the oscillations in GFP (green fluorescent protein) levels. Reproduced from Ellowitz and Leibler, 2000.

Synthetic biology studies of genetic clocks started with the pioneering paper of Elowitz and Leibler (2000), which presented a simple genetic network constructed in a plasmid consisting of three genes, each encoding a repressor, linked together in a daisy-chain network based on genetic repression (Figure 1). This network was named *The Repressilator* , following the trend of *The Brusselator* (Glansdorff and Prigogine, 1971) and *The Oregonator* (Field and Noyes, 1974), which were the first theoretical

models of chemical oscillators, i.e. chemical reactions which could produce sustained oscillations.

The Repressilator was constructed in a multi-copy *E. coli* plasmid, using the genes *lacI*, *tetR* (from the Tn10 transposon), and cI (from phage λ). These genes were expressed from synthetic promoters that had been engineered to contain operator sites to which the three repressors could bind. As a reporter gene the *tetR* promoter was fused to a green fluorescent protein gene (*gfp*), which allowed the observation of the variation of the expression level of *gfp* in single cells by fluorescence microscopy. The Repressilator network showed noisy oscillations in about 40% of the observed cells (Elowitz and Leibler 2000), with the rest of the cells in a population showing constant GFP levels, or no expression of GFP at all. When a population of cells was grown in a culture flask, the fluorescence in samples did not show an oscillatory trend, indicating that the cells that were capable of oscillations were either out of synchrony, or the constant fluorescence signal from not oscillating cells was dominant. Thus, the Repressilator did not function well as a synthetic genetic clock, either at the single cell level or population level. Nevertheless, this device was the first in the field of synthetic gene oscillators, and it inspired a line of research that persists until this day and includes this thesis.

The second attempt to construct a synthetic genetic clock was the work of our laboratory (Atkinson et al, 2003); and the resulting clock is studied in detail in this thesis. I will refer to this device as the Ninfa clock. Other synthetic genetic clocks have been constructed since then in *E.coli* (Fung et al 2005, Stricker et al 2008, Danino et al 2010) and in mammalian cells (Tigges et al 2009). A comparative

analysis of all the published synthetic genetic clocks has been presented (Purcell et al, 2010). Of all these clocks, only the Ninfa clock has been studied in cell populations and produces a very long period on the order of circadian time keeping.

The Ninfa clock

The Ninfa clock is an abstraction of a part of the natural mammalian circadian clock. Figure 2 presents a simplified version of the mammalian circadian clock that illustrates its key features. The "activator" of the mammalian clock is a complex of the Clock and Bmal1 proteins; this complex drives the expression of its activator,

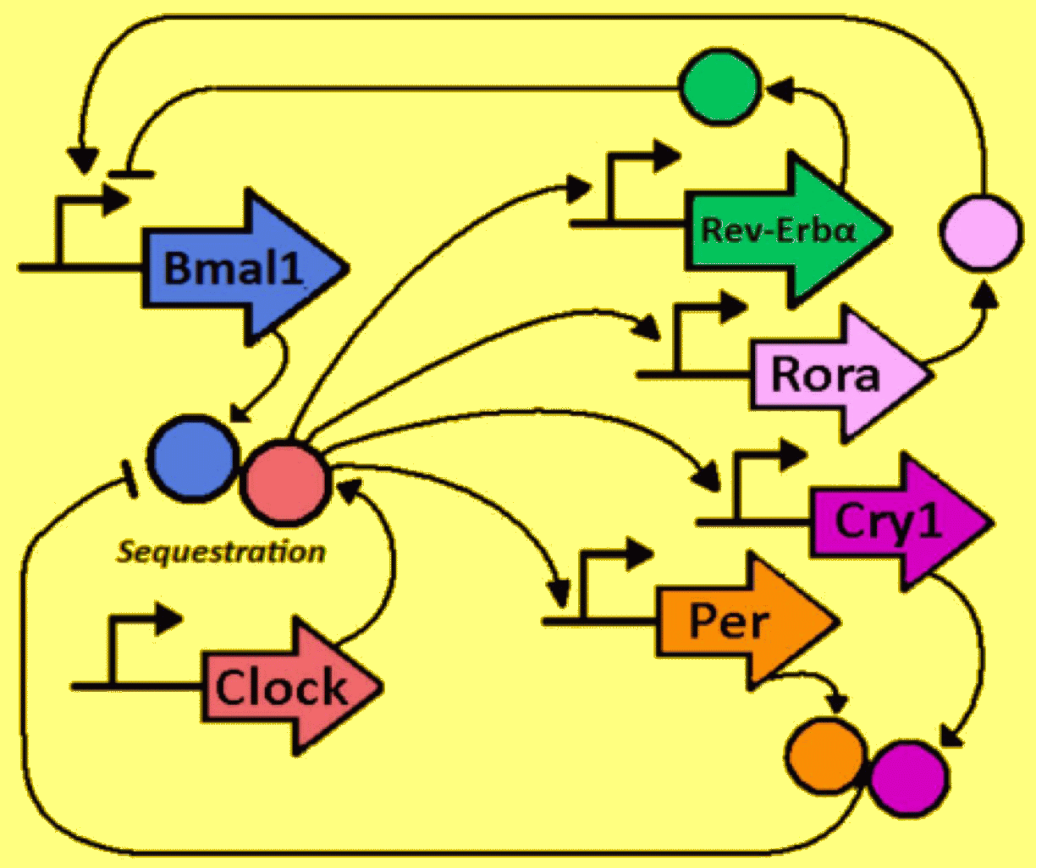


Figure 2. Schematic of the mammalian circadian clock. The colored big arrows correspond to genes, while colored circles represent proteins encoded by the corresponding genes.

Rora, and its repressor, Rev-Erbα. In addition, the complex of the Bmal1 and Clock proteins drives the expression of the Cry and Per proteins; with the Per protein serving to sequester the Clock-Bmal1 complex and inhibit its activity. Thus, the mammalian circadian network consists of a positive feedback loop (Clock-Bmal1 drives expression of Rora), a negative feedback loop based on repression (Clock-Bmal1 drives expression of Rev-Erbα), and a second negative feedback loop based on sequestration (Clock-Bmal1 drives expression of Per). The Ninfa clock is an abstraction of the positive feedback loop and the negative feedback loop based on repression; in the Ninfa clock, the activator activates its own expression and activates the expression of its repressor (Figure 3). Thus, the Ninfa clock is similar to the mammalian clock, but lacks the sequestration-based feedback loop.

The Ninfa clock also bears some similarity to a theoretical clock that has been described (Barkai and Leibler 2000). The theoretical clock of Barkai and Leibler consisted of an activator that activated its own expression and the expression of a protein that brought about its proteolysis. Thus, it consists of linked positive and negative feedback loops, but unlike the Ninfa clock, the negative feedback loop is based on proteolysis of the activator. Although the Leibler and Barkai design has not been implemented, modeling of this network showed a parameter space capable of sustained oscillations.

The Ninfa clock was implemented using elements of the nitrogen assimilation control two-component system in *E. coli*, as well as the gene LacI, the repressor of the *lacZYA* operon (Figure 3). The clock consists of two modules, the activator module and repressor module, that are separately integrated into the *E. coli*

chromosome at two different sites. The activator module is integrated near the origin of replication, within the rbs operon, while the repressor module is located within the glnK region of the chromosome. To construct the activator module; the gene *glnG*, whose product is NRI, the nitrogen-assimilation system transcriptional activator, was fused to a modified *glnA* promoter in which repression by the LacI repressor was engineered by inserting two *lac* operator sequences, one immediately downstream from the start of transcription initiation, as in the *lacZYA* operon, and one upstream

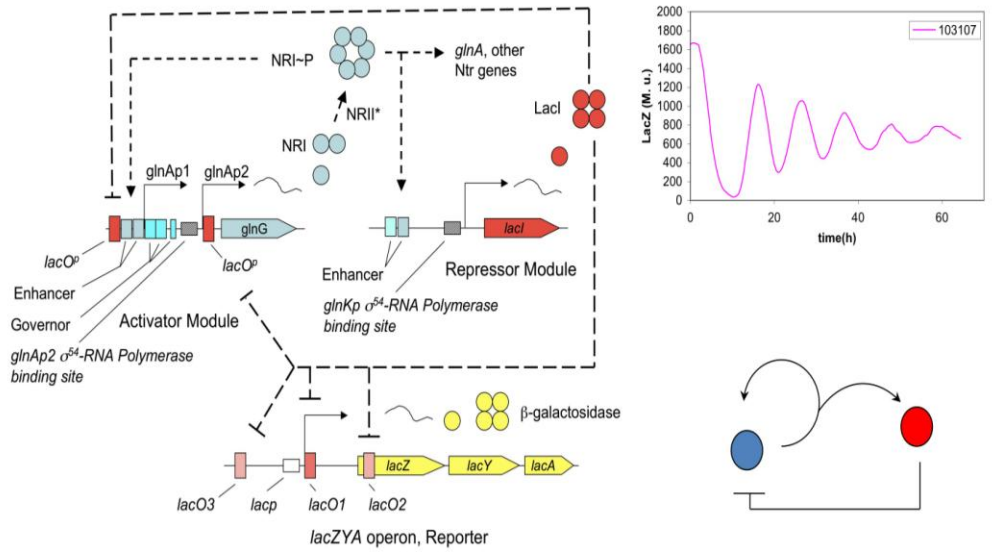


Figure 3. Schematic depiction of the Ninfa clock. The clock consists of linked positive and negative feedback loops (simple topology at bottom right). The positive feedback loop consists of an activator, NRI~P, that activates its own transcription. The activator also activates the transcription of its repressor, LacI. In addition to controlling the expression of activator, repressor controls the expression of the chromosomal *lacZYA* operon of *E. coli*, and expression of the *lacZ* gene product, β-galactosidase, can be used as a reporter of clock function. An example of the oscillations from the clock is shown at the upper right, where LacZ activity from different times of a clock experiment are shown. In the experiment shown, five clear cycles of *lacZ* expression were observed over a time of ~70 hours.

from the enhancer sequences for the binding of the NRI protein. The purpose of the upstream operator site was to allow formation of a repression loop, as at the *lacZYA* operon, and thus facilitate repression. The NRI protein activates its own expression (by binding to enhancer sequences in its own promoter region) when it becomes phosphorylated. In the synthetic genetic clock, NRI protein is always phosphorylated, due to the constitutive expression of a mutant form of NRII, the protein that regulates NRI activity by phosphorylation and dephosphorylation. The mutant form of NRII used in clock studies is unable to bring about the dephosphorylation of NRI~P and brings about the phosphorylation of NRI under all conditions. Since the mutant NRII keeps the NRI protein always phosphorylated, NRI expressed from the synthetic module always autoactivates its own expression as well as drives the expression of repressor. The repressor module of the clock consists of the LacI gene, encoding the *lac* repressor, fused to the *glnK* promoter. This promoter is activated by NRI~P, but it has a weaker enhancer than is found in the activator module. Thus, the activator module is activated first as NRI~P is produced by the clock, and only after the level of NRI~P has become elevated due to its auto-activation is the repressor module activated, bringing about the appearance of repressor.

In addition to the activator (*glnG*) and repressor (*lacI*) modules, the Ninfa clock also includes a reporter module, which consists of the *lacZ* gene under the control of a promoter that is activated or repressed by one of the clock genes. The reporter module, as its name indicates, its necessary in order to measure the oscillations. In the strain used in Atkinson et al 2003 (called NC12), the reporter was not synthetic: it consisted of the wild type *lacZYA* operon, which is repressed by LacI.

The expression of the *lacZ* gene is repressed upon expression of LacI from the repressor module. In this thesis, a variety of reporters will be used, including the *lacZYA* operon (as Atkinson, 2003) and fusion of the *lacZ* gene to the *glnA* and *glnK* promoters.

In order to study clock function, the host cells containing the clock must contain null mutations of the natural copies of *lacI*, *glnG* (encoding NRI) and *glnL* (encoding NRII). This will ensure that the NRI and LacI produced by the activator and repressor modules will be the only source of these proteins in the cell. The wild-type *glnL* gene is replaced with a mutant form of the gene encoding an NRII* protein that phosphorylates NRI under all conditions. In most of our studies, the NRII* protein is expressed from a multicopy plasmid, and it is assumed to be present in excess.

The clock strain used in the experiments presented in Atkinson et al (2003) was called NC12. After that publication, modifications of the activator module were made, producing a strain called NC77 (Ninfa et al, 2009). These modifications to the activator module were as follows: the "governor sequences" mapping between the enhancer sites and the site of polymerase binding were scrambled; this was shown to prevent high concentrations of NRI~P from decreasing promoter activity by binding to the governor sites (Atkinson et al, 2002). Also, the proximal operator (mapping near the site of transcription initiation) was converted from the high-affinity "lacO^{perfect}" sequence to the *lacO$_1$* sequence found at the proximal position of the lac operon. Thus, NC12 had two copies of l*acO^{perfect}* in the activator module, whereas NC77 has one *lacO^{perfect}* and one *lacO$_1$* operator. NC77 produced slightly less

damping in clock oscillations, and therefore it was the primary strain chosen for this study (Ninfa et al, 2009). Due to an unknown bacteriophage that seemed to be airbone and was affecting the clock cultures, a strain identical to NC77 was constructed on a host strain resistant to the unknown bacteriophage. This strain is called NCPR77 and it is the strain used in most experiments throughout this thesis, unless otherwise indicated.

A clock experiment involves measuring the expression of reporters in a cell population growing in continuous culture (Akinson et al, 2003; Ninfa et al, 2007). Briefly, a culture is grown overnight from a single colony of the clock strain, using defined growth medium. The seed cultures had a 250 ml volume and also contained 200 uM IPTG. IPTG was added to the medium in order to inhibit the LacI protein expressed from the repressor module, so in principle the expression of the *glnG* gene in the activator module is at its maximum level of expression in every single cell. This treatment was used to entrain all the cells to be at the same point in their clock cycles, synchronizing the cells at the beginning of each experiment. After overnight growth and entrainment, the culture is washed twice and resuspended in new medium without IPTG, and transferred to a flask where a continuous culture is maintained for several days. After the cells are washed from IPTG in the medium, the cells are expected to lose their internal IPTG by dilution (driven simply by cell division and the pumping of fresh medium), and after that the clock dynamics are expected to show oscillations in gene expression. The level of expression of the enzyme β-galactosidase (*lacZ* product) is measured every half hour from a liquid sample using the standard Miller Assay.

When the clock strains NC12 or NC77 were grown in a continuous culture and their *lacZ* expression measured at half-hours intervals, the results were that the average *lacZ* level in the culture oscillated with a period of ~10 hours, but the oscillations were damped (eg. Figure 3). The *lacZ* levels show 4-5 oscillations with decaying amplitude, until the *lacZ* level reaches an intermediate steady-state expression level.

Given the decay in amplitude, one important question to ask about the Ninfa clock dynamics is why do the oscillations tend to lose their amplitude with time, and whether this decaying can be fixed, in order to have a a genetic network which produces sustained oscillations (as opposed to damped oscillations) in gene expression. This question can have two possible broad answers: the clock design produces damped oscillator dynamics in each single cell, or the clock dynamics produces sustained oscillations dynamics in single cells, but the cells tend to lose synchrony, so the average expression level looks like a damped oscillator. An additional factor when considering the amplitude decay is the mutation rate. During the course of the experiment, if any cell that mutates in one of the clock components (that causes it to lose the oscillation dynamics) develops an environmental advantage in growth rate, the mutated cell would reproduce faster that the cells with intact clock dynamics, eventually becoming the majority of cells in the culture. This would cause the average expression level in the culture to become steady.

The probability that a spontaneous mutation appears in one of the clock modules is low. The mutation rate in *E. coli* is as low as 10^{-10} for a base substitution in a generation. Given this number we could assume a low estimate of 10^{-9} for the

probability of a mutation appearing in the clock modules of a single cell. After 10 generations, this probability would be 10^{-8}. Since the amount of cells in a 100 ml culture is approximately 10^{10} cells, a mutation in at least one of the clock modules is almost certain to appear after a few generations. This mutation could be advantageous to cell growth rate, since presumably the clock synthetic network would add a "drag" to the growth rate of cells. Presumably, if a faster-growing mutant arises and its growth advantage is significant, it will eventually take over the culture. However, we have found that advantageous mutations by themselves can not explain the loss of amplitude in oscillations, because the clock experiments are very repeatable up to the second or third period of oscillating gene expression. Since mutations are essentially random, advantageous mutations are expected to appear at different times during the experiments, and this should prevent the observed repeatability. Also, I will show in subsequent Chapters of this thesis that various manipulations can increase or decrease damping in a reproducible fashion, which also argues against the damping being due mainly to random mutations. It should be noted that one way to investigate the clock dynamics in populations versus dynamics in single cells would be to observe the expression dynamics of single clock cells under a microscope, by switching the reporter from *lacZ* to *gfp* (green fluorescent protein gene). However, due to the lack of a dedicated fluorescence microscope in the lab, and because of very low expression levels of the clock genes and clock reporters, all the experiments in this work were performed at the population level.

In Chapter 2 of this thesis, I present an automated system for conducting clock experiments. This automated system consists of a computer-controlled "turbidostat"

that allows maintenance of a continuous culture at a desired turbidity for indefinite periods, and provides a highly accurate automatic sampling mechanism. This device allowed me to perform several dozen experiments, each of which had a duration of 2-4 days. Without the automated device it would not have been possible to perform this many experiments, due to the need of constant monitoring of the culture.

My goal is to investigate the cause of amplitude damping in the Ninfa clock, with the purpose of decreasing this damping and possibly achieving a clock that shows sustained oscillations. With this purpose in mind, the main work in this thesis was done to characterize the behavior of the clock and its components and thus identify the conditions under which the clock dynamics would show significant variations. By analyzing variations in the clock dynamics due to changes in environmental or genetic conditions, one could find the conditions that alter different aspects of the clock such as damping rate, amplitude and period, as well as synchronization of clocks in individual cells. With such an analysis, conditions could be identified that minimize damping, for example, or conditions that maximize individual clocks synchronization. Identification of conditions that change some parameters could also be useful for future applications in synthetic biology, such as the variation of the period of oscillation.

In Chapter 3 of this thesis I characterized the clock dynamics with various environmental conditions, namely the cell density in the culture, the growth temperature and the concentration of carbon source in the media. It was found that there is a linear correlation between the cell's growth rate and the period of oscillation. However, variations in temperature or carbon source concentration do not affect the

amplitude of oscillations. Also, I found that the amplitude of oscillations is lower than normal when the culture density is low, a finding that may indicate that there is an environmental signal that keeps clocks in individual cells synchronized.

In Chapter 4, I attempt to validate the hypothesis that clocks in individual cells are fully or partially synchronized. Since the cells seem to keep a degree of synchronization for several days it is reasonable to think that they must have a signal in the environment that keeps them synchronized. We found evidence that the metabolite acetate could be such a signal, since cells in which the acetate utilization and formation pathway has been suppressed do not show oscillations in the clock expression.

In Chapter 5, I show that clock dynamics were very sensitive to small changes in the host cell, particularly, to the presence of sites that sequester the activator protein. The clock protein NRI has a set of binding sites in the chromosome with various degrees of binding affinities, which were not thought to influence the clock dynamics. However, due in part to the small number of proteins that each clock module synthesizes, we found that suppression of some of these binding sites has a noticeable effect of clock dynamics, particularly in the damping rate of the oscillations amplitude. However, suppression of binding sites does not necessarily cause less damping, it could also increase damping.

In Chapter 6, I present theoretical work on mathematical models of synthetic genetic oscillators. Specifically, I investigate the preferred way to construct clocks with 3 modules, as a modification of the Ninfa clock. By comparing the likelihood of oscillations in two mathematical models of clocks with three genes, I show that one

of these models, that we call the activilator, has the advantage that it can be "guided" towards an oscillatory regime when oscillations in gene expression are not initially seen. I also outline a couple of way such activilator clock could be implemented in *E.coli*.

Finally, in the appendix of this thesis, I describe synthetic genetic modules that I constructed, which are now being used by other researchers in the laboratory. In particular, I will describe a fusion of the *lacI* structural gene to the NRI-repressed *glnL* promoter of *E. coli*; this construct is currently being used in toggle switch studies by Rohit Singh. I will also describe modules that can be assembled into clocks with multiple activator proteins.

LIST OF REFERENCES CITED IN CHAPTER 1

Anderson JC, Voigt CA and Arkin AP *Environmental signal integration by a modular AND gate*. *Mol. Syst. Biol*. 2007 3, 133

Andrianantoandro E, Basu S, Karig D and Weiss R, *Synthetic biology: new engineering rules for an emerging discipline, Mol. Syst. Biol* 2006 2, p. 0028.

Atkinson MR, Savageau MA, Meyers JT and Ninfa AJ: *Development of genetic circuitry exhibiting toggle switch or oscillatory behavior in Escherichia coli*. *Cell* 2003, 113:597-607.

Atkinson MR, Pattaramanon N, and Ninfa AJ. Governor of the glnAp2 promoter of *Escherichia coli. Mol. Microbiol* 2002 46, 1247-1257.

Bassler BL Small talk. *Cell-to-cell communication in bacteria* Cell. 2002 109(4):421-4.

Barkai N and Leibler S *Biological rhythms: Circadian clocks limited by noise* Nature 2000 403, 267-268

Basu S, Mehreja R, Thiberge S, Chen MT and Weiss R *Spatiotemporal control of gene expression with pulse generating networks*. Proc Natl Acad Sci U S A 2004,101:6355-6360.

Basu S, Gerchman Y, Collins CH, Arnold FH and Weiss R: *A synthetic multicellular system for programmed pattern formation*. Nature 2005, 434:1130-1134.

Benner SA and Sismour AM *Synthetic Biology* nature reviews genetics 2005 6, Pages: 533-543

Chang DE, Leung S, Atkinson MR, Reifler A, Forger DB and Ninfa AJ. *Building biological memory by linking positive feedback loops* PNAS January 5, 2010 vol. 107 no. 1 175-180

Clomburg JM and Gonzalez R *Biofuel production in Escherichia coli: the role of metabolic engineering and synthetic biology* Applied microbiology and biotechnology 2010 Volume: 86 Pages: 419-434

Connor MR and Atsumi S *Synthetic Biology Guides Biofuel Production* Journal of biomedicine and biotechnology 2010 Article Number: 541698

Danino T, Mondragón-Palomino O, Tsimring J and Hasty J *A synchronized quorum of genetic clocks* Nature 2010 463, 326-330

de Lorenzo, V *Systems biology approaches to bioremediation* Current opinion in biotechnology 2008 Volume: 19 Pages: 579-589

Elowitz MB, Leibler S: *A synthetic oscillator network of transcriptional regulators*. Nature 2000, 403:335.

Endy D, *Foundations for engineering biology*, Nature 2005 438, pp. 449–453.

Field RJ and Noyes RM, *Oscillations in Chemical Systems IV. Limit cycle behavior in a model of a real chemical reaction*, J. Chem. Phys. 1974 60, 1877-84.

Fung E., Wong WW, Suen JK, Bulter T, Lee SG, and Liao JC (2005). *A synthetic gene-metabolic oscillator*. Nature 2005 435, 118–122

Gardner TS, Cantor CR and Collins JJ: *Construction of a genetic toggle switch in Escherichia coli*. Nature 2000, 403:339-342.

Ghim CM, Kim T, Mitchell RJ and Lee SK *Synthetic Biology for Biofuels: Building Designer Microbes from the Scratch* Biotechnology and bioprocess engineering 2010 Volume: 15, Pages: 11-21

Glansdorff P and Prigogine I. *Thermodynamic Theory of Structure, Stability and Fluctuations*, 1971 London and New York: Wiley

Golden SS, Ishiura M, Johnson CH and Kondo T *Cyanobacterial circadian rhythms.* Annual review of plant physiology and plant molecular biology. 1997 Volume: 48 Pages: 327-354

Golden SS and Canales SR *Cyanobacterial circadian clocks - Timing is everything* Nature reviews microbiology 2003 Volume: 1 Pages: 191-199

Guet CC, Elowitz MB, Hsing W. and Leibler S *Combinatorial synthesis of genetic networks.* Science 2002 296, 1466–1470

Heinemann M and Panke S *Synthetic biology—putting engineering into biology* Bioinformatics (2006) 22 (22): 2790-2799

Helfrich-Forster C *Drosophila rhythms: From brain to behavior* Seminars in cell & developmental biology 1996 Volume: 7 Pages: 791-802

Kaern M, Blake WJ and Collins JJ, *The engineering of gene regulatory networks* Annu Rev Biomed Eng 2003 5, pp. 179–206.

Keasling JD *Synthetic biology for synthetic chemistry* ACS chemical biology 2008 Volume: 3 Pages: 64-76

King DP and Takahashi JS *Molecular genetics of circadian rhythms in mammals* Annual review of neuroscience 2000 Volume: 23 Pages: 713-742

Kramer BP and Fussenegger M. *Hysteresis in a synthetic mammalian gene network*. Proc. Natl. Acad. Sci. USA 2005 102, 9517–9522.

Loros JJ *Time at the end of the millennium: the Neurospora clock* Current opinion in microbiology 1998 Volume: 1 Pages: 698-706

Loros JJ and Dunlap JC *Genetic and molecular analysis of circadian rhythms in Neurospora* Annual review of physiology 2001 Volume: 63 Pages: 757-794

Mayo AE, Setty Y, Shavit S, Zaslaver A and Alon U. *Plasticity of the cis-regulatory input function of a gene*. PLoS Biol. 2006 4, e45

Neumann H and Neumann-Staubitz P *Synthetic biology approaches in drug discovery and pharmaceutical biotechnology* Applied microbiology and biotechnology 2010 Volume: 87 Pages: 75-86

Ninfa AJ, Selinsky S, Perry N, Atkins S, Song QX, Mayo A, Arps D, Woolf P and Atkinson MR *Using Two-Component Systems and other Bacterial Regulatory Factors for the Fabrication of Synthetic Genetic Devices* Methods Enzymol. 2007; 422: 488–512.

Ninfa AJ, Atkinson MR, Forger DB, Atkins S, Arps D, Selinsky S, Court D, Perry N and Mayo A. (2009) *A synthetic biology approach to understanding biological oscillations: developing a genetic oscillator for Escherichia coli*. pp301-329 In: Bacterial Circadian Programs (Ditty JL, Mackey SR, Johnson CH, eds) Springer-Verlag, Berlin

Porcar M and Moya A. *Paving the way for synthetic biology-based bioremediation in Europe.* Microbial Biotechnology 2010 3:2, 134-135

Purcell O, Savery NJ, Grierson CS and di Bernardo M. *A comparative analysis of synthetic genetic oscillators.* J R Soc Interface. 2010 Nov 6;7(52):1503-24.

Scully AL and Kay SA *Time flies for Drosophila* Cell 2000 Volume: 100 Pages: 297-300

Stricker J, Cookson S, Bennett MR, Mather WH, Tsimring LS and Hasty J: *A fast, robust and tunable synthetic gene oscillator*.Nature 2008, 456:516-519.

Tabor JJ, Salis HM, Simpson ZB, Chevalier AA, Levskaya A, Marcotte EM, Voigt CA and Ellington AD *A synthetic genetic edge detection program*. Cell 2009, 137:1272-1281.

Takahashi JS *Molecular neurobiology and genetics of circadian-rhythms in mammals* Annual review of neuroscience 1995 Volume: 18 Pages: 531-553

Tamsir A, Tabor JJ and Voigt CA *Robust multicellular computing using genetically encoded NOR gates and chemical 'wires'*. Nature 2010 469, 212–215

Tigges M, Marquez-Lago, TT, Stelling J, and Fussenegger M. *A tunable synthetic mammalian oscillator*. Nature 2009 457, 309–312.

Tropynina TS, Golubev OV, Stupak EE and Tchuraev RN: *Construction of an articial digenic network with epigenetic properties*. Mol Biol 2002, 36:472-475.

Weber W and Fussenegger M *The impact of synthetic biology on drug discovery* Drug discovery today 2009 Volume: 14 Pages: 956-963

CHAPTER 2

AUTOMATIZATION OF THE CLOCK EXPERIMENTAL SETUP

Authors note

Portions of this chapter have been previously published in the following two maunscripts:

Ninfa AJ, Selinsky S, Perry N, Atkins S, Xiu Song Q, Mayo A, Arps D, Woolf P, Atkinson MR (2007) Using two-component systems and other bacterial regulatory factors for the fabrication of synthetic genetic devices. *Methods Enzymol* 422, 488-512.

Ninfa AJ, Atkinson MR, Forger DB, Atkins S, Arps D, Selensky S, Court D, Perry N, Mayo A. (2009) A synthetic biology approach to understanding biological oscillations: developing a genetic oscillator for Escherichia coli. pp301-329 In: *Bacterial Circadian Programs* (Ditty JL, Mackey SR, Johnson CH, eds) Springer-Verlag, Berlin

Introduction

To study the function of a synthetic genetic clock in an *E. coli* strain, it is necessary to periodically assay the enzymatic activity of a reporter of clock activity in a continuous culture. Since the typical period of oscillation observed for strains

containing the Ninfa clock was about 10-12 hours, we needed a culture that was alive for several days in order to observe 4-5 oscillations. Depending on growth media and temperature, an *E. coli* culture will reach saturation and begin to deplete the media nutrients, and therefore start to die, in 12-24 hours. Thus, in order to observe several oscillations in a single culture, a continuous culture method must be used. The simplest way to achieve a continuous culture is to periodically dilute a growing culture with fresh medium. Indeed, this method was used in the earliest studies of the Ninfa clock; however, it is not a good method for quantitative analysis, as culture conditions obviously change dramatically when the culture is diluted, unless the dilutions are performed while the culture is at very low turbidity. Another possibility for obtaining a continuous culture would be to pump media into the growth vessel continuously, while at the same time either pumping culture out to keep the volume of the culture constant, or having the culture flow out of the growth flask when its volume exceeds a certain level. This approach requires keeping the flow of media into (and out of) the culture at an appropriate rate, such that the flow is sufficient to prevent the bacteria from becoming saturated while not so much as to cause "wash-out" (the bacteria are dilute away). Either the dilution method or a continuous culture method as described above require continuous human monitoring of the culture, to keep the bacterial concentration at a level where they are growing, and to take samples from the culture periodically in order to assay them for clock activity. Such monitoring is especially important if the rate of growth of the cells changes during the experiment, if one desires to maintain the culture at a certain constant turbidity, or if

the rate of growth is very rapid such that a small miscalculation of the rate leads to wash-out or the culture growing beyond the desired turbidity.

Chemostat methods for maintaining a continuous bacterial culture have been known for some time (Monod 1950, Novik and Szilard 1950), and are used in many settings, such as studies of microbial physiology or directed evolution experiments (Drake and Brodgen, 2002; Dykhuizen and Hartl, 1983; Ferenci, 2008; Kafkewitz et al, 1973). In early studies of the Ninfa clock, a New Brunswick Microfrem chemostat was used to monitor clock function (Atkinson et al, 2003). This chemostat consisted of a peristaltic pump that fed the culture vessel with fresh sterile media; the culture vessel had a port that allowed medium to flow out when the culture volume exceeded a certain limit. This simple set-up kept the volume of the culture constant, while allowing the manual adjustment of the pump speed, thus keeping the culture concentration in check. But, in order to keep the culture concentration within bounds and periodically sample the culture, it was required to manually take a sample from the culture, manually measure its optical density (OD, proportional to concentration within a range), adjust the speed of the pump accordingly, and assay the enzymatic activity of the reporter in the sample. This procedure was required every 45 minutes, in order to avoid big changes in the turbidity of the culture in the vessel.

The requirement of constant monitoring for a continuous culture makes it very difficult to perform an experiment that lasts 4-5 days, as is required to study the Ninfa clock. All the experiments that were performed for the Atkinson et al (2003) publication of the Ninfa clock required at least one person monitoring the experiment at all times, for up to 80 hours in a row. Thus, an experiment could not be performed

by a single person, but required a team of workers to continuously monitor the culture and take samples.

In order to perform many experiments to study different aspects of the Ninfa clock, and avoid the requirement for constant monitoring of the clock experiments, we built an instrument that performs an automated version of the experiment. In this automated system, a computer controls the optical density of the culture by reading optical density periodically and adjusting the flow of media in (and flow of culture out) accordingly. The instrument also can be programmed to take samples from a culture periodically, using an automated sampler, and to store samples at low temperature (~4 C) so that the enzymatic activity in a sample remains stable for one or two days. With this automated device, an experiment could be started at some point, and the samples could be assayed up to two days after they were taken from the vessel, giving the person running the experiment flexibility in time by assaying large sets of samples simultaneously, and thus avoiding the need of constant monitoring and assaying of samples. In other words, the researcher could start an experiment one night, then leave the experiment running by itself, and come back in the morning to assay all the samples taken overnight. Meanwhile the experiment can keep running, and the next morning a new set of samples would be stored, ready to be assayed. This automation not only allowed more experiments to be performed, but increased the accuracy of sampling and reproducibility of the culture conditions.

The technology necessary to construct such an automated experiment is readily available, and it is relatively inexpensive. The system is composed by 3 peristaltic pumps, two diverter valves, one “Z flowcell” for optical density measures,

one fiber optics spectrometer, and one auto sampler. All of the components were purchased from FIALab (Bellevue WA). All of these components can be controlled from the computer, via a serial or USB port.

The automated system

The 250 mL cell culture is maintained in a 1 liter Erlenmeyer flask. One Alitea VS series peristaltic pump (Figure 2-1) is used to pump media into the culture and simultaneously remove culture from the Erlenmeyer, to keep the volume of the culture constant. The speed of this pump can be varied from the computer; the flow driven by the pump is proportional to the pump speed, and also proportional to the width of the tubing used to wrap around the peristaltic pump: The wider the tubing, the more flow driven by the pump. For this pump and the other pumps we used Bio-rad Phar-med TM tubing, which is the more resistant tubing with respect to peristaltic pump wear. For pumping media into the culture flask, we used tubing with internal width 1.6 mm, which allowed a maximum flow of 5.8 mm/minute. For pumping media out we used tubing with internal width 3.2 mm. However, the flow of culture out of the system vessel is not double the flow of media in. The flow of liquid out from the culture is set so the tubing sucks media out of the flask when the culture accumulates a volume higher than 250 ml (Figure 2-2). This arrangement keeps the volume of the culture constant, so the flow of culture out is always the same as flow of media in, regardless of the differences in viscosity between fresh media and culture. As a media reservoir, we use a set of 1 gallon bottles of fresh media that can be connected in a daisy-chain, so one single pump can suck media from all the bottles

Figure 2-1 Alitea VS series peristaltic pump This kind of pump was used to pump media into and out from the culture, and to take a sample of the system every ½ hour. This pump can be controlled from a computer via a serial port.

(Figure 2-3). The waste fluids of culture removed from the culture flask is directed towards a waste container, which is emptied manually.

In addition to the pump that controls medium flow into and out from the growth flask, we used another pump to take samples of the culture and direct the samples to the auto sampler. We also used an air line to flow filtered air into the culture. The reason for flowing air to the culture was to keep the culture aerated at all times in order to avoid changes of metabolism due to changes in the availability of oxygen. A schematic of the culture flask with all the media/culture/sampling lines is shown in Figure 2-2.

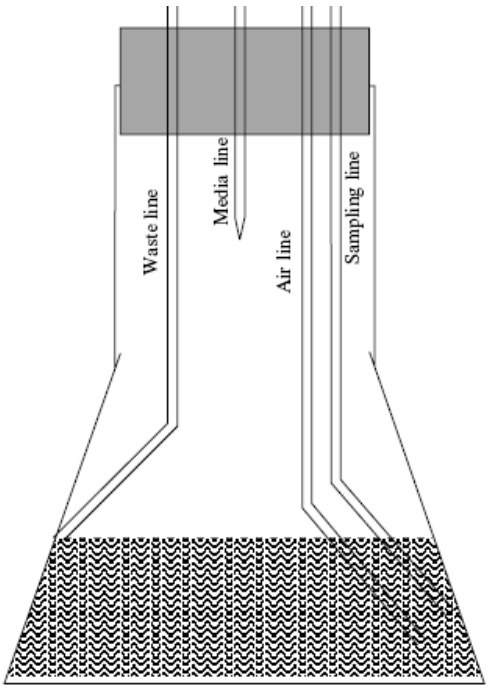


Figure 2-2 Schematics of the culture flask. Waste line and media-in ine are driven with one peristaltic pump. Another pump is used to take samples from the culture periodically and store them in an auto sampler inside a fridge. The air line flows filtered air to culture to keep the culture aerated at all times.

A key component of the instrument is the spectrophotometer that is used to measure culture turbidity, allowing adjustment of nutrient flow to control culture turbidity. To measure culture turbidity, the sampling pump diverts media out of the culture periodically (typically every half an hour) and runs it through the "Z flow cell" , which is a cell that allows a ray of light to be shined across a section of culture of length 1 cm, and measures the amount light that passes through the culture. The ratio of light that passes through fresh media (measured previously) to light that passes unscattered through culture is known as the transmittance. The negative of the

logarithm of the transmittance is the optical density, and this quantity is proportional to the bacterial concentration in the range 0-1. This is exactly the same measurement

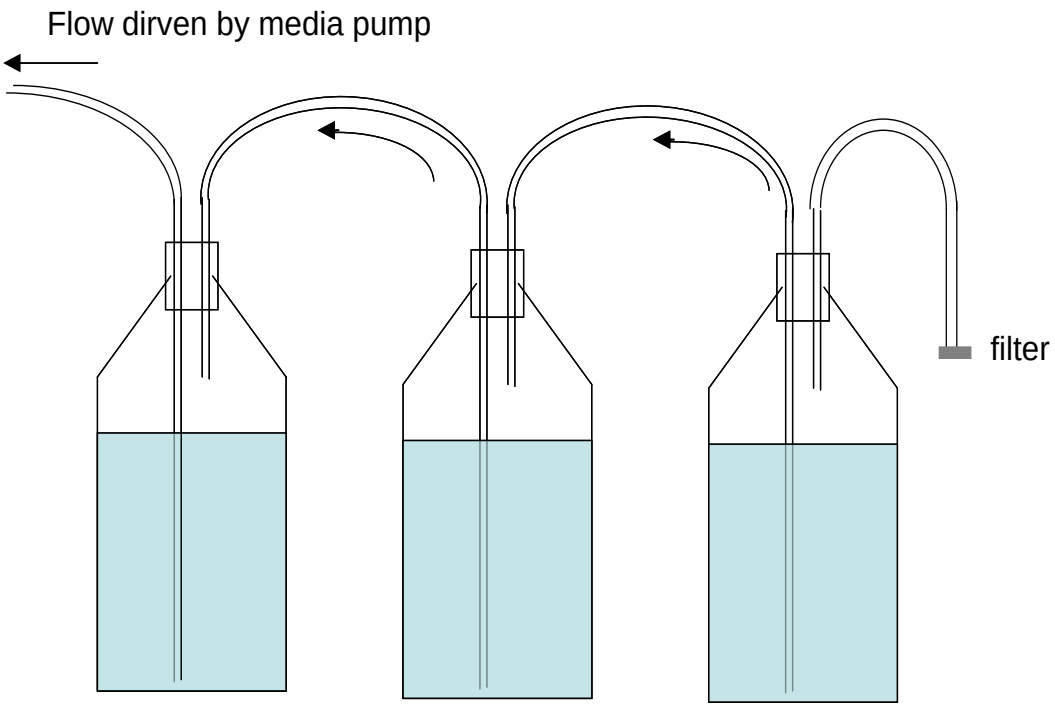


Figure 2-3 Daisy chain of media bottles A daisy chain of bottles allows the media pump to suck media from all bottles at once. A typical experiment required 4 1-gallon bottles of fresh media.

that is done in a regular spectrophotometer. The flow cell, bought from FIAlab, is shown in Figure 2-4. The amount of light that passes through the flow cell is measured via a fiber optic that transmits the signal to the spectrometer. The spectrometer (Ocean Optics SD200) decomposes the signal in the visible spectrum and sends the values to the computer. Then the computer can calculate the absorbance of the culture in the flow cell at any particular wavelength. We used 600 nm, which is the standard wavelength used to measure the density of *E. coli* cultures. The spectrometer SD200 from Ocean optic is pictured in Figure 2-5.

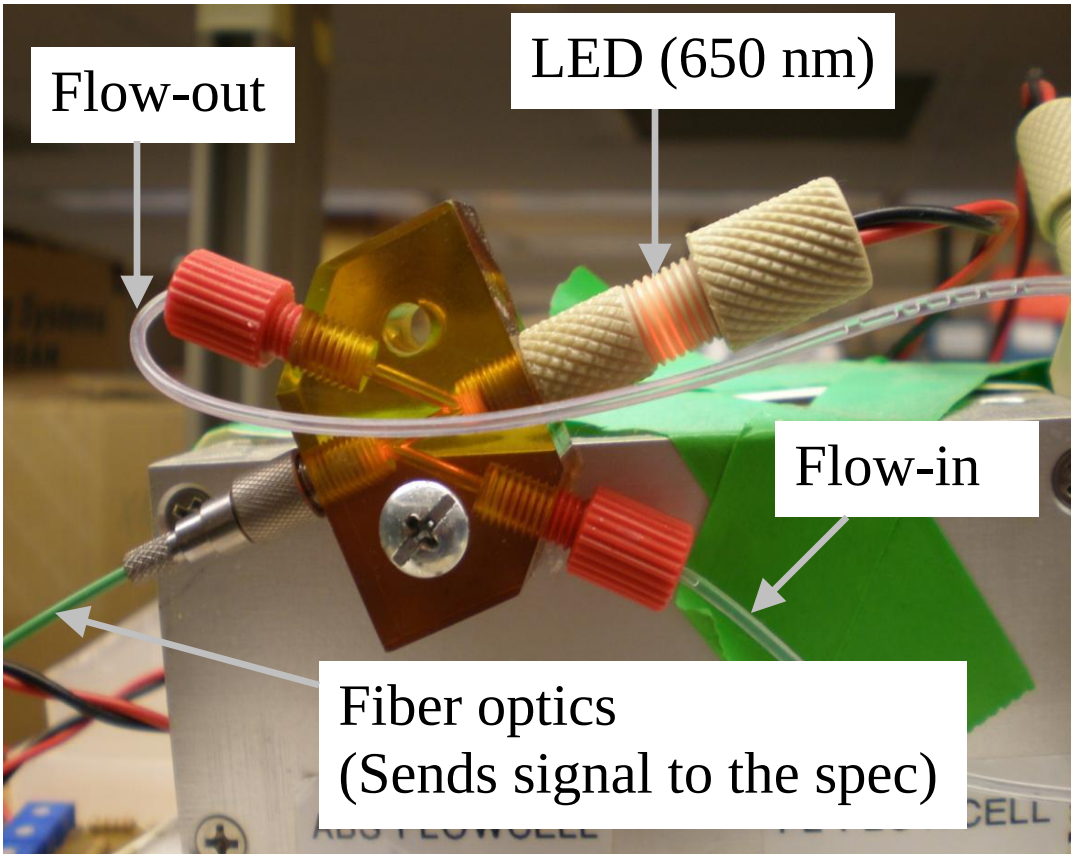


Figure 2-4 Z flow cell from FIAlab. This cell allows the measurement of the transmittance of light trough a section of culture of length 1 cm, and thus the optical density of the culture can be measured the same way a spectrophotometer measures OD in a regular cuvette.

Every half an hour, the sampling pump diverts culture towards an automated sampler located inside a fridge, where the samples are directed to test tubes and can be kept stable for about two days. The auto sampler can be controlled from the computer, so the flow can be directed to a specific test tube in a rack. Schematics of the whole system are shown in Figure 2-6.

A common problem in long-term fermentations involving flow-cell spectrophotometers is fouling of the flow cell by adhering cells from the culture;

bacteria tend to form biofilms on surfaces, which interfere with the spectrophotometer function. To deal with this issue, we explored various cell-washing routines. In our earliest description of the system (Ninfa et al, 2007) we used a water rinse to help minimize fouling of the flow cell. However, we have subsequently developed an

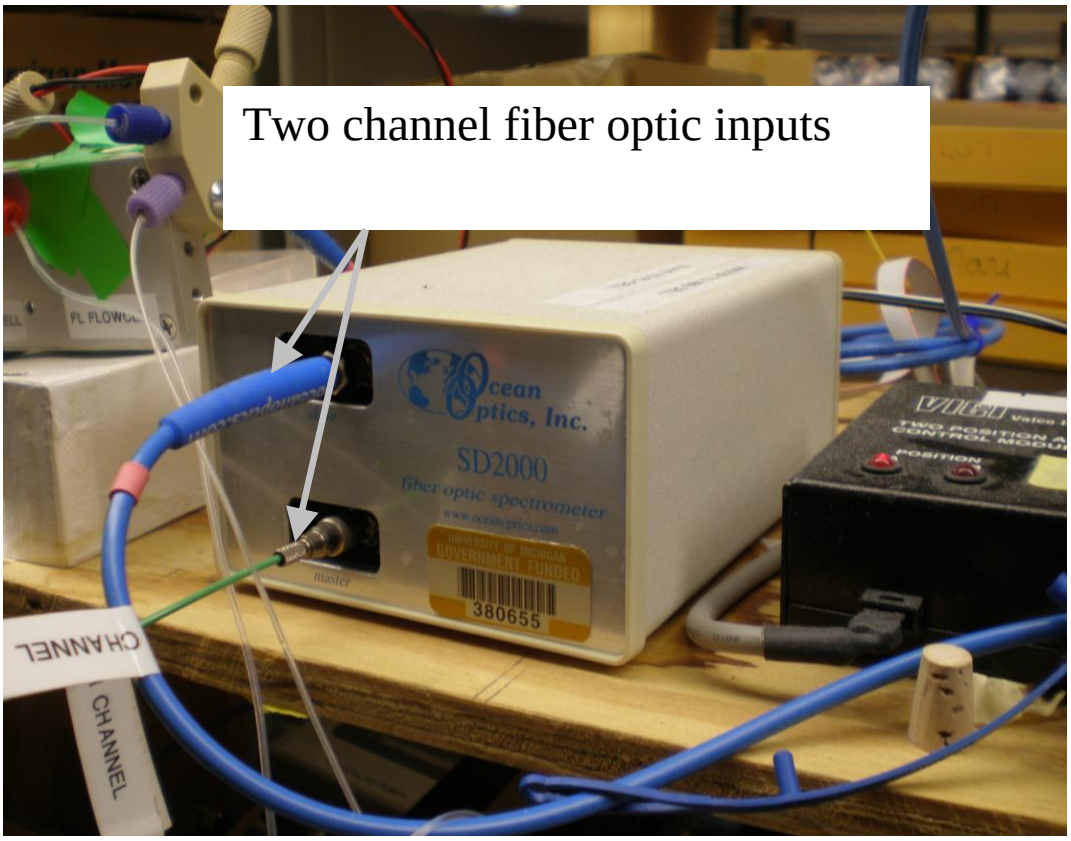


Figure 2-5 Spectrometer used for the automated system.

improved method. We implemented a washing routine that flows 70% ethanol through the flow cell for 15 seconds, followed by flowing sterile water for 1 minute. One peristaltic pump (Alitea S Series) is used to pump water/ethanol toward the flow cell. The "cleanliness" of the flow cell is monitored throughout the whole experiment

by measuring the intensity of light that passes through the flow cell when there is only water in it. This measurement remains fairly constant through the experiment,

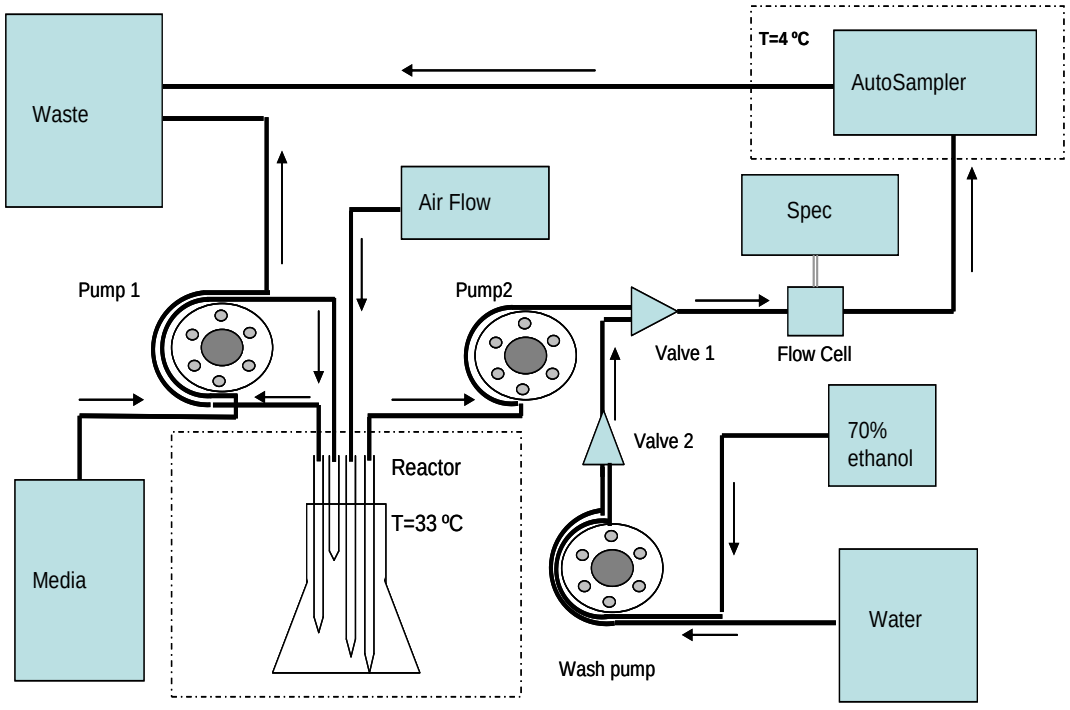


Figure 2-6 Schematics of the automated system. Pump 1 is used to flow media in and culture out. Pum 2is used to take culture out of the flask for OD measurements and sampling. The wash pump is used to wash the flow cell with ethanol and sterile water after each measurement

validating the effectiveness of the cleaning routine. Rarely, but occasionally, some particulate material in the culture or bubbles appearing in the washing line interferes with the measurement of turbidity. Typically, these issues need immediate correction and require intervention of the experimenter. For example, bubbles can be forced from the lines by pushing medium through the lines. Fortunately, these incidents are rare and only a few experiments have been lost because of them; however, they point to areas where further improvements can be made. The system has two computer

controlled diverter valves (VICI Cheminert TM valves) that allow selecting the flow from two possible inputs towards one output. One valve is used to alternate the flow toward the flow cell from the culture flask or ethanol/water from the washing containers. Another valve is used to select from ethanol or water to be directed toward the flow cell.

The automated system worked well most of the time and allowed us to perform several dozens of experiments that constitute the bulk of this thesis. However, since the system has so many components, problems occasionally are encountered. Typical problems include clogs in the tubing, tubing wear, bubbles interfering with the spectrometer measurements, and malfunctioning of the auto sampler. To some extent, the appearance of particulate matter or clumps of cells is a feature of the bacterial strain used in the experiments and cannot be blamed on the instrument. Problems with tubing wear result from our trying to re-use tubing many times to minimize expense. Bubbles in the system are a fairly rare problem, and this problem could possibly be solved by implementing a special washing routine that is triggered by a dramatic change in turbidity readings, such as occurs when a bubble is encountered. Overcoming the malfunction of the autosampler, which is also rare, is somewhat beyond the scope of our studies, and this problem could be solved simply by purchasing a better fraction collector.

Software

All of the components of the system (pumps, valves, spectrometer and autosampler) can be controlled from a personal computer. To implement the experiment routine to measure OD, adjust the flow and take samples we used Labview TM (Texas

instruments) programming language. Labview is the most widely used language to write software that involves the control of machinery. Every instrument that can be controlled from a computer nowadays comes with the associated Labview driver, which is a piece of software that allows the Labview programmer to interact with the instrument, similar to the way a printer driver works. We used the Labview drivers of the spectrometer to send commands and receive information from the spectrometer, as well as Labview drivers to send commands and receive information from the other components via serial ports.

The software implements the following algorithm indefinitely, until the user decides to stop (All the parameters can be set by the user):

1) Initialize by setting the valves in correct position, initializing the auto sampler, setting the media pump speed to 50% (pump speed is measured as a percentage of its top speed), take a sample from the culture with the sampling pump and measure the OD

2) Every 10 minutes, take a sample from the culture with the sampling pump and measure the OD. Adjust the media pump speed (and thus the flow of media in/ culture out) by using the following procedure: first, calculate the growth rate with the formula

$$OD = pOD\ Exp\left\{\left(\alpha - \frac{s}{100}\frac{f_{\max}}{V}\right)\Delta t\right\}$$

where *OD* is the optical density, *pOD* is the optical density in the previous reading, α is the growth rate, *s* is the speed of the pump (measured as a percentage), f_{max} is the

maximum flow (5.8 ml/min), V is the volume of the culture (250 ml) and Δt is the time interval between OD readings (10 minutes) This formula assumes exponential growth of the cells and exponential decay in culture concentration due to dilution, which are both good assumptions. Once the growth rate is calculated, the speed of the pump to be set can be calculated from the formula

$$tOD = OD\ Exp\left\{\left(\alpha - \frac{S}{100}\frac{f_{\max}}{V}\right)\Delta t\right\}$$

where tOD is the target OD set by the user, and S is the new pump speed to be set.

3) Every 30 minutes (every three OD readings), keep the sampling pump running after the taking OD measurement, so the flow is directed toward the auto sampler, where a sample is taken, and stored in a test tube inside a refrigerator.

4) Wash the flow cell by flowing through it 70% ethanol for 15 seconds, followed by flowing sterile water for 1 minute. Measure the light intensity that passes through the flow cell when there is only water in it. This measurement is made to assure that the flow cell is clean after each OD measurement.

Throughout the experiment, the system needs to have a media reservoir, as well as a 70% ethanol reservoir and a sterile water reservoir. One typical experiment requires 4 bottles of 1 gallon of media (16 litters), ½ a litter of 70% ethanol and 1 gallon of sterile water. With the formula for adjusting the flow that we used, the OD of the culture is kept fairly constant, typically within 5% of the target OD. The stored samples are assayed for enzymatic activity within 2 days of the sample being taken.

Automated system accuracy and repeatability

The results of clock runs published in (Atkinson et Al 2003) showed very smooth graphs of enzymatic activity of the clock strain vs. time. Can the automated system reproduce such smooth graphs for the same strain? Unfortunately the graphs resulting from enzymatic activity in the automated system are not as smooth as the old graphs. We believe the noise in the new graphs is not coming from the automated system in itself, but rather from the fact that the samples are being assayed after they have been stored in the refrigerator for a day or two. However, we have found this noise to be random in nature, as when a window average filter is applied to the data the graphs become very smooth. All the graphs presented in this thesis have been filtered with a 5 point size window average. When applying this filter to the data, we have found that the system is capable of good repeatability of an experiment with the same strain and environmental conditions, at least for the first couple of oscillations (Figure 2-7). In the 3 experiments shown in Figure 2-7, 2 of them were from successive weeks, while the third (and most different) of the three was from 7 months later. After the first couple of oscillations, bigger differences appear, particularly for the samples from a different time period. Some part of this divergence is likely due to differences in the assay materials or growth medium components, or slight differences in the temperature of the experiments owing to a lack of precision in our temperature-control mechanism. Our temperature control mechanism consists simply of an old bacteriological incubator with a small hole in the top, allowing fluid and air lines to be connected to the growth flask. Our stirring mechanism consists simply of

a stir bar in the growth flask and a magnetic stirrer placed under the flask. Small differences in temperature or stirring from run to run are therefore to be expected, and could be minimized with some further investment in the instrument.

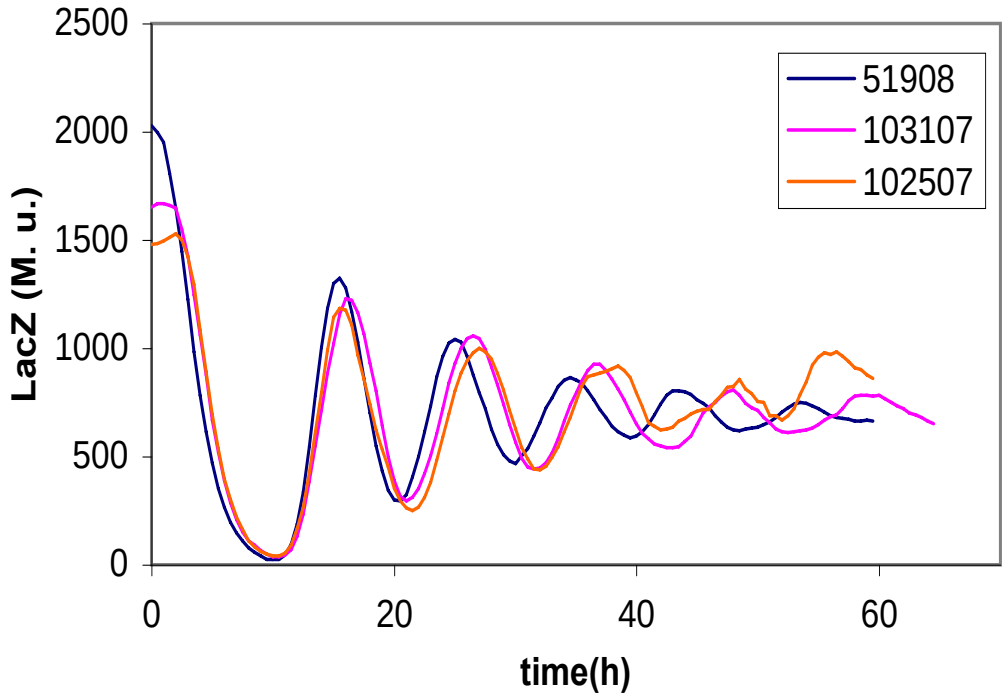


Figure 2-7 Repeatability of clock experiments. The three graphs shown correspond to three different experiments done with the same clock strain (NCPR77), under the same conditions, as a test of the repeatability of the experiment. The pink and orange graphs were done one week apart, while the blue graph was done seven months later than the other two. The graph shows that despite inevitable small differences in media composition, temperature, culture induction, etc. the repeatability is quite good, especially in the first two cycles.

Summary

We developed a turbidostat instrument that allows bacterial cultures to be maintained at constant turbidity for many generations. Unlike a typical chemostat used for continuous culture, where growth rate is limited by a key nutrient, our device allows continuous cultures to be maintained at constant turbidity at conditions where

cells are growing rapidly. This is because our device measures culture turbidity using a spectrophotometer, and uses this information to adjust the flow of nutrients into the continuous culture. Thus, our device will be useful for some continuous culture experiments that cannot be performed using a simple chemostat. Although we found several different discussions of turbidostats on the internet and mathematical models of cell growth and metabolism in a turbidostat, we did not find any source for a similar device. Thus, we believe that our device is unique. Our motivation in producing the instrument was to enable a single person to conduct experimental studies of the Ninfa clock. However, the device has already been put to additional uses by other researchers in our laboratory. For example, Rohit Vyas and Elizabeth LaSalle, undergraduate researchers in the lab, are using the instrument to monitor the slow kinetics of induction and de-induction of cells containing a synthetic genetic toggle switch. For this purpose, a culture is maintained at constant turbidity over many generations, while inducer concentration is gradually increased or decreased. An additional application of the device could be in studies of the evolution of novel bacteria in response to various challenges.

LIST OF REFERENCES CITED IN CHAPTER 2

Atkinson M R, Savageau MA, Meyers J and Ninfa AJ. *Development of a genetic circuitry exhibiting toggle switch or oscillatory behavior in Escherichia coli.* Cell 2003 113:597-607.

Dykhuizen D and hartl D. *Selection in Chemostats* Microbiology Reviews 1983, P.150-168 Vol. 47, No. 2

Monod J. *La technique de culture continue. Théorie et applications*. Ann. Inst. Pasteur (Paris). 1950;79:390–410.45.

Novick A and Szilard L. *Experiments with the chemostat on spontaneous mutations of bacteria*. Proc. Natl. Acad. Sci. USA. 1950;36:708–719.

Ferenci T. *Bacterial physiology, regulation and mutational adaptation in a chemostat environment*. Adv Microb Physiol. 2008;53:169-229.

Drake DR and Brodgen KA *Continuous-Culture Chemostat Systems and Flowcells as Methods to Investigate Microbial Interactions*. In *Polymicrobial Diseases*. Brogden KA and Guthmiller JM, editors. Washington (DC): ASM Press; 2002.

Kafkewitz D, Iannotti EL, Wolin MJ and M. P. Bryant *An Anaerobic Chemostat That Permits the Collection and Measurement of Fermentation Gases* Appl Microbiol. 1973 April; 25(4): 612–614.

CHAPTER 3

VARIATION OF CLOCK DYNAMICS WITH ENVIRONMENTAL CONDITIONS

Introduction

The first step in characterizing the behavior of the Ninfa clock was to evaluate its dynamics under different environmental conditions. We evaluated the clock dynamics by varying of three key environmental conditions: the "target OD" (the OD at which the turbidostat keeps the culture), the temperature, and the richness of the growth medium (percentage of casamino acids, which serve as carbon and nitrogen source).

The Ninfa clock was developed using well-understood transcription factors and cis-acting DNA sequences, but its function is not well isolated from other cellular systems. The host strain had the genes involved in clock dynamics (*glnL*, *glnG* and *lacI*) deleted such that the clock modules contained the only copy of these genes within the cell. In addition, the *glnL** mutant allele is used to provide the cell with an altered NRII* protein that brings about the phosphorylation of NRI regardless of cellular nitrogen status. However, the gene expression from the clock modules can be affected by other factors, known and unknown, such as amino acid biosynthesis, and global regulators [cAMP, acetyl phosphate (see Chapter 4)]. The products of the clock modules can in turn affect the expression of other genes (see Chapter 5 for the

effect of *glnG* on genes not involved in clock function). For this reasons, the complete isolation of the clock function from other aspects of metabolism is not possible. Since the internal metabolism of the cells changes with different environmental conditions, we wanted to explore the possibility that changes in environment could change the clock dynamics. Changes in clock dynamics could include changes in the clock period, in the amplitude of oscillations, the number of oscillations, or the steady state expression level. By examining different conditions, we hoped to identify conditions in which the clock would perform better, for example by having less damping of the oscillation. On the other hand, variations in the parameters which do not necessarily make the clock 'better" can also give insight into how parameters are correlated with environmental conditions (e.g. temperature), thus giving insight into the way oscillations arise in a synthetic network.

Methods

All the experiments performed in this chapter were performed with W-salts based growth media. W-salts medium consists of 10.5 g/L K_2HPO_4, 4.5 g/L KH_2PO_4, 0.16 g/L $MgSO_4$-$7H_2O$ and 0.04 g/L thiamine. W-salts need to be supplemented with a carbon source and a nitrogen source. For the experiments is this chapter, we used 0.5% (W/V) casamino acids as a carbon and nitrogen source, and also had 0.1% (w/v) glutamine present as another nitrogen source. The media was also supplemented with 0.04 g/L tryptophan, as some strains (used in other chapters) were tryptophan auxotrophs, as well as 100 ug/L ampicillin and 34 ug/L chloramphenicol. In the case

of experiments with different amounts of casamino acids, the concentration was reduced to 0.1% (w/v). The clock strain NCPR77 was used in all experiments.

For the experiments with different target ODs, the ODs were fixed simply by setting the target OD parameter in the controlling software of the turbidostat. For these experiments, the temperature was set to 33 ºC.

For the experiments with different temperatures, temperatures were set using the temperature control of the incubator that houses the growth chamber. This incubator is a bacteriological incubator with a small hole added to the top that serves as a port for the tubing necessary to connect the growth chamber to the other turbidostat components. The temperature control of the incubator does not allow an exact temperature to be set, but rather it gives a level of radiated heat inside the incubator; temperatures were measured with a thermometer and have an error range of about 1 ºC. The temperature inside the incubator had a slight vertical gradient, due to the hole at the top which serves as the tubing port. For the experiments with different concentrations of casamino acids, the temperature was set at 33 ºC.

Results

The set of experiments done with different target ODs is shown in Figure 3-1. We performed experiments with target OD_{600} of 0.08, 0.35 and 1.1. The experiments show differences in the amplitude of the oscillations, but no significant differences in the period. The doubling time was always in the range of 68-72 minutes. The oscillations in the experiment with target OD 0.08 showed small amplitude, and only two clear oscillations were observed before the clock reached a steady state. On the

other hand, the experiments with target OD 0.35 and 1.1 showed the expected number of oscillations (4-5, same as experiments in Atkinson et al 2003), and these two experiments showed oscillations with similar and higher amplitudes than were obtained when the target OD was 0.08. These experiments show that clock dynamics were similar when the target OD is 0.35 or 1.1, but the clock did not function as well

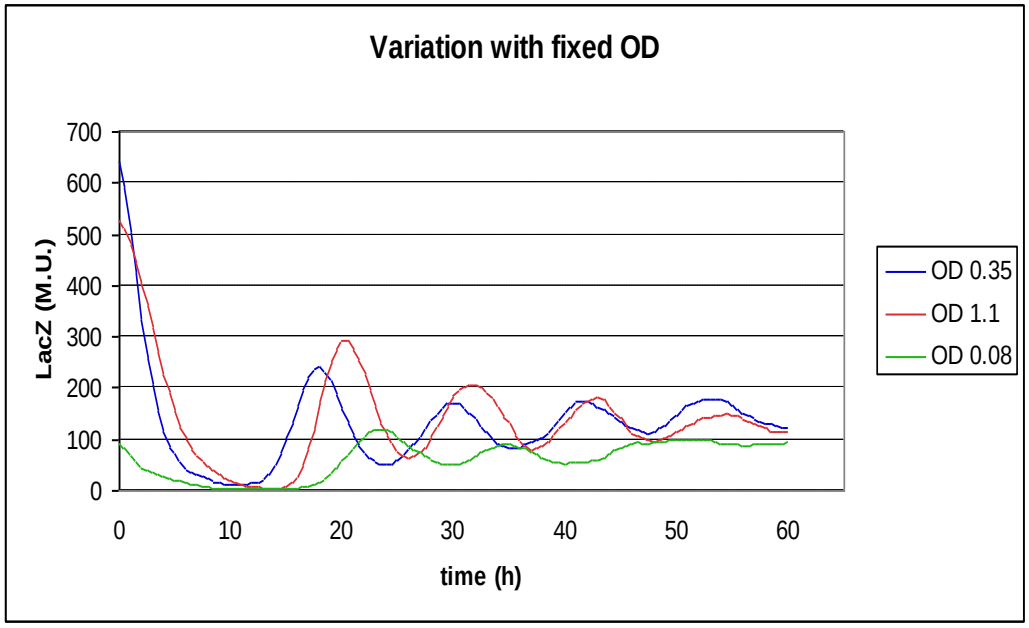


Figure 3-1 Comparison of clock dynamics with target OD. The clock dynamics seem to be better with higher target OD. With a target OD of 0.08, the clock dynamics show less oscillations with smaller amplitud.

when the target OD was reduced to 0.08. It should be noted that by setting the target OD between 0.08 and 1.1, we are covering the full range of accurate measurement of OD_{600} using our flow-through spectrophotometer cell. Any measurement significantly higher than 1 is not linear with concentration and any measurement lower than ~0.05 is not accurate due to intrinsic limitations of the optical measurement. However, this range of OD_{600} tested only covers the higher end in the

spectrum of possible bacterial concentrations in a liquid culture. In the clock medium, an OD of 1.0 corresponds to a concentration of ~10^8 cells/mL, whereas an OD of 0.1 corresponds to ~10^7 cells/mL. In terms of order of magnitude, the range 0.08-1.1 in OD only covers 1 order of magnitude of the 8 orders of magnitude of possible cell densities. Taking this limitation into account, the results seem to show that the dynamics are better with higher OD. This fact is used to support the theory of a signal that partially drives the synchronization of individual clocks in Chapter 4.

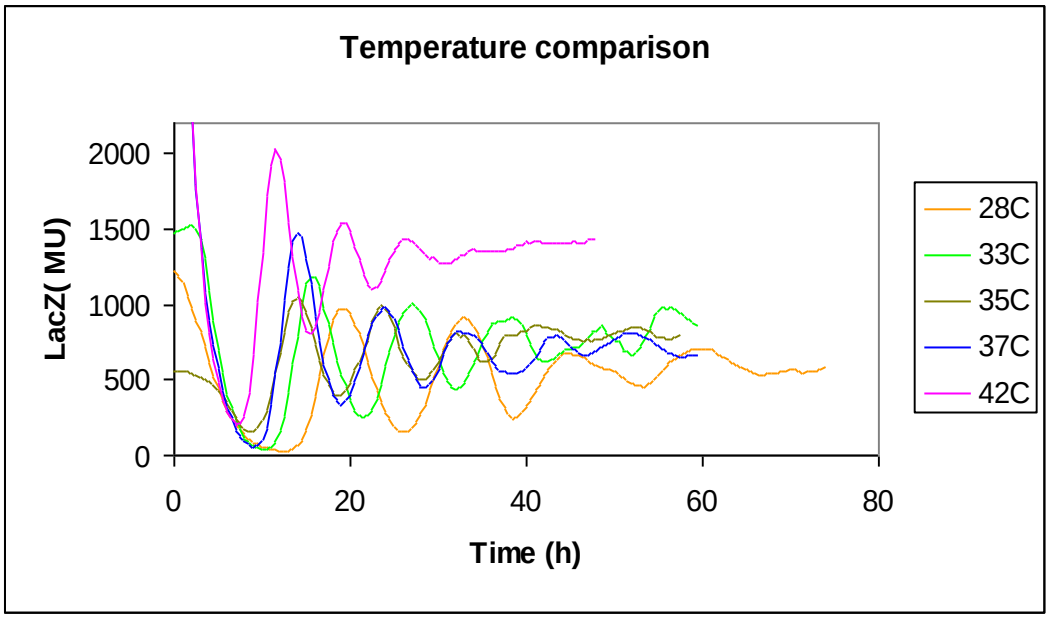


Figure 3-2. Clock Dynamics changes with temperature. The clock period is inversely proportional to the temperature used. The amplitude of oscillations does not change much with temperature, except when the temperature is 42 ºC. The steady state expression is also not changed with different temperatures, with the exception of 42 ºC.

The next set of experiments was done by varying temperature inside the incubator, while keeping the medium components and OD_{600} constant. The minimum temperature possible inside the incubator was 27 ºC, as this corresponds to room temperature plus a couple of degrees given by the heat from the stirring plate used to

provide mixing of the bacterial culture. As mentioned before, the temperatures measured have an error of +/- 1 ºC. Besides 27 ºC, the experiment was performed at 33 ºC, 35 ºC, 37 ºC and 42 ºC. Comparison of all five experiments is shown in Figure 3-2.

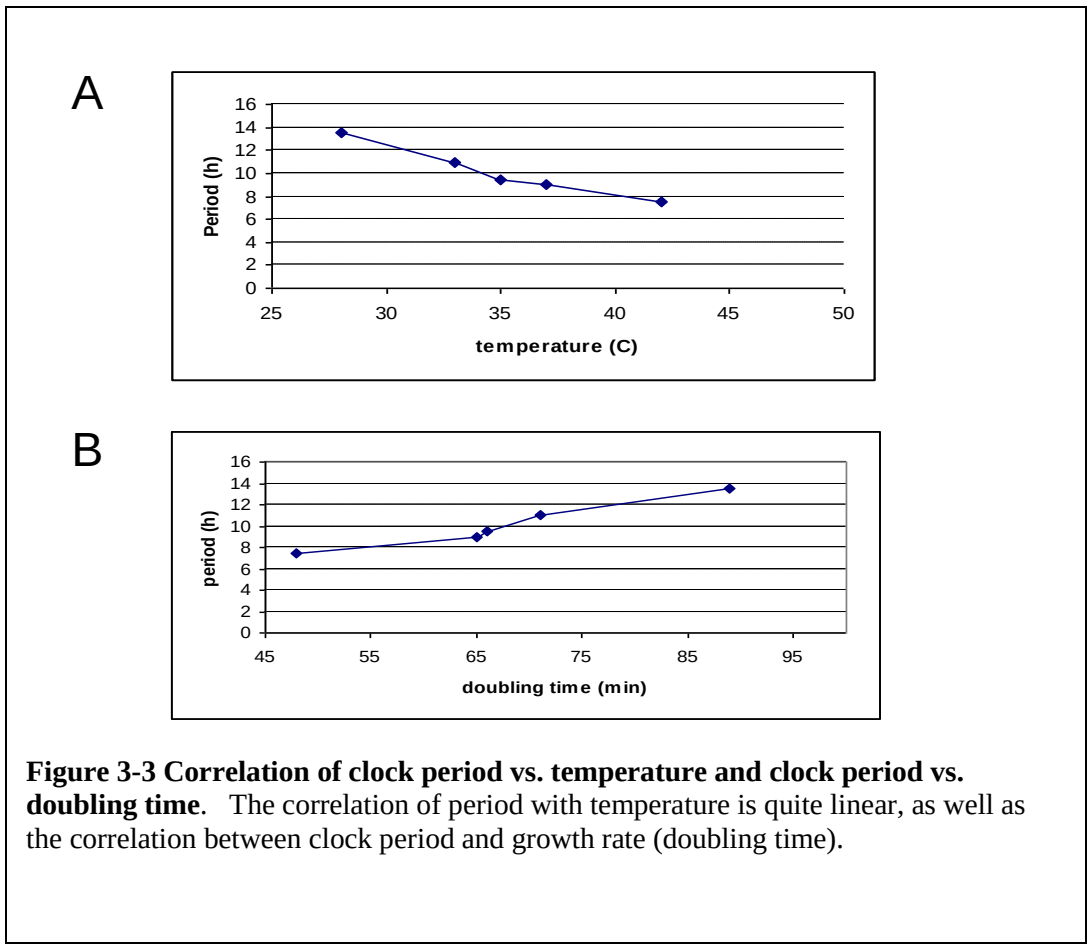


Figure 3-3 Correlation of clock period vs. temperature and clock period vs. doubling time. The correlation of period with temperature is quite linear, as well as the correlation between clock period and growth rate (doubling time).

The temperature used is highly correlated with the clock period. The clock period observed (+/- 0.5 hours) is 13.5 hours for 27 ºC, 11 hours for 33 ºC, 9.5 hours for 35 ºC, 9 hours for 37 ºC and 7.5 h for 42 ºC. This relationship between period and temperature appear to be linearly correlated; the relationship between period and growth rate (doubling time) is also quite linear (Figure 3-3), consistent with the fact that temperature and doubling time are linearly correlated in *E.coli* growth for a range

of temperatures in the interval 30ºC-37 ºC. The amplitude of oscillations did not vary dramatically between temperatures, with the exception of the experiment at 42 ºC. The experiment at 42 ºC is also an exception when comparing the steady state expression level; while the clock dynamics do not show much variation in steady state level for most temperatures examined, the experiment at 42 ºC had a steady state level more than twice the level of the other experiments. Possible reasons why the experiment at 42 ºC appears as an outlier are presented in the discussion section.

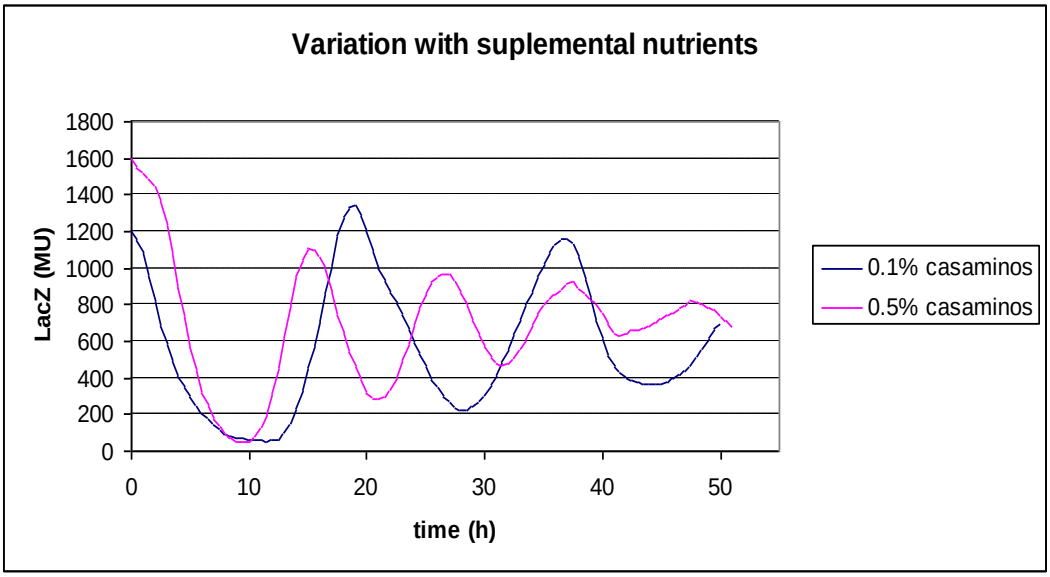


Figure 3-4. Comparison of clock dynamics with different concentrations of carbon source. When reducing the concentration of carbon source the clock period is increased, while the oscillations amplitude and steady state expression are not changed. This is consistent with the variation of period with doubling time obtained in the experiments with temperature variation.

In order to test the variability of clock period with growth rate, an experiment was performed by reducing the concentration of carbon source (casamino acids) from 0.5% to 0.1%, at a temperature of 33 ºC. The result was an increase in period from 11

hours to 18 hours (63% increase), while the doubling time increased from 70 minutes to 108 minutes. When this point is added to the clock period vs. growth rate trend, the linearity of the correlations was well maintained (Figure 3-5), supporting the trend that clock period is directly proportional to doubling time (and linearly inverse to growth rate).

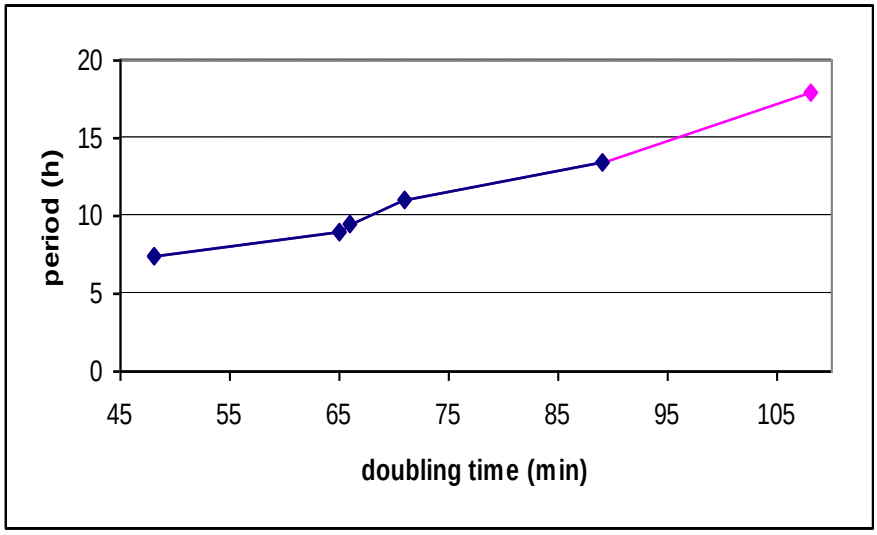


Figure 3-5. Linearity of clock period vs. doubling time. The linear trend in clock period vs. doubling time obtained in the experiments with temperature variation is maintained when adding the point corresponding to the experiment with low concentration of carbon source.

Discussion

In the set of experiments with different target ODs, it can be seen that the clock does not work well at low target OD. When the OD_{600} was 0.08, the clock oscillations had low amplitude and the expression settled to a steady state after just two oscillations. In contrast, the experiments with target OD 0.35 and 1.1 showed the expected number of oscillations (4-5) and the amplitude of oscillations was not very different between

them. Given that the amplitude of oscillations corresponds to the average *lacZ* expression per cell, the above observations suggest that more cells are oscillating in synchrony at higher target OD. When more cells have an oscillation peak simultaneously, the measured oscillation peak of the population is higher. In the case of a low OD_{600}, the lower amplitude of the oscillation peaks either means that the oscillations in individual cells are not synchronized, or that some cells do not express the reporter gene. In either case a measured peak in the population would be much lower, reflecting an average of expression between some cells at the peak of an oscillation and some cells in other phases of the oscillation cycle, or cells that do not express *lacZ*. If the lower peaks of oscillation measured in clock experiments with low culture OD are a result of individual clock de-synchronization, this would be consistent with a secreted signal that keeps individual clock synchronized (at least partially), in a way analogous to quorum sensing (Bassler, 2002). This possibility is explored further in Chapter 4.

When the dynamics of the clock were examined at different temperatures, the results strongly suggested an inversely proportional relationship between the clock period and temperature. This is consistent with the fact that, over a certain range of temperatures, *E. coli* grows faster with increasing temperature. The clock dynamics depend on cell growth for the dilution of activator and repressor proteins; this dilution is the main mechanism for decay of these proteins. With higher temperature, there would be faster growth and thus faster decay rates for activator and repressor, resulting in faster dynamics and a shorter oscillatory period. However, this

relationship is limited to temperatures where *E.coli* can grow; the metabolism of the cells at temperatures higher than ~42 ºC is significantly different, due to the heat shock response (Neidhardt et al, 1984; Strauss et al, 1987). Rather than a direct consequence of changes in temperature, changes in the clock period with temperature appear to be directly correlated to changes in growth rate. This was supported by the observed period of oscillation and growth rate when the cells were grown in a low carbon source concentration. Figure 3-5 shows a strong linear correlation between period and growth rate (doubling time), in all experiments with temperature variation and carbon source variation. It is known that temperature and growth rate are linearly correlated when cells grow in a particular medium in the range 30 ºC-37 ºC. Figure 3-2 shows the differences in clock dynamics due to changes in temperature. While the period of oscillation changes according to temperature, it can be seen that the changes in temperature do not greatly affect the amplitude of oscillations. This could mean that temperature is not a key factor in synchronization of the clocks in individual cells.

One particularity of Figure 3-2 is the experiment at 42 ºC. This experiment shows a higher steady state level of expression than the experiments at other temperatures, as well as a higher damping rate since only 3 oscillations are observed. The doubling time observed for this experiment was 48 minutes, which is lower than the doubling time of 65 minutes at 37 ºC. This contrasts with a wild type strain that would show larger doubling time at temperatures higher than 37 ºC, which is the optimal temperature for *E. coli*. This behavior may be explained by the fact that our genetic clock places a stress upon *E. coli* (by causing the expression of unnecessary

genes) and that this stress is reduced partially at 42 ºC. At 42 ºC the heat shock response is activated (Neidhardt et al 1984; Strauss et al 1987), and this causes the expression of a set of genes that includes chaperones and proteases. Perhaps the induction of a protease causes a shorter high life of the LacI protein, thus increasing the overall *lacZ* expression. The very high steady state in the experiment at 42 º C is consistent with reduced LacI activity at high temperature.

The variation of clock dynamics with growth rate is consistent with the fact that gene expression is correlated with growth rate in *E.coli* (For a recent analysis of this phenomenon, see Scott et al 2010). Faster growth rate (increased by temperature or nutrient conditions) leads to faster transcription rates that lead to faster clock dynamics, making the period shorter.

In summary, we found a strong linear correlation between doubling time (growth rate) and clock period. This linear correlation in the Ninfa clock contrasts with natural genetic clocks, which have a period of ~24 hours regardless of temperature changes, a characteristic known as temperature compensation (Johnson et al, 2008; Jolma et al, 2010). This difference, however, may prove useful for synthetic biology applications that require clock with different periods.

LIST OF REFENCES CITED IN CHAPTER 3

Atkinson M R, Savageau MA, Meyers J and Ninfa AJ. *Development of a genetic circuitry exhibiting toggle switch or oscillatory behavior in Escherichia coli.* Cell 2003 113:597-607.

Bassler BL *Small talk. Cell-to-cell communication in bacteria.* Cell 2002 May 17; 109(4):421-4.

Johnson CH, Mori T and Xu Y. *A cyanobacterial circadian clockwork*. Curr Biol. 2008 Sep 9;18(17):R816-R825

Jolma IW, Laerum OD, Lillo C and Ruoff P. *Circadian oscillators in eukaryotes* Wiley Interdiscip Rev Syst Biol Med. 2010 Sep-Oct;2(5):533-49.

Neidhardt FC, VanBogelen RA and Vaughn V *The Genetics and Regulation of Heat- Shock Proteins* Annual Review of Genetics 1984 Vol. 18: 295-329

Scott M, Gunderson CW, Mateescu E, Zhang Z and Hwa T *Interdependence of cell growth and gene expression: origins and consequences* Science 2010 330: 1099-1102

Straus DB, Walter WA and Gross CA *The heat shock response of E. coli is regulated changes in the concentration of sigma-32* Nature 1987 329, 348 - 351

CHAPTER 4.

ACETATE AS A SYNCHRONIZATION SIGNAL

Introduction

A remarkable feature of the Ninfa clock is that it works so well in large cell populations. Under our typical experimental conditions, the period of the clock is ~ 10-12 hr, and we are able to observe 4-5 oscillations before the amplitude of the oscillations becomes very low. This means that the cells in the population oscillate synchronously, or partially so, for about 3 to 4 days after the start of an experiment. This is a long time, corresponding to about 60 generations. By comparison, keeping a culture synchronized in cell cycles is impossible after a few cycles (Carl, 1970; Helmstetter and Cummings, 1963; Ferullo et al, 2009). Since the period of oscillation is correlated with doubling time (Chapter 3), the fact that cells remain partially synchronized for days brings up the question: is there a way in which the cells "signal" to each other to maintain synchrony?

Quorum sensing mechanisms allow cell-cell communication in bacteria. One way in which bacterial cells are known to communicate with one another is by quorum sensing (Bassler, 2002). Originally observed in *Vibrio fischeri* , quorum sensing is a mechanism by which bacteria communicate using a secreted signal, known as the autoinducer, which triggers a response from the same bacteria when the level of autoinducer in the environment reaches a certain threshold. In *Vibrio fischeri*,

the quorum sensing mechanism is used by the bacteria to produce luciferase (and thus light) when there is a dense bacterial population. If the bacterial population is low, the cells don't express the enzyme luciferase, whereas if the cell density reaches a threshold, all the cells express luciferase in unison to produce light. This mechanism consists of an autoinducer molecule (homoserine lactone) that is constantly secreted by the bacteria, plus a receptor protein that is able to measure the amount of autoinducer in the environment, and trigger an internal response whenever the autoinducer concentration reaches a certain threshold. Quorum sensing has been observed widely observed in many bacterial species as a mechanism to respond to various environments (Ahmer, 2004; Sauer et al, 2002; Chang et al, 2011). In *E. coli*, "synthetic" quorum sensing has been developed by introducing the genes for autoinducer synthesis and recognition from *Vibrio fischeri*, to engineer pattern formation (Basu et al, 2005) and more recently to construct a genetic clock that shows synchronized oscillations at the population level (Danino et al, 2010). Since quorum sensing is used by the bacteria to synchronize their gene expression, we wanted to evaluate if the Ninfa clock had a synchronization signal, analogous to the autoinducer in quorum sensing, that allows the population to oscillate in synchrony for a few days.

Acetate is a global signalling molecule in *E. coli.* The most likely candidate for a signaling molecule in *E. coli* would be acetate. Acetate is a molecule that can diffuse freely in and out of the cell, and it has been proposed as a global signal in bacterial gene regulation (McCleary et al, 1993). Acetate inside the cell is produced mainly from acetyl-coA by the *pta-ackA* pathway (Figure 4-1). Acetyl-CoA is transformed

reversibly to acetyl phosphate by the enzyme phosphotransacetylase (*pta*), and acetyl phosphate is transformed reversibly to acetate by the enzyme acetate kinase (*ackA*).

Acetyl phosphate is a phosphate donor that can phosphorylate many enzymes. Many regulatory responses to environmental conditions in bacteria are achieved by phosphorylation-dependent signal transduction pathways, known as two-component

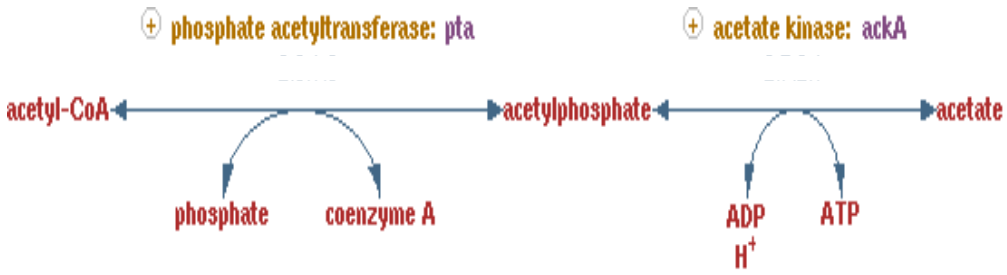


Figure 4-1. Acetate formation pathway. Taken from Ecocyc

systems. These systems consist of an enzyme (histidine kinase) that senses an environmental signal and an enzyme that acts as a response regulator by mediating the cellular response to the signal. The activation and inactivation of the response regulator is achieved by its phosphorylation and dephosphorylation, by means of the histidine kinase. In many cases, stimulatory effectors serve to reciprocally regulate these two activities of the histidine kinase protein, such that the response regulator kinase activity is increased by the stimulatory effector, while the response regulator~P phosphatase activity is simultaneously decreased by the signal. Alternatively, the stimulatory effector may increase the phosphatase activity and inhibit the kinase activity of the histidine kinase protein. Examples of well characterized two-component systems in *E. coli* are the phosphate starvation response system, the nitrogen assimilation control system, the osmolarity control system, and

the chemotaxis system. In each of these cases, acetyl phosphate is able to directly phosphorylate the response regulator proteins in the absence of the histidine kinase protein. Thus, acetyl phosphate can act as global regulator by interfering with the two-component systems environmental response. The purpose of such global regulation is unclear, but a few possible hypotheses for such signaling have been discussed (Mc Cleary 1993, Wolfe 2005). For example, acetyl phosphate may serve as a signal of internal stresses that cause the TCA cycle to function slowly, relative to glycolysis. A relatively slow flux through the TCA cycle may result in "spillover" from excessive glycolysis being directed to the acetyl phosphate pathway, with excretion of acetate. Acetate as a global signal for cell-cell communication has been proven to work in a synthetic network constructed for that purpose (Bulter et al, 2004), and there is evidence that acetate can be a global signal involved in biofilm formation in *E. coli* (Wolfe et al, 2003).

In the particular case of the nitrogen starvation response of *E. coli*, from which the activator module of the clock is based on, the histidine kinase is NRII, the product of *glnL*, and the response regulator is NRI, the product of *glnG*. Upon sensing low nitrogen levels inside the cell (by sensing the internal glutamine to ketoglutarate ratio), NRII phosphorylates NRI, and NRI~P activates the expression of genes involved in nitrogen assimilation. Conversely, in cells containing ample nitrogen, NRII dephosphorylates NRI~P. NRI can also be phosphorylated by acetyl phosphate, therefore activating the nitrogen starvation response in cells that lack NRII (Feng et al, 1992). Indeed, if cells contain wild-type NRII, the phosphatase activity of NRII in cells containing ample nitrogen results in the dephosphorylation of NRI~P,

regardless of the original source of phosphoryl groups on NRI~P. This causes a large decrease in cellular acetyl phosphate, that is dependent on both NRII and cellular nitrogen excess. Thus, the nitrogen status of the cell is signaled to other two component systems by NRII and its effect on the acetyl phosphate pool (Feng et al, 1992).

In the context of the Ninfa clock, the possible phosphorylation of NRI by acetyl phosphate implies that the auto-activation of the activator module could be driven in part by NRII, and in part by acetyl phosphate. If a significant percentage of the activator protein (NRI) phosphorylation was done by acetyl phosphate, then the theory that acetate could act as a synchronization signal could be envisioned: exogenous acetate could drive a few cells which are out of synchrony towards oscillating in synchrony with the majority. When the clock in the majority of cells goes to high level of expression, the phosphate in acetyl phosphate has been taken by NRI, then the majority of cells would secrete acetate and a "wave" of exogenous acetate could enter into cells that are out of synchrony and therefore their acetate levels are lower than average. This exogenous acetate can be converted into acetyl phosphate by the action of the acetate kinase (AckA) enzyme, and then acetyl phosphate would phosphorylate NRI, driving its own expression level to a high point, same as the majority of cells (Figure 4-2). According to this hypothesis, the AckA enzyme would act as a sensor of exogenous acetate, in an analogous way to the autoinducer receptor protein in *Vibrio fischeri*.

The idea that acetyl phosphate and NRII might work together to bring about the phosphorylation of NRI~P in the Ninfa clock is consistent with observations that

these two mechanisms to phosphorylate NRI~P play a role in the utilization of arginine as a nitrogen source. Growth of wild-type *E. coli* on arginine as sole

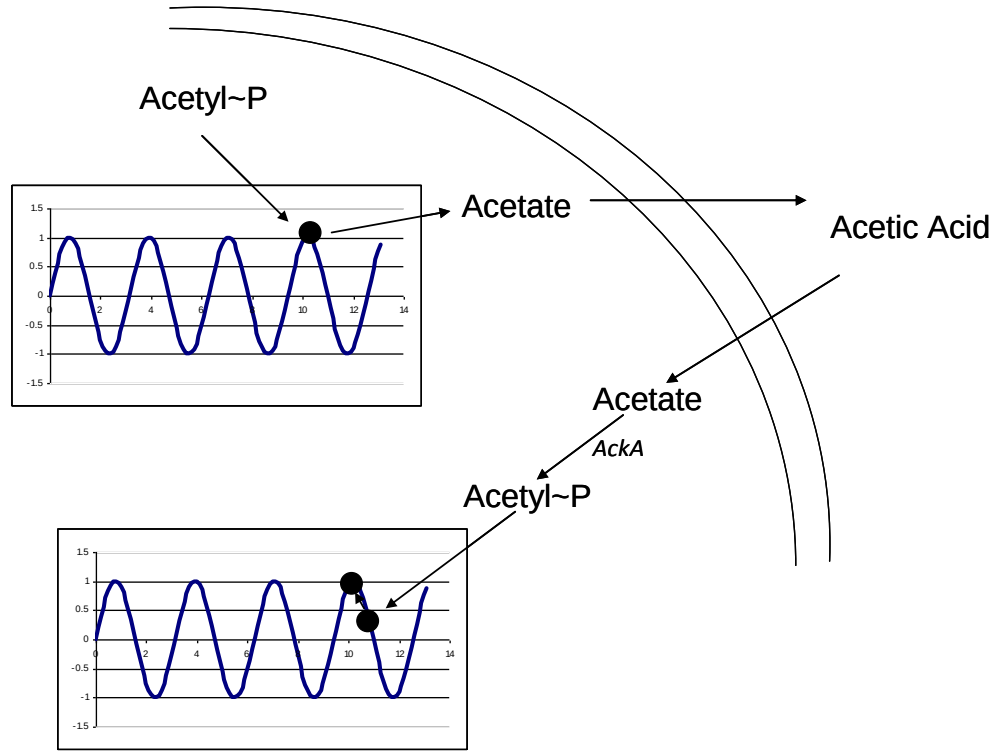


Figure 4-2. Acetate as a synchronization signal in the Ninfa clock. Acetate could drive the synchronization of expression in the clock activator module, with the enzyme *AckA* acting as a sensor of exogenous acetate.

nitrogen source requires both NRII and the ability to produce acetyl phopsphate (Feng et al, 1992). Presumably, the combination of both acetyl phosphate and NRII leads to higher NRI~P than either NRII or acetyl phosphate alone, and the activation of the *astC* operon required for growth on arginine requires a very high level of NRI~P.

In order to test the theory of acetate as a synchronization signal, I manipulated the genes of the acetate formation pathway in the clock strain, and modified environmental conditions that could favor or impede such a proposed cell-cell communication.

Methods

Mutations in the genes *ackA*, *pta* and a deletion mutation inactivating both genes (Δ*ackA-pta)* were introduced into the clock strain NCPR77, via P1vir transduction. The *ackA* null mutation was transuced from strain AJW1939 (Kumari et al, 2000), the *pta* null mutation was transduced from strain BW16167, and the *pta-ackA* deletion was transduced from strain BW16169 (Wanner and Wilmes-Riesenberg, 1992). Strain 3.30L*G was constructed by replacing the *glnL* deletion in strain 3.30LG (Atkinson et al, 2003) with a mutant *glnL* allele coding for an altered NRII that lacks the ability to dephosphorylate NRI~P (called *glnL**; Pioszak and Ninfa, 2003) , so the *glnALG* operon in strain 3.30LGPR becomes *glnAL*ΔG*. This replacement was done in two steps: first, the *glnA* deletion from strain TH16 (Reitzer and Magazanik, 1986) was transduced into strain 3.30LGPR to produce the strain 3.30ALGPR. The allele *glnL2302* (denoted here as *glnL**) from strain AP1009 (Piozak and Ninfa, 2003) was then transduced into strain 3.30ALGPR to produce strain 3.30L*GPR. The strain 3.30L*GPR was used then as the host to construct a new clock strain by transducing the activator and repressor modules from strain *NCPR77*. Strain NCPR101, with the *glnL** expressed from the chromosome, can be used in the clock

experiment without being transformed with the plasmid p3Y15. The complete set of strains used in this chapter is listed in Table 1.

Minimal media for clock experiments performed in this chapter was supplemented with different amounts of potassium acetate or pyruvate. Both of these chemicals were purchased from Sigma-Aldrich.

Strain	Relevant genotype	Reference
AJW1939	AckA51:kan	Kumari et al 2000
BW16167	pta-200 zej-223::Tnl0	Wanner and Wilmes-Riesenberg 1992
BW16169	Δ(pta ackA hisQ hisP) zej-223::Tnl0	Wanner and Wilmes-Riesenberg 1992
TH16	glnA::tn5	Reitzer and Magazanik 1986
AP1009	glnL*	Piozak and Ninfa 2003
3.30LGPR	ΔglnLG, phage resistant	This study
3.30ALGPR	glnA::tn5 ΔglnLG	This study
Ncpr77	Clock strain	This study
Ncpr101	As ncpr77, glnL*	This study
Ncpr77 ackA-	As ncpr77, AckA51:kan	This study
Ncpr77 pta-	As ncpr77, pta-200 zej-223::Tnl0	This study
Ncpr77 pta- ackA-	As ncpr77, Δ(pta ackA hisQ hisP) zej-223::Tnl0	This study

Table 4-1 Strains in chapter 4

Results

The first set of experiments related to the theory of acetate as a synchronization signal was already shown in Chapter 3. These experiments are the comparison of clock dynamics depending on the fixed culture turbidity (optical density, OD_{600}) for each run. I fixed the optical density of clock experiments by using the software of the turbidostat instrument to force bacterial cultures to be maintained at the desired OD indefinitely. In order to make a first evaluation to see if the there is a global signal that allows the cell to communicate and keep the synchrony of their gene expression for a few days, in a way similar to quorum sensing, I tested to see if the clock

performed “better” (in terms of amplitude or decay rate) in a culture with high OD than in a culture with low OD. If acetate acts as a quorum sensing signal, then cell-cell communication below a certain threshold of cell density would not work, since the signal is not strong enough, but it would work with cell densities higher than the threshold. In principle, the clock should work better with high cell densities and less good with low cell densities in a gradual way; in other words, the higher the cell density, the better the oscillations.

As shown in Figure 3-1 (Chapter 3), we compared clock runs with fixed OD_{600} of 0.08, 0.35, and 1.1. We could not try experiments with fixed OD lower than 0.08, since an OD_{600} lower than that could not be read accurately in our flow-cell spectrometer. The experiment at OD 1.1 showed a first peak of height 293 MU (Miller units), whereas the experiment with fixed OD of 0.35 showed a first peak of 240 MU, and the experiment with fixed OD of 0.08 showed a first peak of 117 MU. It can be seen from the graph that the clock strain run at 0.08 OD only showed 2 small oscillations before settling to a steady state, while the strains run at OD 0.35 and 1.1 showed the typical 4-5 oscillations of reporter activity. Since Miller units are an expression of specific activity, that is, the average expression per cell, the results show that at the highest point in expression the cells in the experiment at 0.08 OD have less than half of overall expression than the cells in the experiment at OD 0.35. This likely indicates that in the experiment at fixed OD 0.08, a bigger percentage of cells are not showing high expression, i.e. not “participating” in the synchronized clock oscillations. The difference in heights of the first peak in runs with fixed OD

0.35 and 1.1 is smaller, but this little difference also supports the idea that cells are in better synchrony with higher density (Chapter 3).

The next step in testing the theory of acetate as a global synchronization signal was to introduce exogenous acetate in the growth medium. If acetate is present in the growth medium, then the signaling could be weaker or even lost since the exogenous acetate that the cells "sense" would not correspond to the activity of the majority of cells. Depending on the concentration of exogenous acetate, the changes in the level of acetate concentration due to the clock activity of the cells would be less noticeable with higher level of acetate in the growth medium, and a high level of acetate in the growth medium would suppress completely the cell-cell communication. We tried clock experiments with strain NCPR77 and with exogenous acetate concentrations of 0.05% (0.5 g/l) or 1% (10 g/l). A low concentration of exogenous acetate (0.05%) showed no drastic effect in the clock dynamics when compared to the dynamics with the regular clock media (Figure 4-3). However, the run with low exogenous acetate showed lower amplitude of the expression peaks (overall ratio of corresponding expression peaks of ~1.4), which is consistent with the theory of acetate as a communication signal, since the exogenous acetate would lessen the effect of acetate concentration changes and reduce the overall expression in the peaks, as explained above. It should be noted that the period in the run with exogenous acetate is a little higher, but this was expected since the growth rate of the cells would be lowered by the exogenous acetate.

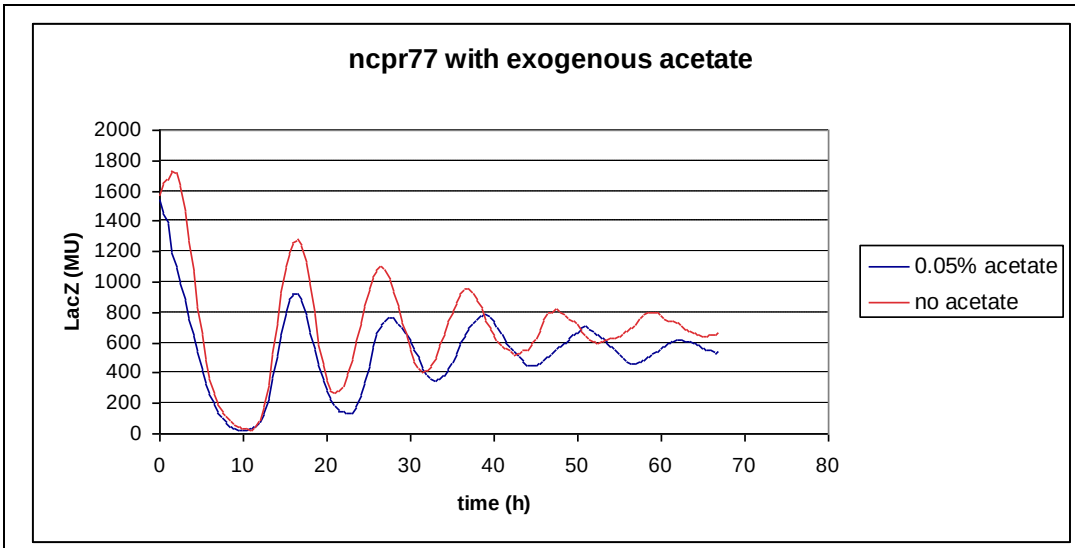


Figure 4-3. Comparison of clock dynamics with or without 0.5 g/l exogenous acetate in the growth media. Lower amplitudes in the run with acetate in the media are consistent with the theory of acetate as a communication signal, since exogenous acetate would produce less synchronization between individual cells, thus lowering the overall expression in the peaks

When a high concentration of acetate was introduced in the growth media, the oscillations were almost completely suppressed (Figure 4-4). The growth rate with 1% of exogenous acetate in the media was noticeable decreased (the doubling time changed from ~70 min to 110 min). It should be noted that the decrease of the growth rate alone cannot be the cause of loss of oscillations, since slow growth rates in NCPR77 (due to media composition or low temperature) did not "damage" the clock oscillations, as shown in Chapter 3. The very small oscillations in this run are consistent with the theory that a saturating concentration of exogenous acetate in the growth medium would eliminate any cell synchrony, since the changes in acetate concentration due to clock activity would be minimal.

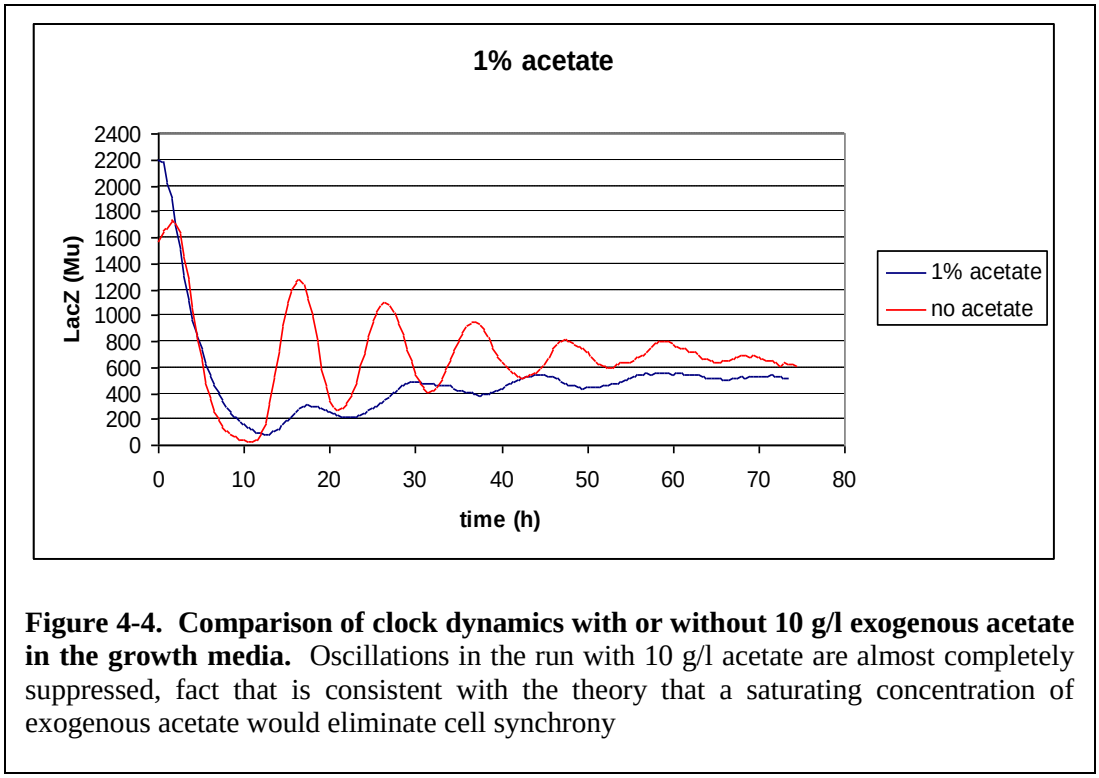


Figure 4-4. Comparison of clock dynamics with or without 10 g/l exogenous acetate in the growth media. Oscillations in the run with 10 g/l acetate are almost completely suppressed, fact that is consistent with the theory that a saturating concentration of exogenous acetate would eliminate cell synchrony

We then manipulated the genes in of the acetate formation pathway, *ackA* and *pta*, in strain NCPR77, to see if these genes affected clock dynamics. According to the hypothesis, *ackA* would affect the clock by phosphorylating exogenous acetate to produce acetyl phosphate, which in turn would increase the expression of the activator protein. According to this hypothesis, the deletion of *ackA* would eliminate the cell synchrony.

We found that when the *ackA* gene is deleted from strain NCPR77, the clock oscillations were completely suppressed (Figure 4-5). After the start of the clock experiment, the level of expression in the cells went from a high level toward a steady state level quickly, without producing a single oscillation. The steady-state level that is obtained corresponded to the peaks of LacZ activity in the related strain that had the wild-type *ackA* gene, thus loss of *ackA* seems to result in low activator and low

expression of the lacI repressor. This is consistent with the proposed role of *ackA* as an exogenous acetate sensor. As a confirmation of this experiment, the strain NCPR77-*ackA* was transduced with wild type phage and selected for growth on acetae (as a way to recover the wild-type *ackA* gene). The resulting strain, genetically identical to NCPR77, showed the same oscillations as NCPR77 (not shown). This series of experiments showed that *ackA* was necessary for oscillations by the Ninfa clock.

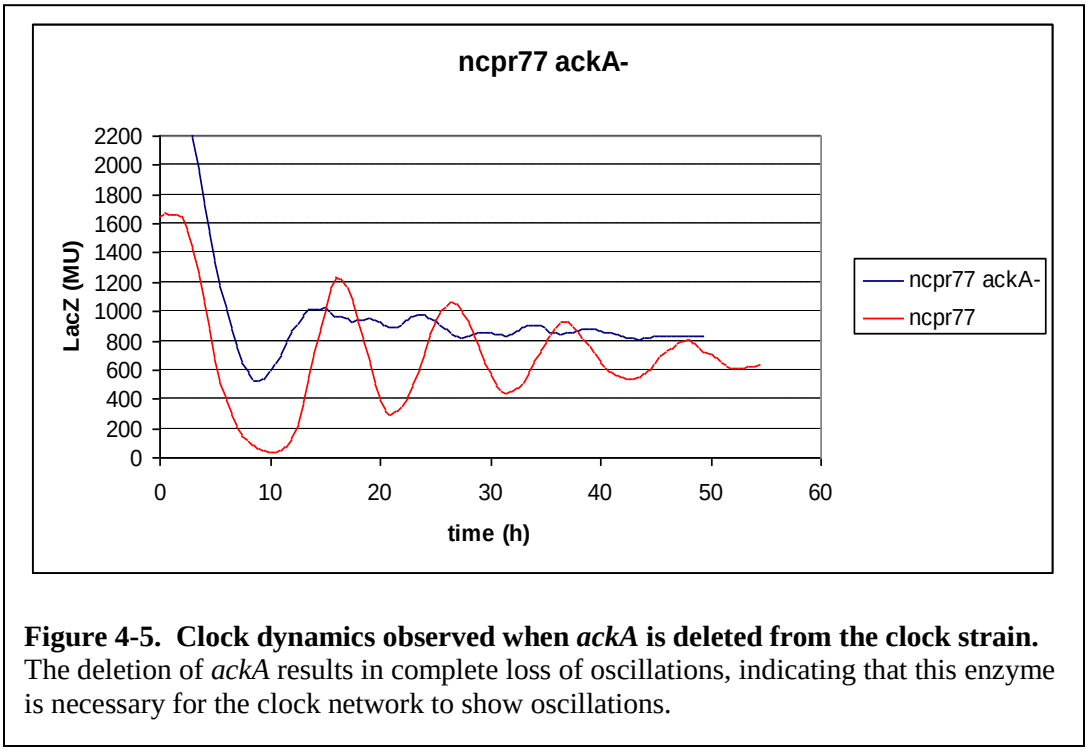


Figure 4-5. Clock dynamics observed when *ackA* is deleted from the clock strain. The deletion of *ackA* results in complete loss of oscillations, indicating that this enzyme is necessary for the clock network to show oscillations.

When the gene *pta* was deleted from strain NCPR77, the oscillations where not lost (Figure 4-6). However, the clock dynamics changed significantly. With the *pta* deletion, the amplitude of the peaks did not decay exponentially as before, but instead it goes up after the first peak only to go back down again. Another major

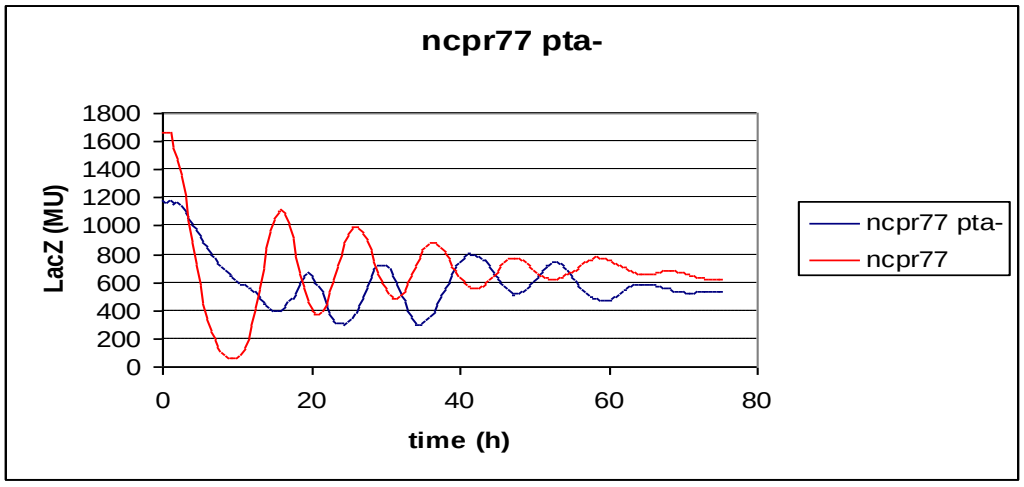


Figure 4-6. Clock dynamics observed when *pta* is deleted from strain NCPR77. The oscillations show a quite different dynamics than NCPR77, but oscillations were not lost, and the same number of oscillations were observed. This indicates that, contrary to *ackA*, the enzyme encoded by *pta* is not necessary for the clock network to show oscillations

difference is that the expression level did not go to zero after the start of the experiment, as it was the case with NCPR77. But, like with NCPR77, five clear

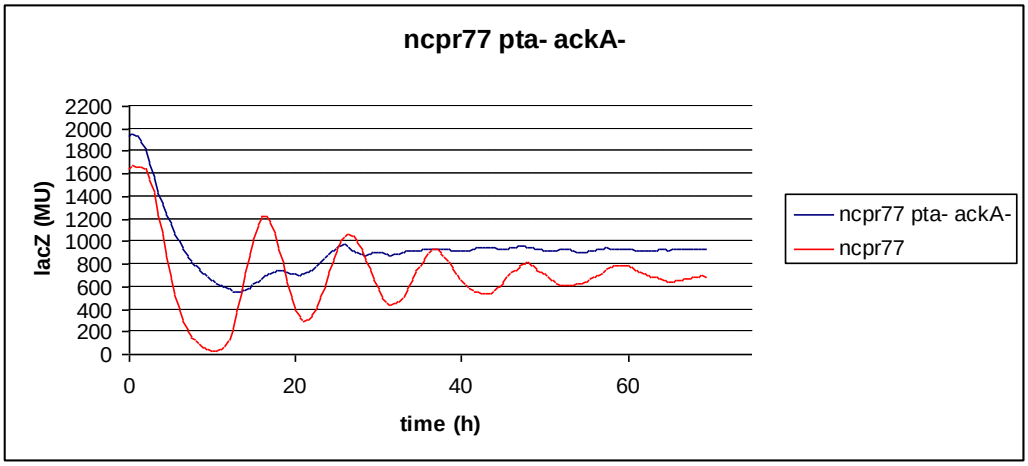


Figure 4-7. Clock dynamics observed when both *pta* and *ackA* are deleted from strain NCPR77. In this case, similar to the deletion of *ackA* alone, the double mutation completely suppressed clock oscillations. This is consistent with the observation that *ackA* is necessary for oscillations.

oscillations in reporter activity were obtained. These results show that *pta* was not strictly required for oscillations, although it obviously plays some role.

When both *ackA* and pta were deleted, no oscillations were present and the result of the clock experiment was a graph similar to the one with the *ackA* deletion (Figure 4-7). This again shows that wild-type *ackA* was required for oscillations by the Ninfa clock. **Small reduction in the level of NRII* did not eliminate oscillations.** According to our hypothesis, the deletion of both *ackA* and *pta* enzymes would eliminate synchrony by dramatically reducing acetyl phosphate inside the cells, since the deletion of *pta* assures that acetyl phosphate cannot be obtained from acetyl coA and the *ackA* mutation assures that it cannot be obtained from acetate. Without acetyl phosphate in the cells, the activator protein would lose one of its sources for phosphorylation. We wandered if the loss of a source of phosphorylation could be a main factor in the loss of clock oscillations, as observed in the strain NCPR77 *ΔackA-pta*. To evaluate this possibility, we compared the clock dynamics of strain NCPR77 with the strain NCPR101, which is a strain whose only difference with the clock strain NCPR77 is that the *glnL** allele (only kinase activity, no phosphatase activity) is recombined into the chromosome, as opposed to the strain NCPR77 that was always run with the multicopy plasmid p3y15 that expresses the *glnL** allele constitutively. The strain NCPR101, with *glnL** in the chromosome, would express less NRII than NCPR77 with the plasmid p3y15, simply by the multiple copies of the plasmid inside a single cell, as opposed to the single copy in the chromosome. Thus the strain NCPR101 would have less phosphorylation of the activator enzyme than NCPR77 containing p3y15.

The idea that we could control the extent of NRI phosphorylation by moving NRII* from a plasmid to the chromosome comes from the work of Dr. Dong-Eun Chang in this laboratory (Chang et al, 2010). Chang investigated the hysteresis of synthetic

genetic toggle switches based upon the activator module of the Ninfa clock. He observed significantly greater hysteresis for one of these toggle switches when NRII* was present on p3Y15, as compared to when NRII* was on the chromosome (Chang, unpublished data). Thus, it seems that greater phosphorylation was obtained when the kinase was on a multi-copy plasmid, which is as expected.

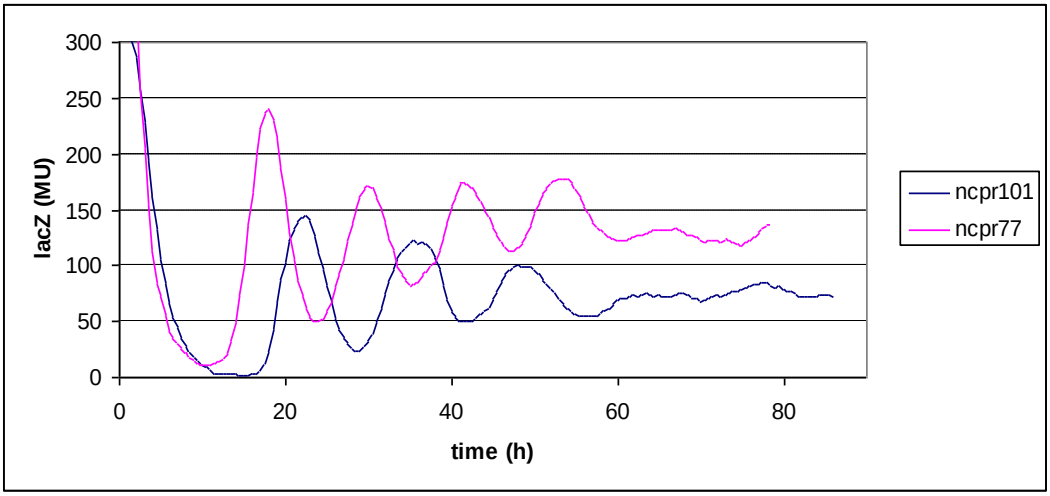


Figure 4-8. Comparison of clock dynamics when the concentration of NRII* inside the cell is changed by placing *glnL in the chromosome (NCPR101) or on a multi-copy plasmid (NCPR77).** The strain NCPR101 has less NRII* inside the cell, so the rate of NRI phosphorylation is diminished. A reduction in the phosphorylation rate of NRI (activator protein) does alter the oscillations negatively, since the overall height of the peaks is lowered, but such a reduction did not eliminate the oscillations.

We found that with the *glnL** allele in the chromosome the dynamics of the clock changed, but not dramatically (Figure 4-8). This indicates that loss of some

phosphorylation of the activator enzyme cannot be the main cause of the loss of oscillations in strain NCPR77 *ackA pta* or NCPR77 *ackA*.

The theory of acetate as a communication signal says that cells keep synchrony due to acetate; however the cells are only partially synchronized, as we see in every single clock experiment that the expression eventually settles to a steady state. If the theory of acetate as a signal holds, could it be possible to improve synchronization by acetate, thus keeping the system oscillating for a longer time? One possible way to do this would be to eliminate the phosphorylation of NRI by NRII, making the phosphorylation of NRI only possible with acetyl phosphate. To evaluate this possibility we performed a clock run with NCPR77 without the plasmid p3y15, so there is no NRII* involved in NRI phosphorylation. In addition to running

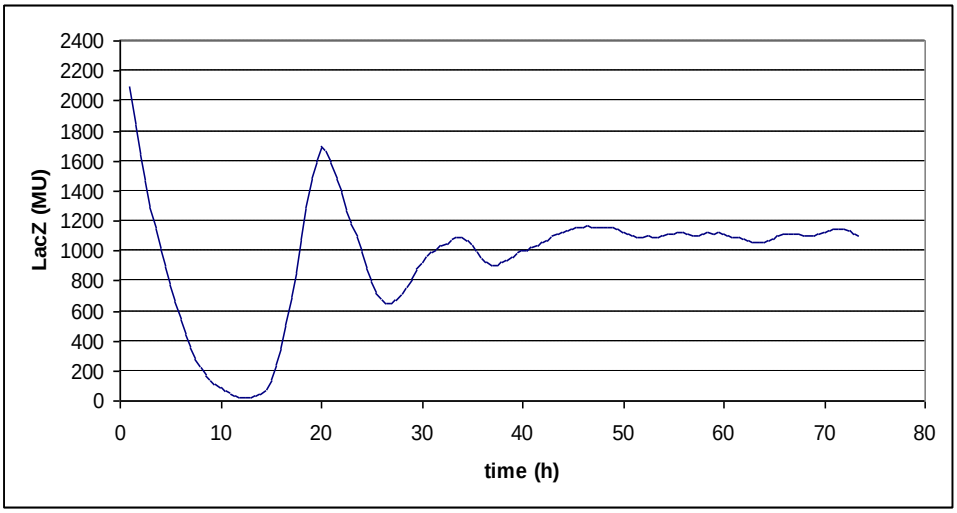


Figure 4-9. Clock dynamics with the strain ncpr77 without the plasmid p3y15 there is no glnL* in the clock network and all NRI phosphorylation is due to acetyl phosphate. To provide excess of internal acetate, the growth medium was supplemented with 0.4% sodium pyruvate.

the strain without NRII*, we supplemented the media with 0.4% sodium pyruvate, so the cells would have excess of acetate in the medium (due to the pyruvate oxidative pathway). The result of this experiment was a clock with one oscillation followed by a low amplitude oscillation before the system settles to a steady-state (Figure 4-9). The results were not as expected in the sense that the oscillations did not improve. However, this run shows that oscillations are possible without the need to have *glnL** in the clock network. The phosphorylation of NRI by acetyl-phosphate was enough to produce a couple of oscillations.

Discussion

The experiments performed with different fixed ODs gave the first evidence that the bacteria growing in the culture may have some signaling molecule analogous to the signal in quorum sensing, by showing that the clock dynamics are "better", in terms of the oscillating amplitude and number of oscillations, under conditions of higher cell density. Consistent with a signaling molecule driving synchronization, the clock would perform better by having more signaling molecule in the environment, as would be the case with higher cell density. It would have been desirable to measure the clock activity at lower ODs than 0.08, to see it the oscillations were completely gone at low ODs, however this was not possible, since lower ODs than 0.08 are not possible to measure accurately in our flow-cell spectrometer. This prevented us from using the automated system to keep a culture growing at such OD. An OD of 0.1 corresponds roughly to a cell density of 10^8 cells/ml, so the spectrometer measurements leave out the possibility of measurements of cell density for about 8 orders of magnitude, while the ODs measured in the spectrometer only cover 1 order

of magnitude of cell density. This was an intrinsic limitation of the experimental methods.

We decided to investigate the possible role of acetate as a signaling molecule due to previous evidence of acetate playing a role in the collective behavior of *E.coli* biofilms (Wolfe et al 2003) and the previous knowledge that acetyl-phosphate could phosphorylate NRI, independently of NRII (Feng et al, 1992; Mc Cleary et al, 1993), as well as the fact that an synthetic cell-to-cell signaling was previously constructed (Bulter et al 2004). If acetate was the signaling molecule needed for synchronization, then adding exogenous acetate to the growth media would interfere with cell-to-cell signaling by reducing or eliminating the correlation of acetate in the medium with cell density. As shown in the results section, we obtained results consistent with the theory: when adding a small concentration of potassium acetate (0.05% = 5 mM) to the media the amplitude of oscillations was reduced, consistent with a lower fraction of cells showing oscillations in synchrony (since we assume that the expression at the peak of an oscillation in the culture correspond to most cells expressing the peak level of the reporter), and when adding a large concentration of potassium acetate (1% = 100 mM) the oscillations were almost completely suppressed, consistent with saturation of the signal in the medium leading to loss of synchrony. These results support the hypothesis of acetate as a synchronization signal, however it should be noted that adding acetate to the medium alters cell metabolism and is deleterious to cell growth (Roe et al, 2002). Given this, the possibility that clock oscillations are modified or suppressed by metabolic changes cannot be excluded. However, as seen in Chapter 3, slow growth alone cannot be the case for loss of oscillations. Slow

growth seems to increase the period of oscillation, but it does not change the amplitude of oscillations.

Given the initial evidence to support the theory of acetate as signaling molecule, we proceeded to manipulate the clock strain by manipulating the two genes involved in the acetate formation pathway. The deletion of *ackA* completely suppressed oscillations, while transducing back wild type *ackA* to the clock strain with *ackA* deletion resulted in the original oscillating dynamics. This indicates that *ackA* is necessary for oscillations, consistent with the theory of acetate as a signal, which postulates *ackA* as the sensor of exogenous acetate. Without *ackA* the cells would lose the possibility of communication.

When *pta* was deleted from the clock strain, oscillations were not lost, but the oscillation dynamics were significantly changed. This indicates that while *ackA* is necessary for oscillation, the acetate formation pathway as a whole is not, even though cells without the acetate formation pathway have a somewhat altered metabolism. The loss of the acetate formation pathway means that the cells lose the main source of acetate, so how could acetate be a signal if is not being produced by the cells? The reality is that acetate can be produced by other means, such as the pyruvate oxidative pathway and some amino acid synthesis pathways (Bulter et al 2004). The fact that the cells without *pta* lose the acetate formation pathway does not contradict the theory of acetate as a signal, because acetate (and acetyl phosphate through acetate conversion by *ackA*) can be produced by the cells by other means.

The deletion of *ackA-pta,* eliminating both genes, confirmed the observation that *ackA* is necessary for oscillations; however in this case, different from the *ackA*

only deletion, the cells cannot produce acetyl phosphate. This means that the phosphorylation rate of NRI is reduced, and is only achieved by NRII*, something that could harm the dynamics of the clock network in individual cells, and not only harm the synchronization of clocks.

Since the manipulation of the *ackA* and *pta* genes alters the internal concentration of acetyl phosphate in the cells, the rate of phosphorylation of NRI is also altered, due to the availability of acetyl phosphate as a phosphate donor. We wanted to explore the possibility that changes in the phosphorylation rate of NRI could produce loss of oscillations or drastic changes in dynamics. To do this, we compared clock strains with the *glnL** allele inserted in the chromosome or expressed from a multi-copy plasmid. Due to the difference in the copy number of *glnL** such strains, the strain with multiple copies (plasmid) would have a higher internal concentration of NRII, thus a faster rate of NRI phosphorylation. Given the results in clock dynamics show in Figure 4-8, it can be said that the changes in NRI phosphorylation rate do change the clock dynamics, but not in a drastic way, and it would be hard to imagine that a small loss of phosphorylation rate due to acetate could be the cause of complete loss of oscillations.

Given good evidence that acetate could be a synchronization signal, we thought about a way to improve the synchronization mechanism. One way would be to remove NRII from the clock genetic network, thus making the phosphorylation of NRI dependent on acetyl phosphate only. This in principle would make phosphorylation of NRI more responsive to the level of acetate in the medium, and that would produce better synchronization between cells. As shown in Figure 4-9, the

results of NRII removal from the network were not as expected, since the experiment only produced two oscillations. However this experiment shows that phosphorylation of NRI due to acetyl phosphate is far from negligible, so this reinforces the role of acetate in the clock dynamics.

Overall, this set of experiments present strong evidence that acetate and acetate metabolism are an important part of the clock genetic network, something that was not intended when the clock was designed. The evidence for acetate as a signaling molecule is somewhat less strong; the fact that *ackA* is necessary for oscillations while *pta* is not does support the theory, but the effect of the manipulation of *ackA* and *pta* could well be indirect, as we noted that acetyl phosphate could interfere with many different two component systems, therefore many aspects of metabolism. It would have been desirable to have a way to assay the acetate in the media at various stages of the different experiments, to observe that in fact the level of acetate in the medium was oscillating in correlation with the clock oscillations. We researched for ways to measure acetate in liquid samples, but we found that all of the available technologies for such are measurement were cost-prohibitive to implement in the lab.

LIST OF REFERENCES CITED IN CHAPTER 4

Ahmer BM *Cell-to-cell signalling in Escherichia coli and Salmonella enterica* Mol. Microbiol. 2004 52(4), 933–945

Atkinson M R, Savageau MA, Meyers J and Ninfa AJ. *Development of a genetic circuitry exhibiting toggle switch or oscillatory behavior in Escherichia coli.* Cell 2003 113:597-607.

Bassler BL *Small talk. Cell-to-cell communication in bacteria.* Cell 2002 May 17; 109(4):421-4.

Basu S, GerchmanY, Collins CH, Arnold FH and Weiss R. *A synthetic multicellular system for programmed pattern formation* Nature 2005434, 1130-1134

Bulter T, Lee SG, Wong WW, Fung E, Connor MR and Liao JC *Design of artificial cell–cell communication using gene and metabolic networks* Proc Natl Acad Sci U S A. 2004 February 24; 101(8): 2299–2304

Carl PL. *Escherichia coli mutants with temperature-sensitive synthesis of DNA.* Mol Gen Genet. 1970;109:107–122

Chan KG, Atkinson S , Mathee K , Sam CK, Chhabra SR , Camara M , Koh CL and Williams P *Characterization of N-acylhomoserine lactone-degrading bacteria associated with the Zingiber officinale (ginger) rhizosphere: Co-existence of quorum quenching and quorum sensing in Acinetobacter and Burkholderia*. BMC Microbiology 2011 11: 51. doi:10.1186/1471-2180-11-51.

Danino T, Mondragón-Palomino O, Tsimring L and Hasty J *A synchronized quorum of genetic clocks* Nature 2010 463, 326-330

Ecocyc, E.coli bioinformatics database. Nucleic Acids Research 37:D464-70 2009 http://ecocyc.org

Feng J, Atkinson MR, McCleary W, Stock JB, Wanner BL and Ninfa AJ *Role of phosphorylated metabolic intermediates in the regulation of glutamine synthetase synthesis in Escherichia coli.* J Bacteriol. 1992 October; 174(19): 6061-6070

Ferullo DJ, Cooper DL, Moore HR, and Lovett ST *Cell cycle synchronization of E. coli using the stringent response, with fluorescence labeling assays for DNA content and replication* Methods 2009 48: 8–13

Helmstetter CE, Cummings DJ. *Bacterial synchronization by selection of cells at division* Proc Natl Acad Sci U S A. 1963;50:767–774

Kumari, S, Beatty CM, Browning DF, Busby SJ, Simel EJ, Hovel-Miner G and Wolfe AJ *Regulation of acetyl coenzyme A synthetase in Escherichia coli*. J.Bacteriol. 2000 182:4173-4179

Lewis Sauer K. Camper A, Ehrlich G, Costerton J and Davies D. *Pseudomonas aeruginosa displays multiple phenotypes during development as a biofilm*. Journal of Bacteriology 2002 184 (4): 1140–1154. doi:10.1128/jb.184.

McCleary WR, Stock JB and Ninfa AJ *Is acetyl phosphate a global signal in Escherichia coli?* J Bacteriol. 1993 May; 175(10): 2793–2798.

Pioszak AA and Ninfa AJ. *Genetic and biochemical analysis of phosphatase activity of Escherichia coli NRII (NtrB) and its regulation by the PII signal transduction protein* J Bacteriol. 2003 Feb;185(4):1299-315

Reitzer LJ and Magasanik B *Expression of glnA in Escherichia coli is regulated at tandem promoters.* Proc. Natl. Acad. Sci. USA 1985 82:1979-1983. .

Roe AJ, O'Byrneb C, McLagganD and BoothIR *Inhibition of Escherichia coli growth By acetic acid: a problem with methionine biosynthesis and homocysteine toxicity* Microbiology 2002 148 , 2215-2222

Wanner BL and Wilmes-Riesenberg MR . *Involvement of phosphotransacetylase, acetate kinase, and acetyl phosphate synthesis in control of the phosphate regulon in Eschenichia coli*. J. Bacteriol. 1992 174:2124-2130.

Wolfe AJ, Chang DE, Walker JD, Seitz-Partridge JE, Vidaurri MD , Lange CF, Prüß BM, Henk MC, Larkin JC and Conway T *Evidence that acetyl phosphate functions as a global signal during biofilm development* Molecular Microbiology 2003 48(4), 977–988

Wolfe AJ *The Acetate Switch* Microbiol Mol Biol Rev. 2005 March; 69(1): 12–50.

CHAPTER 5

VARIATION OF CLOCK DYNAMICS WITH THE REPORTER MODULE AND OTHER CHROMOSOMAL SITES

Introduction

During the development of the Ninfa clock by Ninfa and Atkinson, it was observed that the clock seemed to behave differently when integrated into different *E. coli* host cells. However, those early experiments were conducted manually, since the instrumentation developed in Chapter 2 of this thesis had not yet been developed. Given the difficulty in conducting the experiments manually, and the resulting noisiness of the data, it was not clear whether these early observations were valid or not. Here, with the advantage of the instrumentation developed in Chapter 2 and with well-established experimental conditions developed in Chapter 3, I return to the issue of whether or not the clock functions differently in different host cells, and if so, how this behavior can be explained.

The Ninfa clock is expected to operate with very low numbers of activator and repressor proteins per cell; and in the case of the activator protein it has been shown through immunoblotting experiments that the number of activator molecules per cell is lower than the range in wild-type cells of 20-200 subunits (Atkinson et al, 2003). Thus, one mechanism by which the clock might behave differently in different cell

lines is that either activator or repressor protein, or both, might be affected by sequestration by sites on the chromosome distinct from the clock modules.

Since the reporter used to assess the clock function is a site that binds either activator or repressor protein, we first investigated the variation of the clock dynamics with changes in the reporter module. Our purpose was to evaluate whether the reporter module had an influence on the decay of oscillations. The reporter module in the strain NCPR77 is the wild type *lacZYA* operon, at its normal chromosomal location. This operon has a total of three operator sequences (sequences to which the LacI repressor binds; Jacob and Monod, 1961; Pfahl et al, 1979). Two of these operator sites are in the promoter region, and one operator is within the *lacZ* gene (Oehler et al, 1990). These three operators bind to functional repressor proteins, thus diminishing the amount of free repressor proteins that are available to bind to other sites. In the clock strain, the clock dynamics are understood to arise from the dynamics of repressor and activator proteins binding to specific binding sites in the chromosome. The activator proteins bind to the promoter regions of the clock activator and repressor modules, while the repressor proteins bind to the promoter region of the activator module. By including a *lacZYA* reporter module, the repressor proteins have two potential binding regions in the chromosome (activator module and reporter module), thus the binding sites on the reporter module can be thought as sites that sequester repressor proteins from the synthetic promoter of the activator module and therefore potentially affect the clock dynamics. Since the number of functional repressor and activator proteins is expected to be low, in the order of a few dozens (Reitzer and Magazanik, 1986; Gilbert and Müller-Hill, 1970),

and the fact that the repressor binding sites in the *lacZ* promoter region and in the activator synthetic promoter have similar strength (binding affinity) it would be expected that by sequestering some of the repressor molecules, the presence of the reporter module would result in different dynamics of the clock. That is, a clock design with a reporter module may have different dynamics than a clock design without a reporter module. Of course, this latter case is only useful when considering abstract clock network designs.

This consideration of the sequestration effects of binding sites on the clock dynamics would also be valid for binding sites of the activator proteins (enhancers), and therefore it would be expected that we should be able to alter clock dynamics by manipulating activator binding sites away from the activator and repressor modules. Thus, in one set of experiments, I compare the effects of using *lacZYA* as the clock reporter (as in NCPR77) with using a fusion of the *glnAp2* promoter or the *glnK* promoter to *lacZYA* structural genes as the reporter. I will show that the amplitude and damping of oscillations were different when different reporters were used.

We decided then to also explore the changes in the clock dynamics due to alteration of distal activator binding sites. In a second set of experiments, I examined the effect of deleting enhancers that bind to NRI~P at chromosomal genes far from the clock modules. I will show that in some cases I could discern changes in clock function due to such deletions.

Note on background: When designing the Ninfa clock, the activator module was constructed from the genes in the nitrogen starvation response, because of the lab familiarity with the *glnALG* operon, as well as because the *glnALG* operon was

known to have an auto-activation loop (the *glnG* product, NRI, when phosphorylated, activates the expression of its own promoter). This autoactivation loop is capable of producing rapid changes in the expression level of the *glnA*, *glnG* and *glnL* genes.

In the original clock strains produced by Atkinson and Ninfa, the genetic clock is not well isolated from normal cellular functions. The original clock host strain, designated 3.300LG, lacks endogenous copies of *glnG*, *glnL*, and *lacI*, so the versions of these genes associated with the clock are the only copies in the cell. But, *glnG*-encoded NRI functions as a transcriptional activator, and NRI~P can bind and regulate 17 known operons (as noted in Ecocyc). Thus, analysis of the clock dynamics, for example in a mathematical model or an analysis of dynamics such as this document, should consider the roles of these binding sites.

Methods

To construct clock strains with different reporter modules, I constructed strains with two synthetic reporter modules, in which the *lacZYA* operon was removed from the chromosome and new reporter modules with a promoter activated by the activator protein was fused to the *lacZ* gene. The reporter module with the promoter from the gene *glnK* driving *lacZYA* expression was constructed previously (Pioszak and Ninfa, 2003), and the reporter module with the *glnAp2* promoter driving *lacZYA* was constructed in a similar way as part of this study. The *glnAp2* promoter (this correspond to the *glnAp* promoter, minus the theoretical cAMP binding site; this reduces the basal level of expression, see Farmer and Liao, 2001) was PCR amplified from the *glnAp* promoter region using the primers GlnAp2-US and GlnAp2-DS (see primers table) and cloned into the plasmid pRS551 (Simons et al, 1987) as a EcoRI-

BamHI fragment, to produce the plasmid pGlnAp2LacZ. This plasmid can be used to recombine the module *glnAp2-lacZYA* into the chromosome of the strain TE2680 (Elliot 1992), which is a widely used strain to construct reporter modules with different promoters. The resulting strain was designated NP03. The new reporter modules were transduced into clock strains, after removing the natural *lacZYA* operon, with P1vir grown on strains YMC21kΦ and NP03.

To delete the *lacZYA* region on the chromosome of clock strains, we used the method described in Datsenko and Wanner, 2000. The whole *lac* region, from before the *lacI* promoter to the end of *lacA* was deleted by constructing a PCR product with homologous regions to the regions on the chromosome just before the *lacI* promoter and just after *lacA* flanking a kanamycin resistance cassette. The primers lac-KO-US and lac-KO-DS (see primers table) were used to PCR the ftr-flanked kanamycin cassette in plasmid pKD4 (Datsenko and Wanner, 2000). This PCR product was electroporated into the strain BW25113 pKD46 (Datsenko and Wanner 2000) to produce the strain NP02. P1vir grown on strain NP02 was used to transduce clock strains with the *lac* region deletion. The kanamycin cassette could be removed from a clock strain after such transduction by using the plasmid pCP20 (Datsenko and Wanner 2000); this plasmid causes recombination between the ftr sites flanking the kanamycin resistance gene. In a similar way to the *lac* deletion, a *glnALG* deletion was constructed to remove the gene *glnA* from the clock strains, including its promoter region. The primers glnALG-KO-UP and glnALG-KO-DN were used to PCR the kanamycin cassette in plasmid pKD4. The PCR product was electroporated

Strain	Relevant phenotype	Reference/construction
TE2680	trpDC::putPA1303::[KanS-CamR-lacZ]	Elliot 1992
BW25113	wild type	Dasenko and Wanner 2000
3.300	lacI22	Pardee et al 1959
DE1021	as BW25113, Δeda:kan	BW25113 x Δeda:kan (PCR product)
YMC21kΦ	trpDC::putPA1303::[KanR-glnKp lacZ]	Pioszak and Ninfa 2003
3.30LG	as 3.300, ΔglnLG	Atkinson et al 2003
3.30LGPR	as 3.30LG, resistant to unknown phage	Selected from resistant colonies of 3.30LG
3.30ALGK	as 3.300, ΔglnALG:Kan	3.30LG pKD46 x ΔglnALG:Kan (PCR product)
NP02	as BW25113, Δ(lacI-LacZYA):Kan	BW25113 pKD46 x ΔlacIZYA:Kan (PCR product)
NP03	trpDC::putPA1303::[KanR-glnAp2 lacZ]	TE2680 x pGlnAp2LacZ/ pstI
NP16	as DE1021 ΔKan	DE1021 x pCP20
NP22	as 3.30LGPR, glnAp:tetp:Kan	3.30LGPR pKD46 x glnAp:tetp:Kan (PCR product)
NP23	as NP16, glnHp:tetp:Kan	NP16 pKD46 x glnHp:tetp:Kan (PCR product)
NC77	clock strain	Ninfa lab
NCPR77	as NC77, resistant to unknown phage	constructed on 3.30LGPR
NCPR107	as NCPR77, ΔglnALG:Kan	NCPR77 x 3.30ALGK P1vir
NCPR120	as NCPR77, ΔlacIZYA:Kan	NCPR77 x NP02 P1vir
NCPR121	as NCPR77 ΔlacIZYA	NCPR120 x pCp20
NCPR122	as NCPR121, trpDC:putPA1303:[KanR-glnKp lacZ]	NCPR121 x YMC21kΦ P1vir
NCPR123	as NCPR121, trpDC:putPA1303:[KanR-glnAp2 lacZ]	NCPR121 x NP03 P1vir
NCPR124	as NCPR120, ΔglnALG:Kan	NCPR120 x 3.30ALGK P1vir
NCPR125	as NCPR120, ΔglnALG	NCPR124 x pCP20
NCPR126	as NCPR125, trpDC:putPA1303:[KanR-glnKp lacZ]	NCPR125 x YMC21kΦ P1vir
NCPR127	as NCPR126, trpDC:putPA1303:[KanR-glnAp2 lacZ]	NCPR125 x NP03 P1vir
NCPR140	as NCPR121, glnAp:tetp:Kan	NCPR121 x NP22 P1vir
NCPR141	as NCPR121, glnAp:tetp	NCPR140 x pCP20
NCPR142	as NCPR141, trpDC:putPA1303:[KanR-glnKp lacZ]	NCPR141 x YMC21kΦ P1vir
NCPR143	as NCPR142, trpDC:putPA1303:[KanR-glnAp2 lacZ]	NCPR141 x NP03 P1vir
NCPR144	as NCPR141, glnHp:tetp:Kan	NCPR141 x NP23 P1vir
NCPR145	as NCPR141, glnHp:tetp	NCPR144 x pCP20
NCPR147	as NCPR145, trpDC:putPA1303:[KanR-glnAp2 lacZ]	NCPR145 x NP03 P1vir

Table 5-1 Strains used in this chapter

into the strain 3.30LG pKD46 to produce the strain 3.30ALGK, which then was used as a donor strain to transduce the deletion to the clock strains.

To remove the *glnA* promoter region from clock strains while still maintaining a constitutive level of expression of *glnA*, the promoter region of the *tetA* gene in the plasmid pBR322, known to be constitutive of medium strength (Bolivar et al 1977), was placed in the chromosome as a replacement on the *glnA* promoter region. To

accomplish this, a PCR product was formed with homologous regions to just before the *glnA* promoter region and just after the start of transcription of promoter *glnAp2* flanking a kanamycin cassette fused to the *tetA* promoter. The primers glnALG-KO-US and tetp-glnA-DS were used to PCR the kanamycin cassette of plasmid pKD4. The PCR product was electroporated into the strain 3.300 pKD46 to produce strain NP22, which then was used as a donor strain to transduce clock strains. As above, the kanamycin cassette could be deleted after the transduction by using the plasmid pCP20 to promote recombination between the ftr sites. Similarly to the replacement of the *glnA* promoter region with the *tetA* promoter, the *glnHPQ* promoter region was replaced with the *tetA* promoter region (to express the operon *glnHPQ* constitutively) by using the primers glnH-KO-US and tetp-glnH-DS to PCR the kanamycin cassette in the plasmid pKD4, electroporate the PCR product into strain NP16 pKD46, and produce the strain NP23.

Plasmid	Reference/construction
pRS551	Simons et al 1987
pBR322	Bolivar et al 1977
p3Y15	Atkinson et al 2003
pGlnAp2lacZ	pRS551 x glnAp2 / EcoRI-BamHI
pKD4	Dasenko and Wanner 2000
pKD46	Dasenko and Wanner 2000
pCP20	Dasenko and Wanner 2000

Table 5-2 Plasmids used in chapter 5

Primer	Sequence 5'-3'
glnALG-KO-UP	TACAAAACAGGATCACAAACATCCTCCGCAAACAAGTATTGCAGA GTGTAGGCTGGAGCTGCTTCG
glnALG-KO-DN	ATCATACTGAACTTATCGGAACAGTAAAGCGTAAAATACCAGCAA ATGGGAATTAGCCATGGTCCA
GlnAp2-US	GGGGAATTCTCTAGAAATAACAGGCCTGCTGGTAATCGCAGGCCT TTTTATTTGGCAAAGGTCATTGCACCAACATGG
GlnAp2-DS	GGGGAAGCTTAATTAA CATATG TTAACTCTCCTGGATTGGTCAT GG
lac-KO-US	CAGATCGAAGAAGGGGTTGAATCGCAGGCTATTCTGGTGGCCGG AAGGCGGTGTAGGCTGGAGCTGCTTCG
lac-KO-DS	TCGTTGGGCTGATGATCATAACCCTGCGTTTTGCACCAGTACGTTT TCCGATGGGAATTAGCCATGGTCCA

Table 5-3 Primers used in chapter 5

Results

As described in the methods section, the clock strain NCPR77 was deleted of its *lacI-lacZYA* region to produce the strain NCPR121. The strain NCPR122 was constructed by transducing the reporter module *glnKp-lacZYA* into NCPR121. Strain NCPR123 was constructed by transducing the reporter *glnAp2-lacZYA* into NCPR121. Thus, strains NCPR77, NCPR122 and NCPR123 have the same activator and repressor modules of the clock; they only differ in the reporter module used. These strains showed marked differences in clock dynamics (Fig 5-1). While strain NCPR122 showed low level of expression with respect to the steady state expression level of strain NCPR77, and a barely measurable oscillations (see inset in Fig 5-1), strain NCPR123 showed a high level of expression with respect to the strain NCPR77, with a steady state expression level of ~1000 miller units, compared to ~650 miller units of steady state expression. Besides these differences in steady-state expression, one marked difference between strains NCPR123 and NCPR77, visible in Fig 5-1, was the decay rate of the peaks amplitude. The decay rate of the strain NCPR123 is noticeable higher than the decay of strain NCPR77. Measured as the rate of exponential decay of the peaks amplitude towards the steady state, the decay rate of strain NCPRR123 was 0.055 h^{-1}, while the decay rate of strain NCPR77 was 0.037 h^{-1}. The period of oscillation in all three strains was the same (10-11 hours), which is consistent with the fact that all three strains showed about the same growth rate (doubling time 68-72 minutes).

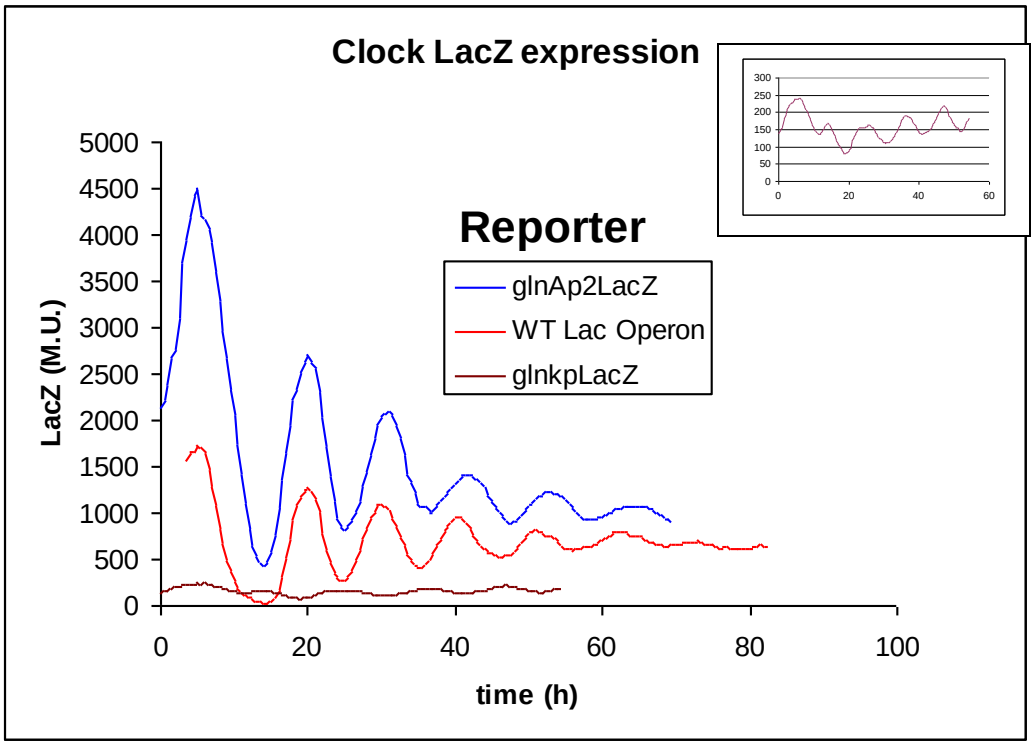


Figure 5-1 Experiment results for three clock strains with different reporter modules, but otherwise with the same genetic background and under the same environmental conditions**.** Strain NCPR77 was the "regular" clock strain with the wild type *lac* operon as reporter, while the strain NCPR122 has the reporter *glnKp-lacZYA* and strain NCPR123 has the reporter *glnAp2-lacZYA*. Inset: blow-up of the clock strain NCPR122.

Since we observed that modifying the reporter module could have a big impact in the observed decay rate, and taking into account that modifying the reporter is equivalent to manipulating other binding sites, then we decided to remove the gene *glnA* (including its promoter region) from the chromosome of strains NCPR77, NCPR122 and NCPR123. Removing the promoter *glnAp* from the chromosome of clock strains would remove a strong binding site of the activator protein. The results of removing the gene *glnA* and its promoter region are shown in Fig 5-2, 5-3, and 5-5.

The gene *glnA* and its promoter region were deleted from strain NCPR77 to produce strain NCPR107. When run under identical conditions, strains NCPR77 and NCPR107 did not show significant differences in growth rate, period of oscillation or peaks amplitude (Fig 5-2). The only observable difference in the clock dynamics was that the lowest levels of expression in the oscillations were about 200 miller units higher than in NCPR77.

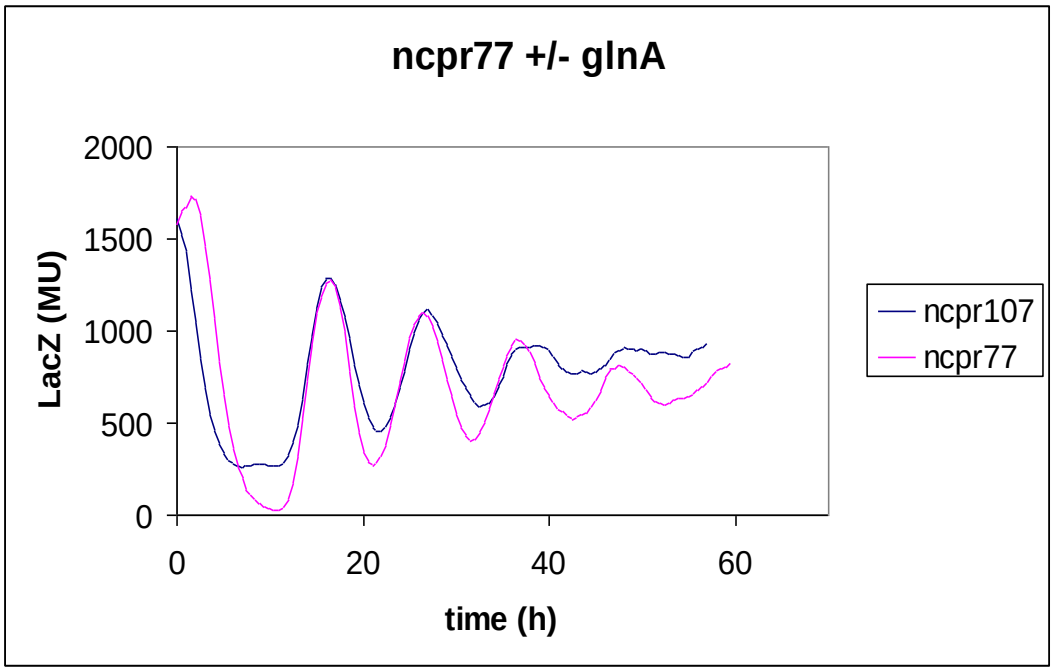
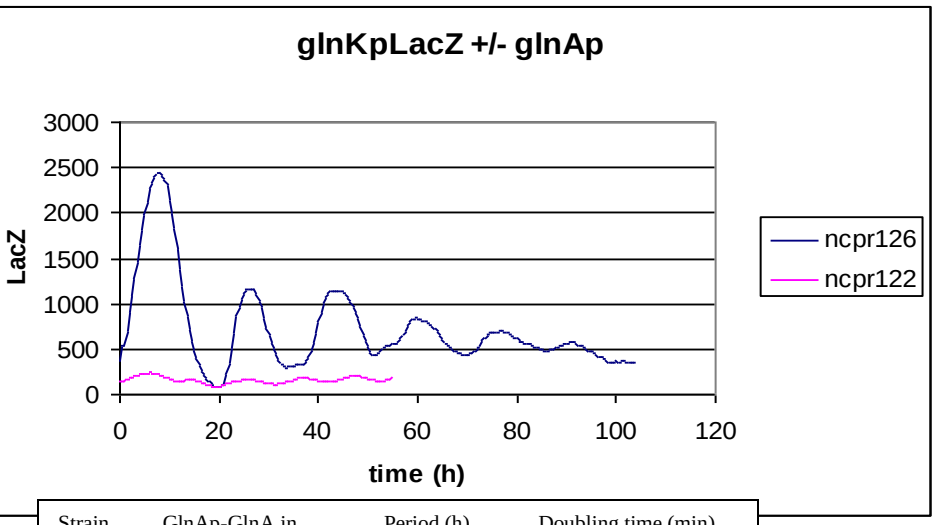


Figure 5-2. Comparison of clock dynamics with and without the gene *glnA* and its promoter region. The deletion of the gene *glnA* from strain NCPR77 did not produce a change in the period of the clock or in the peaks amplitude. The only difference in dynamics seems to be that the lower points in the clock oscillations in NCPR107 are somewhat higher than the original strain NCPR77.

When removing *glnA* and its promoter from the strains NCPR122 and NCPR123, the results were markedly different than what happened when removing *glnA* and its promoter from NCPR77. In the case of NCPR122, the deletion of *glnA*



Strain	GlnAp-GlnA in chromosome	Period (h)	Doubling time (min)
ncpr126	-	16-17	88-90
ncpr122	+	11-11.5	70-72

Figure 5-3 Comparison of clock dynamics with and without the gene glnA and its promoter region in strain NCPR122. Very small oscillations in strain NCPR122 are drastically changed when the *glnA* gene is removed. The new oscillations in strain NCPR126 are closer to oscillations observed in the original clock strain NCPR77

and its promoter (resulting in strain NCPR126) resulted in clock activity very different from strain NCPR122; NCPR122 clock activity was almost flat at low level, whereas NCPR126 produced noticeable oscillations, similar to those in other clock strains, such as NCPR77. The strain NCPR126 also grew significantly slower than the strain NCPR122, and this was reflected in a larger period of oscillation (16 hours, compared to 11 hours in NCPR122). The big difference in dynamics between strains NCPR122 and NCPR126 made us doubt about the repeatability of the experiment with NCPR122. We found that the low level flat dynamics exhibited by NCPR122 were not exactly repeatable. Several runs with the same strain gave the same low level of steady state repression, however small amplitude oscillations were not always discernible, and the amplitude of the first peak (where cells presumably have IPTG in

the cytoplasm) could go very high before settling to the low steady state. The promoter of the *lacZ* gene in NCPR122 was sequenced, and no mutations were found in it, indicating that the difference in amplitude of the first peak was not likely due to genetic differences in the reporter module. When strain NCPR122 was reconstructed by transducing the *glnKp-lacZYA* module to strain NCPR121, using fresh P1vir with such module (producing strain NCPR122_2), the results were different again: this time the strain NCPR122 showed two oscillations before abruptly settling to a low steady state, similar to the steady states observed before. In summary, strain NCPR122 showed different transient dynamics in different runs, but they always settled to a low steady-state level of expression. Different dynamics of different runs of strain NCPR122 are shown in Fig 5-4.

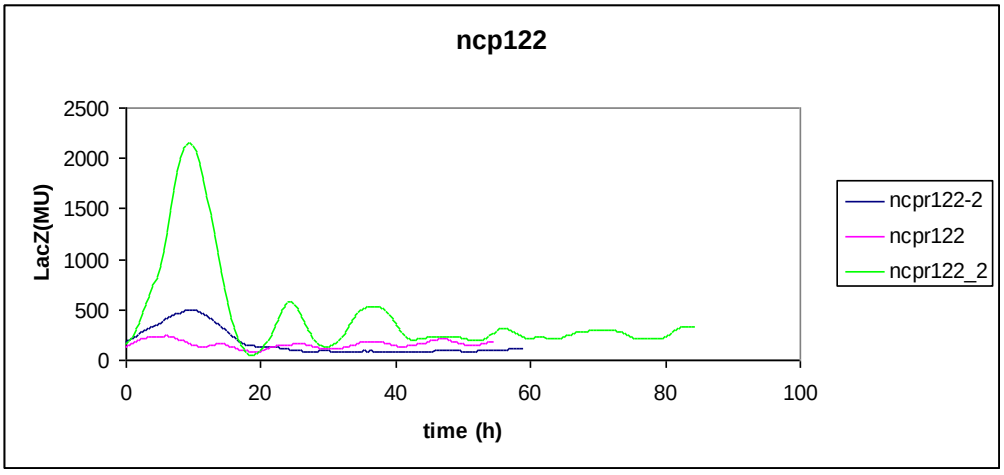
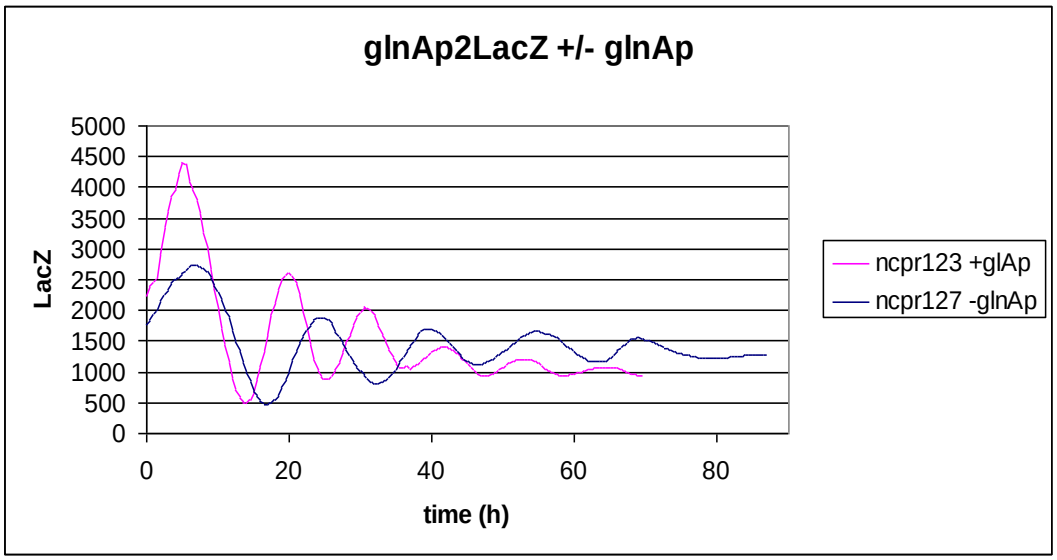


Figure 5-4 Comparison of different runs of strain ncp122. The runs show different transient dynamics but similar steady-state values. Different transient dynamics did not appear to be due to genetic differences

When deleting the gene *glnA* and its promoter from strain NCPR123, producing NCPR127, the results were also very different from the initial strain NCPR123. Strain NCPR127 showed slower growth rate and longer period , similar to NCPR126, but strain NCPR 127 also "improved" the clock dynamics, in the sense that the decay rate of the peaks amplitude became 0.018 h^{-1}, about one third the decay rate for NCPR 123, and less than the one for NCPR 77.

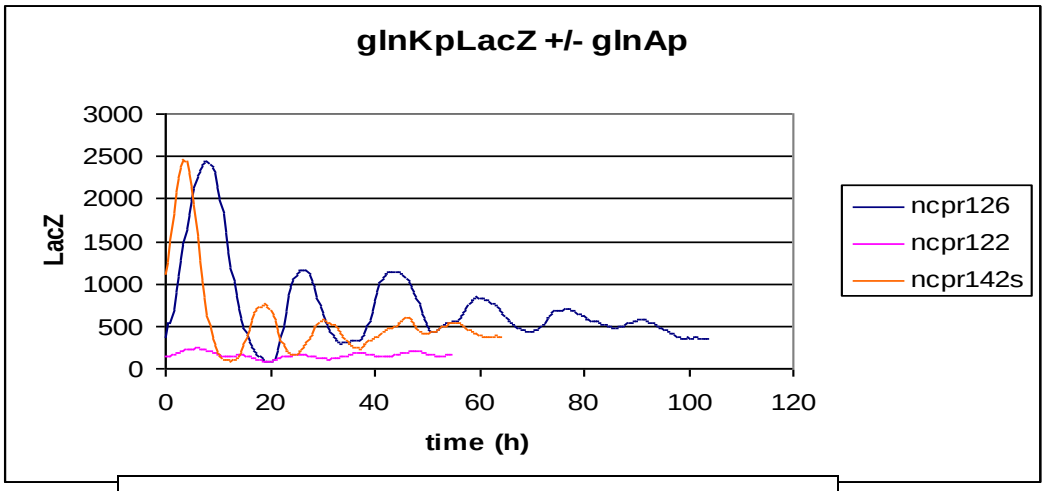


Strain	GlnAp2 in chromosome	Period (h)	Doubling time (min)
ncpr127	-	15-16	85-87
ncpr123	+	11-11.5	70-72

Figure 5-5 Comparison of clock dynamics with and without the gene glnA and its promoter region in strain NCPR123. In contrast to strain NCPR107, the *glnA* deletion causes slower growth rate, reflected on a larger period in the same environmental conditions. The deletion of *glnA* also reduces the damping rate by a factor of 2.

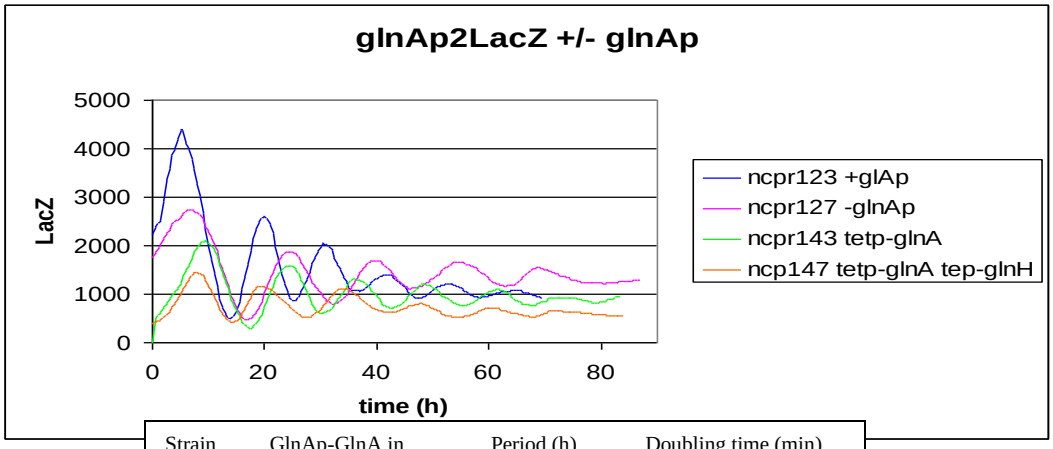
Since both strains NCPR126 and NCPR127 showed slower decay rates compared to NCPR122 and NCPR123 respectively, we investigated the effect of

removing the *glnAp* promoter from the chromosome, while still maintaining expression of the gene *glnA*. This was accomplished by replacing the *glnAp* promoter in the chromosome with a low-strength constitutive promoter, without modifying the *glnA* gene, so *glnA* would be expressed at low level, similar to when cells grow in nitrogen-rich conditions. We chose the promoter of the *tetA* gene in the plasmid pBR322 (Bolivar et al 1977), and we replaced the *glnAp* promoter region in strains NCPR122 and NCPR123 with the *tetAp* promoter to produce strains NCPR142 and NCPR143, respectively. In both cases the strains with the *tetAp* promoter in lieu of the *glnAp* promoter (NCPR142 and NCPR143) showed faster growth rate (and shorter period) than the strains with the *glnAp-glnA* region deleted (NCPR126 and NCPR127), but this grow rate was still slower than that of the strains with intact *glnAp-glnA* region. (See figures 5-6 and 5-7). In the case of strain NCPR143, although the growth rate was "improved" with respect to NCPR127, the decay rate of the peaks amplitude was increased with respect to NCPR127 by a factor of two (Fig 5-7). An additional strain was constructed, in which the *glnHp* promoter (another NRI binding site) was replaced with the *tetAp* promoter, in a similar fashion as the *glnAp* promoter was replaced with *tetAp* in NCPR143. When the *tetAp* promoter was used to replace the *glnHp* promoter in strain NCPR143, the result was named NCPR147. The strain NCPR147 was constructed with the purpose of "freeing up" more activator protein from binding sites not related (in principle) to the clock design. The result of NCPR147, when compared to strain NCPR143 was that the clock dynamics were not "better": The growth rate became slower (larger period) and the decay rate was slightly increased (Fig 5-7).



Strain	GlnAp-GlnA in chromosome	Period (h)	Doubling time (min)
ncpr126	-	16-17	88-90
ncpr122	+	11-11.5	70-72
ncpr142	tetp-glnA	12-13	78-80

Figure 5-6 Comparison of clock dynamics of strains with the glnKpLacZ reporter module. When the *glnA* promoter region is replaced with the constitutive promoter *tetp*, oscillations become more noticeable, and there is only a small reduction in growth rate, compared to a larger loss in growth rate when *glnA* and its promoter region are deleted. This suggests a role of the *glnA* protein in clock dynamics.



Strain	GlnAp-GlnA in chromosome	Period (h)	Doubling time (min)
ncpr127	-	16-17	88-90
ncpr123	+	11-11.5	70-72
ncpr143	tetp-glnA	12-13	78-80
ncpr147	tetp-glnA tetp-glnH	14-15	85

Figure 5-7 Comparison of clock dynamics of strains with the glnAp2LacZ reporter module. When the *glnA* promoter region is replaced with the constitutive promoter *tetp*, oscillations become less damped. When the *glnA* and its promoter region are deleted the oscillations become even less damped, but the growth rate is significantly decreased. When additionally the *glnH* promoter region is replaced with *tetp*, the damping rate does increase with respect to the strain NCPR127 (*tetpglnA* only), indicating that removal of NRI binding sites does not necessarily "improve" the clock dynamics.

Discussion

The experiments presented in this chapter show that manipulation of activator and repressor protein binding sites in the chromosome, that were not intended to be part of the clock design, can actually alter the clock dynamics significantly. When the *lacZYA* operon in strain NCPR77 was removed and replaced with another reporter module, the dynamics changed considerably, as seen in Fig 5-1. This shows that the addition and/or eliminating activator and repressor binding sites could have effects in changing the decay rate of the peaks of expression, as well as the level of the steady-state expression. When the reporter module was switched from lacZYA to *glnKp-lacZYA*, the steady state was reduced by a factor of 3, approximately, and when the reporter module was switched from *lacZYA* to *glnAp2-lacZYA* the steady state was increased by a factor of 1.6. However, the period of oscillation was not changed when the reporter was switched. This is consistent with the hypothesis that the period is proportional to doubling time (Chapter 2), since the measured growth rate was the same in the three strains with different reporter modules. The changes in steady-state expression level show that the changes in binding sites can cause big changes in the expression levels of activator and repressor proteins. This is presumably due to changes in the level of activation and repression in the clock modules, in turn due to the changes of available free activator and repressor proteins as a consequence of changes in the binding sites. Since the changes in expression dynamics observed, due to changes in binding sites, did not change the period of oscillation, could this mean that the period of oscillation does not change with changes in the expression level of activator and repressor proteins? In Chapter 2 and with the changes in the *glnAp*

promoter we observed changes in the period of oscillation, but these changes in period were always accompanied with changes in doubling times. Are the changes in period the consequence of changes in the growth rate only? All the experiments presented in this chapter are consistent with such a conclusion, since every time the period changed, this could be attributed to changes in growth rate. Rejection of this hypothesis would require two experiments with the same growth rate, but with changes in period, due to changes in the binding sites of activator and/or repressor proteins. No such pair of experiments was observed.

When removing the whole *glnAp-glnA* region, it was observed that the period of oscillation (and the doubling time) was increased in strains NCPR126 and NCPR127 (which have synthetic promoters), but it was not changed in strain NCPR107 (with the wild type *lac* operon as reporter). This result is puzzling: what makes the case of NCPR107 different to NCPR126/127? In principle, the deletion of the gene *glnA* should have no effect on growth rate if there is glutamine supplied in the growth medium; however, in strains NCPR122 and NCPR123 the deletion of *glnA* slowed the growth significantly and even the replacement of the *glnA* promoter with a constitutive promoter slowed the growth, indicating that the expression and regulation of *glnA* in the strains NCPR122/123 is a factor in the growth rate, with the given medium, something that is not the case for strain NCPR77. Given that the change in reporter alone cannot be the cause of the slower growth, since strains NCPR77, NCPR122 and NCPR123 all showed similar growth rates, and the deletion of *glnA* alone cannot be the cause of slower growth since strains NCPR77 and NCPR107 showed similar growth rates, the reason for the slower growth of strains

NCPR126/127 must lie in some cooperativity between the deletion of *glnA* and the change in reporter. One difference between the reporter modules in NCPR77 and NCPR122/123, besides the difference in promoters, is that the synthetic promoters in NCPR122/123 were constructed in such a way that they recombine in the chromosome in the *trp* operon (*trpABCDE*), so the *trp* enzymes are not expressed. Since the growth media is always supplied with tryptophan, the expression of these enzymes should not make a difference in growth. One possible construction to test the cooperativity would be to construct the synthetic reporters in the same chromosomal locus as *lacZYA*, so no other parts of the chromosome are affected when switching promoters. This could be accomplished by replacement of the promoter region of the *lac* operon in strain NCPPR77 with a synthetic promoter. The way to do such a replacement would be to construct a piece of DNA in which regions of homology to the region previous to *lacI* and to the start of *lacZ* flank the sequence of an antibiotic marker followed by the synthetic promoter. This would be analogous to the way in which the *glnA* promoter region was replaced with the *tetAp* promoter in strains NCPR126/127.

The effect of changing the *glnAp-glnA* region of strains NCPR122/123 in the clock damping factor was not as clear as we have hoped for. In strains NCPR122, NCPR126 and NCPR142 , strains that have the *glnKp-lacZYA* promoter, the effects of manipulation of *glnAp-glnA* region could not be studied, since this strains showed some variability in behavior (Fig 5-4), and the oscillations did not show exponential decay, or some decay close to it. In NCPR122, for example, one experiment showed very small oscillations that did not seem to decay in amplitude, while other

experiment showed a couple of similarly sized oscillations before the expression becomes abruptly flat. NCPR126 showed two similarly sized oscillations before decaying to smaller oscillations. We therefore concluded that the *glnKp-lacZYA* reporter module may not be appropriate to show regular oscillations, even if damped, as we would like a genetic clock to show. However, in this set of strains we can observe how the deletion (or manipulation) of a single binding site in the chromosome, far away from the clock modules, can drastically alter the clock expression dynamics.

In strains NCPR123/127/143/147, strains that have the *glnAp2-lacZYA* reporter module, we can see regular oscillations and therefore see more clearly the effects of manipulating the *glnAp-glnA* region. When deleting the whole region from strain NCPR123 to produce NCPR127, the decay rate went from 0.055 h^{-1} to 0.018 h^{-1}, a drop by a factor of 3. However, since strain NCPR127 grows slower than NCPR123, it would be expected that by the sole fact that it grows slower the decay rate would be less, since decay rate has units of inverse time. When comparing the decay rate times the doubling time, the results are doubling-time-corrected decay rate of 0.066 for NCPR123 and 0.027 for NCPR127, showing that the deletion of the *glnAp-glnA* region showed improvement in the decay rate, regardless of the slower doubling time. When the *glnAp* promoter was replaced with the *tetAp* promoter in strain NCPR123 to produce strain NCPR143, the decay rate was .037 h^{-1}, and the doubling-time-corrected decay rate was 0.05, which indicates that the replacement of the *glnAp* promoter (with its binding sites) did indeed improved the clock dynamics by reducing the decay rate, but the decay rate was reduced much more by deleting the

whole region, suggesting that the gene *glnA* and its regulation have an impact in clock dynamics, beyond slowing the growth rate, and apart from the removal of binding sites. The deletion of another NRI binding (*glnHp*) in strain NCPR143 to produce NCPR147, showed a decay rate of 0.045 h^{-1} and a doubling-time-corrected decay rate of 0.063, roughly the same as NCPRR123, indicating that the deletion of a second NRI binding site did not improve the clock dynamics as it did the deletion of the stronger binding site (in *glnAp2*). This shows that the removal of binding sites, leading to the increase of activator protein available for the clock, is not necessarily good for the clock performance, as we had hoped for, and the clock performance would not necessarily benefit from further removal of NRI binding sites. Clock performance can be influenced also by other gene products which are not known to be directly related to clock dynamics, such as *glnA*.

It should be noted that the variation in clock dynamics due to deletion of binding sites could be considered as an example of the recently introduced concept of retroactivity (Del Vecchio et al, 2008; Saez-Rodriguez et al, 2008) in which a downstream target of a network (a binding site in this case) can affect the dynamics of an upstream module. Examples of this retroactivity include the regulation of MAPK phosphorylation by some of its substrates (Kim et al 2011) and changes in sensitivity to stimulus in signal transduction by a substrate (downstream target), in a synthetic system of modified *E.coli* proteins (Ventura et al 2010). However, due to the nature of the interaction of the clock proteins and distal chromosomal binding sites, it is difficult to classify the binding sites as "downstream targets", since the binding sites in the promoters of clock modules have the same physiological function

as those binding sites called “downstream targets”. There is no functional difference between one group of binding sites and another. In this sense, is not possible to separate the clock modules and the rest of the chromosome as different “network modules”. Future research in synthetic clocks and other synthetic constructions must account for this lack of modularity.

LIST OF REFERENCES CITED IN CHAPTER 5

Atkinson M R, Savageau MA, Meyers J and Ninfa AJ. *Development of a genetic circuitry exhibiting toggle switch or oscillatory behavior in Escherichia coli*. Cell 2003 113:597-607.

Bolivar F, Rodriguez RL, Greene PJ, Betlach MC, Heynecker HL and Boyer HW *Construction and characterization of new cloning vehicles. II. A multipurpose cloning system*. 1977 Gene, 2, 95-113.

Datsenko A and Wanner BL *One-step inactivation of chromosomal genes in Escherichia coli K-12 using PCR products* PNAS 2000 vol. 97 no. 12 6640-6645

Del Vecchio D, Ninfa AJ and Sontag ED *Modular cell biology: retroactivity and insulation* Molecular Systems Biology 2008 4:161

Ecocyc, E.coli bioinformatics database. Nucleic Acids Research 37:D464-70 2009 http://ecocyc.org

Elliot T. *A method for constructing single-copy lac fusions in Salmonella typhimurium and its application to the hemA-prfA operon*. J Bacteriol. 1992;174:245–253.

Farmer WR and Liao JC *Acetate-inducible protein overexpression from the glnAp2 promoter of Escherichia coli*. Biotechnology and Bioengineering 2001 75:504-9

Gilbert W and Müller-Hill B. *The lactose repressor*. In: Beckwith JR, Zipser D, editors. The lactose operon. NY: Cold Spring Harbor Press; 1970. p. 93-109

Jacob F and Monod J *Genetic regulatory mechanisms in the synthesis of proteins*. Journal of molecular biology 1961 Volume: 3 Pages: 318-

Kim Y, Paroush Z, Nairz K, Hafen E, Jimenez G and Shvartsman SY *Substrate-dependent control of MAPK phosphorylation in vivo* Molecular Systems Biology 2011 7:467

Oehler S, E R Eismann ER, H Krämer H and Müller-Hill B *The three operators of the lac operon cooperate in repression*. EMBO J. 1990 April; 9(4): 973–979

Pardee, AB, F. Jacob F and Monod J. *The genetic control and cytoplasmic expression of "inducibility" in the synthesis of B-galactosidase by E. coli*. J.Mol.Biol. 1959 1:165-178

Pfahl M, Gulde V and Bourgeois S. *"Second" and "third operator" of the lac operon: An investigation of their role in the regulatory mechanism* Journal of Molecular Biology 1979 Volume 127, Pages 339-344

Pioszak AA and Ninfa AJ *Genetic and Biochemical Analysis of Phosphatase Activity of Escherichia coli NRII (NtrB) and Its Regulation by the PII Signal Transduction Protein* Journal of bacteriology 2003, Vol. 185, No. 4 p. 1299–1315

Reitzer LJ and Magazanik B *Expression of glnA in. Escherichia coli is regulated at tandem promoters.* Proc. Natl. Acad. Sci. USA 1985 82:1979-1983.

Saez-Rodriguez J, Gayer S, Ginkel M and Gilles ED *Automatic decomposition of kinetic models of signaling networks minimizing the retroactivity among modules*. Bioinformatics 2008 24: i213–i219

Simons RW, Houman F and Kleckner N. *Improved single and multicopy lac-based cloning vectors for protein and operon fusions*. Gene.1987;53(1):85-96.

Ventura AC, Jiang P, Van Wassenhove L, Del Vecchio D, Merajver SD and Ninfa AJ *Signaling properties of a covalent modification cycle are altered by a downstream target* PNAS June 1, 2010 vol. 107 no. 22 10032-10037

CHAPTER 6

PREDICTING WHICH NETWORK DESIGNS ARE MOST LIKELY TO SHOW SUSTAINED OSCILLATIONS

Introduction

When building synthetic genetic oscillators, one would like to maximize the likelihood that an artificial genetic network would exhibit sustained oscillations. It is reasonable to assume that the network structure (here referred to as topology) will be a major factor in determining whether oscillations will appear. Thus, the builder of a genetic clock network wants to know which topology gives a better chance of oscillations over another topology. A general answer for this question is not easy to find, in particular for designs with different number of variables and parameters. Variables and parameters may, in fact, play a key role in determining the behavior of oscillators (Conrad et al 2008). In this chapter I present a case study that represents a simplified version of this problem, by comparing two simple three-gene clock designs: The repressilator, which was the pioneer of genetic clock designs introduced by Ellowitz and Leibler (2000) , and another design what we term the activilator, which is a network with three genes in which the first gene activates the expression of the second gene, the second gene activates the third gene, and the third gene represses the expression of the first gene, to close a feedback loop. The activilator is the same

as the repressilator with the repression of the first gene on the second and the repression of the second gene on the third replaced with activation. Both are overall negative feedback loops, which are designs where oscillations can be achieved. Schematics of the two genetic network designs are shown in figure 6-1

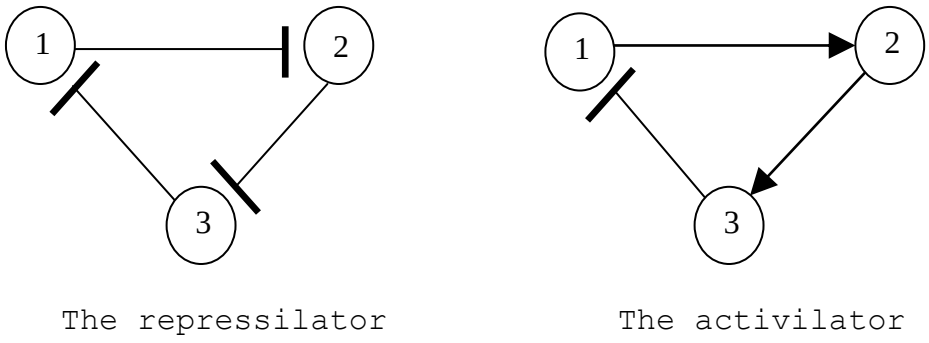


Figure 6-1 . Schematics of two three-module synthetic clock designs. The two networks represent an overall negative feedback loop. The difference between the two networks is the way in which gene 1 activates its own repression. In the repressilator, gene 3 is activated by two serial repressions, whereas in the activilator gene 3 is activated by two serial activations

Since these two designs are quite similar, a direct comparison of their likelihood of oscillation given a set of parameters is possible. In fact, by considering the biochemical reactions that give rise to activation and repression of gene expression, we show a direct correspondence between the parameters of the two systems (see appendix). By this, we mean they are composed of the same basic biochemical reactions, with the possible exception of switching activation for repression.

These two designs may seem may equivalent at first, since two activations seem to have the same overall effect as two repressions, which is activation of the last gene. However, transcription regulation saturates, and this saturation is different for

activation and repression. The level of saturation is a key factor in whether oscillations can be seen since it defines the sensitivity of transcription regulation which has previously been shown to be a key factor in whether oscillation are seen. Although we find both designs are capable of oscillations, the set of parameters for which each design oscillates is different. The repressilator is most likely to oscillate when the genes are under moderate expression, whereas the activilator is most likely to oscillate when the genes are close to fully activated. This later predictions is quite useful in that, if built, the activilator can be "guided" to an oscillating state by simply increasing gene expression.

Before introducing the mathematical formalism, I review the concept of sensitivity of a given function, which is the main concept used to evaluate the likelihood of oscillations in the two synthetic clock designs considered.

Sensitivity of activation and repression

Two forms of activation of a gene in a biochemical/genetic network are commonly found. Direct activation, as when a modified RNA polymerase can bind to a promoter that was not previously active, or derepression, when an enzyme or an external factor represses a constitutive repressor of a gene. The classical example of direct gene activation is the activation of the AraBAD genes in E.coli (whose product is necessary to metabolize arabinose) by the AraC protein in the presence of arabinose. This model of gene activation was proposed by Englesberg et al (1965). For a recent review see for example Schleif (2002). The classical example of derepression is the activation of the LacZYA genes in E.coli (whose product is necessary to metabolize

Lactose) by the gene LacI in the presence of Lactose. This model of gene activation was the pioneer model in gene activation as was proposed by Jacob and Monod (1961). For a more modern review see for example Muller-Hill (1996). When considering these forms of activation at the network level, they are often considered as equivalent, because their overall effect is gene activation. Here we point to a very significant difference between these two forms of activation, at the network level: The sensitivity of the two motifs to input signals are quite different.

The sensitivity of a continuous function $f(x)$ is defined as $\frac{dLog(f(x))}{dLog(x)} = \frac{df(x)}{dx}\frac{x}{f}$ and it measures the "exponent" of the power law that approximates the function $f(x)$ locally. If *f(x)* is the rate of a chemical reaction, then the sensitivity of $f(x)$ measures the order of such a reaction. Let us consider the widely used functions for the biochemical reactions of genetic activation and genetic repression, namely, the hill functions. The rate of transcription of a gene that can be repressed, given the concentration of repressor protein *p*, is

$$\frac{dm}{dt} = \alpha \frac{1}{1+(p/K)^n}$$

Where α and K are constants, and n is an integer exponent. Similarly, the rate of transcription of a gene that can be activated, given the concentration of activator protein *p*, is

$$\frac{dm}{dt} = \alpha \frac{(p/K)^n}{1+(p/K)^n}$$

These functions are normally called Hill functions, since they are the same as the famous Hill equation, used in biochemistry to quantify the rate of ligand binding to a protein. The Hill functions can be derived from basic mass-action kinetics, by considering the reactions of repressor/activator binding to a promoter (see appendix). To see the difference in sensitivity between activation and derepression, let us consider the difference in sensitivity between two activations in series and two repressions in series. That is, consider the case of three genes, A,B and C, in which A activates C either by two step direct activation or by two repressions in series, either case causing activation of C by A (figure 6-2)

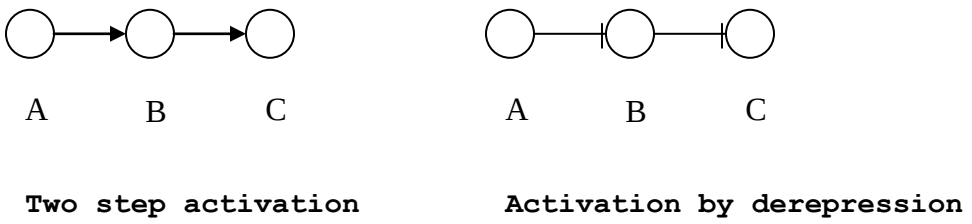


Figure 6-2. Direct activation vs. activation by derepression. A significant difference between these two forms of activation can be seen in the sensitivity of the activation and repression functions.

Denoting the function for repressible transcription rate as (setting α and *K* equal to 1, and n=3, as an example)

$$R(p) \equiv \frac{1}{1+p^3}$$

And the function for transcription by activation as

$$A(p) = \frac{p^3}{1+p^3}$$

we have that the overall action of gene A on gene C is either $A(A(p))$ or $R(R(p))$. Considering the response of gene C to a periodic signal in gene A, say *Sin(x)+1* (figure 6-2), we can see how different the responses will be. In the case of two activations in series, the response is very low for low input signal, and for intermediate to high input signal the response varies smoothly. In the case of derepression, the response is also very low for low input, but the response varies more abruptly between low and high plateaus for low and high input.

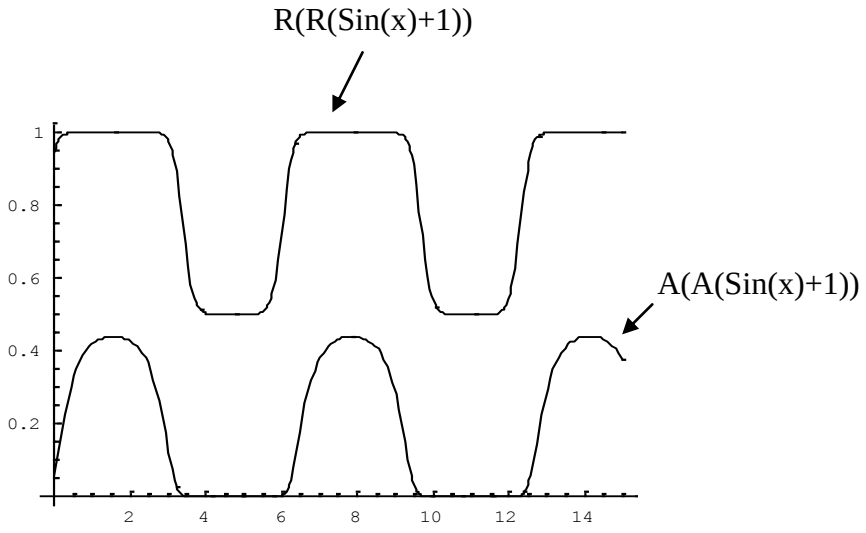


Figure 6-3 Response of double activation vs. double repression to an input signal. In the case of double activation, the response to the input signal *Sin(x)+1* varies smoothly for intermediate-to-high input signal, whereas for double repression the response varies more abruptly between low and high response plateaus for low and high input signal.

With the use of the definition of sensitivity of activation and repression functions, we compare the likelihood of oscillations between the repressilator and the activilator, and show how it may be easier to achieve oscillations with the latter.

Mathematical Models

Defining the functions for repression and activation as previously found,

$$R[p] \equiv \frac{1}{1+p^n}$$ Repression function

$$A[p] \equiv \frac{p^n}{1+p^n}$$ Activation function

We can write the mathematical models for the repressilator and activilator as (assuming linear protein and mRNA degradation and linear translation) with the same set of basic reactions

Repressilator		Activilator	
$\frac{dm_1}{dt} = t_1 R[p_3] - d_1 m_1$	$\frac{dp_1}{dt} = l_1 m_1 - g_1 p_1$	$\frac{dm_1}{dt} = t_1 R[p_3] - d_1 m_1$	$\frac{dp_1}{dt} = l_1 m_1 - g_1 p_1$
$\frac{dm_2}{dt} = t_2 R[p_1] - d_2 m_2$	$\frac{dp_2}{dt} = l_2 m_2 - g_2 p_2$	$\frac{dm_2}{dt} = t_2 A[p_1] - d_2 m_2$	$\frac{dp_2}{dt} = l_2 m_2 - g_2 p_2$
$\frac{dm_3}{dt} = t_3 R[p_2] - d_3 m_3$	$\frac{dp_3}{dt} = l_3 m_3 - g_3 p_3$	$\frac{dm_3}{dt} = t_3 A[p_2] - d_3 m_3$	$\frac{dp_3}{dt} = l_3 m_3 - g_3 p_3$

Where the variables of the system are $\{m_i\}$, mRNA levels for each gene, and $\{p_i\}$ protein levels. The parameters are $\{t_i\}$ transcription rates, $\{l_i\}$ translation rates, $\{d_i\}$ mRNA degradation rates and $\{g_i\}$ protein degradation rates.

For these models, we can easily calculate equations for the fixed point for each system

For the repressilator

$$\varphi_1 R(\varphi_3 R(\varphi_2 R(p_1^*))) = p_1^*$$
$$\varphi_2 R(\varphi_1 R(\varphi_3 R(p_2^*))) = p_2^*$$
$$\varphi_3 R(\varphi_2 R(\varphi_1 R(p_3^*))) = p_3^*$$

For the activilator

$$\varphi_1 R(\varphi_3 A(\varphi_2 A(p_1^*))) = p_1^*$$
$$\varphi_2 A(\varphi_1 R(\varphi_3 A(p_2^*))) = p_2^*$$
$$\varphi_3 A(\varphi_2 A(\varphi_1 R(p_3^*))) = p_3^*$$

Where $\varphi_i \equiv \frac{t_i l_i}{d_i g_i}$

Likelihood of Oscillations and the Secant condition

The local condition for oscillations in a model given a parameter set can be calculated by evaluating the eigenvalues of the Jacobian Matrix for the system, assuming that the instability of the fixed point leads to oscillations. For simple systems like the repressilator and the activilator chaotic behavior can be ruled out (Mallet-Paret and Smith 1990; El Samad et al 2005), so instability of the fixed point always leads to oscillations. For these systems, calculating the eigenvalues requires solving a sixth degree polynomial, something that cannot be done analytically (although easily done numerically). However, a necessary (not sufficient) condition for oscillation can be obtained. This condition is called the Secant condition (Othmer 1976; Tyson and Othmer 1978, Thron 1991). It applies to any biochemical network that consists of an overall negative feedback loop. As it applies to the repressilator and the activilator it can be written as

A necessary condition for oscillations of the repressilator equations is

$$\left|S_R(p_3^*)\right| S_R(p_1^*) S_R(p_2^*) \geq Sec(\frac{\pi}{6})^6 \cong 2.37$$

A necessary condition for oscillations of the activilator equations is

$$\left|S_R(p_3^*)\right|S_A(p_1^*)S_A(p_2^*) \geq Sec\left(\frac{\pi}{6}\right)^6 \cong 2.37$$

Where p_i^* are the fixed points for each protein and S_R , S_A are the *sensitivities* of the repressor and activation functions, respectively

$$S_R(p) \equiv \frac{d\,Log(R(p))}{d\,Log(p)} = \frac{dR(p)}{d\,p}\frac{p}{R(p)} = \frac{-np}{1+p^n}$$

$$S_A(p) \equiv \frac{d\,Log(A(p))}{d\,Log(p)} = \frac{dA(p)}{d\,p}\frac{p}{A(p)} = \frac{1}{1+p^n}$$

Note that $\left|S_R(p)\right|$ is an increasing function and $S_A(p)$ is a decreasing function.

Our question is whether the left hand side (lhs) of this secant condition can be increased by increasing the fixed point p_3^*, which is equivalent to increasing the value of $\varphi_3 = \frac{t_3 l_3}{d_3 g_3}$.

Consider the functions

$$RR(p) \equiv \varphi_3 R(\varphi_2 R(p))$$

$$AA(p) \equiv \varphi_3 A(\varphi_2 A(p))$$

The sensitivity of these functions is

$$S_{RR} = \frac{d\,RR(p)}{dp}\frac{p}{RR(p)} = \frac{dR(\varphi_2 R(p))}{d(\varphi_2 R(p))}\frac{\varphi_2\,dR(p)}{dp}\frac{p}{R(\varphi_2 R(p))} = \frac{dR(\varphi_2 R(p))}{d(\varphi_2 R(p))}\frac{\varphi_2 R(p)}{R(\varphi_2 R(p))}\frac{dR(p)}{dp}\frac{p}{R(p)}$$

$$= S_R(\varphi_2 R(p))S_R(p)$$

$$S_{AA} = \frac{dAA(p)}{dp}\frac{p}{AA(p)} = \frac{dA(\varphi_2 R(p))}{d(\varphi_2 R(p))}\frac{\varphi_2 dA(p)}{dp}\frac{p}{A(\varphi_2 R(p))} = \frac{dA(\varphi_2 R(p))}{d(\varphi_2 R(p))}\frac{\varphi_2 A(p)}{A(\varphi_2 R(p))}\frac{dA(p)}{dp}\frac{p}{A(p)}$$

$$= S_A(\varphi_2 A(p))S_A(p)$$

Since $p_1^* = \varphi_1 R\left(p_3^*\right)$ and $\begin{cases} p_2^* = \varphi_2 R(p_1^*) & \text{For the repressilator} \\ p_2^* = \varphi_2 A(p_1^*) & \text{For the activilator} \end{cases}$

We have

$$S_{RR}(\varphi_1 R(p_3^*)) = S_R(p_2^*)S_R(p_1^*)$$

$$S_{AA}(\varphi_1 R(p_3^*)) = S_A(p_2^*)S_A(p_1^*)$$

The secant condition can then be written as,

$$\left|S_R(p_3^*)\right|S_{RR}(\varphi_1 R(p_3^*)) \geq Sec(\frac{\pi}{6})^6 \cong 2.37 \qquad \text{For the repressilator}$$

$$\left|S_R(p_3^*)\right|S_{AA}(\varphi_1 R(p_3^*)) \geq Sec\left(\frac{\pi}{6}\right)^6 \cong 2.37 \qquad \text{For the activilator}$$

Now, we'll prove that the function $S_{AA}(p)$ is an decreasing function. We also note that $S_{RR}(p)$ is not monotone.

Consider

$$\frac{dS_{RR}(p)}{dp} = \varphi_2 S_R'(\varphi_2 R(p))R'(p)S_R(p) + S_R(\varphi_2 R(p))S_R'(p)$$

We know that $S_R(p) < 0$, $S_R'(p) < 0$ and $R'(p) < 0$, so we have 3 negative signs in the first term and two negative signs in the second term, so we cannot say that $\frac{dS_{RR}(p)}{dp}$ is definitively positive or negative. But in the case of the activilator

$$\frac{dS_{AA}(p)}{dp} = \varphi_2 S_A'(\varphi_2 A(p))A'(p)S_A(p) + S_A(\varphi_2 A(p))S_A'(p)$$

We know that $S_A(p) > 0$, $S_A'(p) < 0$ and $A'(p) > 0$, so $\frac{dS_{AA}(p)}{dp}$ is always negative, i.e. $S_{AA}(p)$ is a decreasing function.

Figure 6-4 shows the functions $S_{RR}(p)$ and $S_{AA}(p)$ when $\varphi_3 = 1, \varphi_2 = 1$

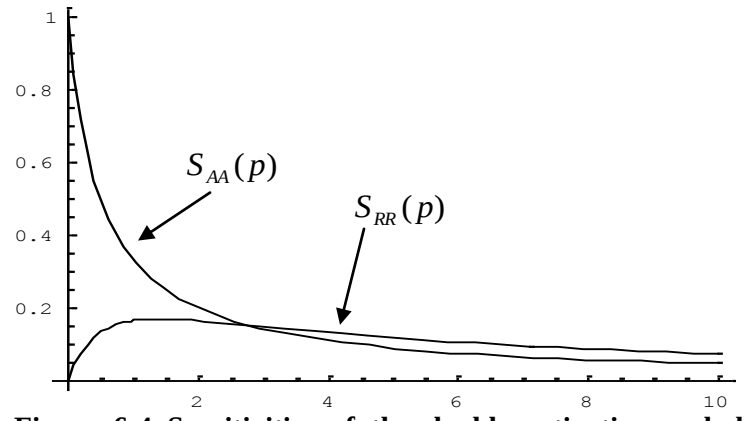


Figure 6-4 Sensitivities of the double activation and double repression functions. The sensitivity of double activation is a decreasing function, while the sensitivity of double repression is not monotonous

Since $S_{AA}(p)$ is decreasing, $S_{AA}(\varphi_1 R(p))$ is an increasing function, and so is the lhs of

the secant condition (which is not the case for the repressilator)

$\left|S_R(p_3^*)\right| S_{AA}(\varphi_1 R(p_3^*))$ Increases with p_3^*

The easiest way to increase the value of p_3^* is increasing φ_3. These can be seen in the equation for calculating the fixed point p_3^*

$$\varphi_3 A(\varphi_2 A(\varphi_1 R(p_3^*))) = p_3^*$$

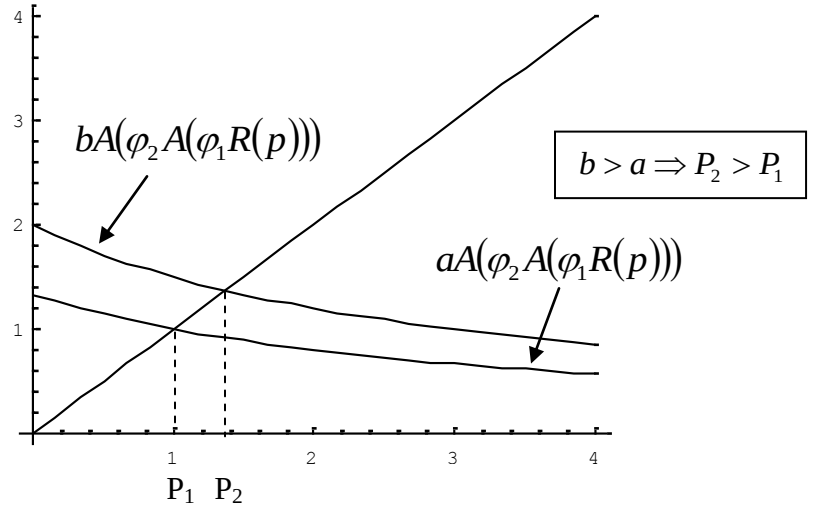


Figure 6-5. Fixed point for protein 3 in the activilator. The fixed point of protein 3 (p_3^*) increases with increased parameter φ_3

This means that by increasing the value of φ_3 experimentally (for example by making the decay rate of the protein or mRNA smaller i.e. protein or mRNA more stable) the lhs of the secant condition can be maximized, thereby maximizing the likelihood of oscillations by manipulating only one gene, without knowing the exact values of the parameters. This represents a clear advantage for the genetic clock constructor over the repressilator design.

Conclusion

We have shown how the lhs of the secant condition can be increased by increasing gene expression (either by increasing transcription/translation rates or by reducing mRNA/protein turnover rates) of gene 3 in the activilator, independent of the expression levels of gene 1 and 2. While this is true for the activilator, it is not the case for the repressilator. An increase in the lhs of the secant condition increases the probability of an instable fixed point of the system, resulting in sustained oscillations in the expression of all the system genes.

By allowing the likelihood of oscillations to increase by manipulation of a single network gene, the activilator network design offers a significant advantage over the repressilator network design. One common problem when constructing synthetic genetic clocks is that once the network of genes is assembled, the parameter set of the system falls outside of the region in parameter space that produces sustained oscillations. Then the question becomes how to modify the parameters of system in order to fall in the oscillatory regime. In the activilator, the answer to this question becomes simple: just increase the expression of gene 3 and eventually the system will fall in the oscillatory regime. In practice, gene expression can be modified, for example, by increasing the strength of a promoter or ribosome binding site with the help of published libraries. In the repressilator, the effect on the lhs of the secant condition by manipulating a single gene cannot be known without knowing the system parameters; therefore it is not easy to manipulate the system to make it oscillate once a particular construction falls outside the oscillatory regime.

Note on building an activilator: To construct an activilator in E.coli cells, we wanted to insert an additional module into the design of the Ninfa clock. Such a module would be activated by the activator module, and it would activate in turn the repressor gene. The resulting topology would be identical to the activilator (with the exception of the auto activation loop of the activator module). The new modules constructed for this purpose were glnKp-nifA and nifHp-LacI (see thesis appendix) for one possible implementation, and glnKp-AraC and AraBp-LacI (see thesis appendix) for another implementation. The recombination of these modules into the chromosome was not finalized, but the modules in plasmid form are available for future research in the lab.

Appendix: Biochemical reactions of repression and activation

We now show how the biochemical reaction of a protein binding to a DNA site gives rise to the same parameter (maximum rate) for activation and repression, given mass-action dynamics.

Consider the binding of a protein (transcription factor), denoted T_f, to a DNA site, denoted S as a mass-action reaction

$$S + T_f \underset{k_{-1}}{\overset{k_1}{\leftrightarrow}} ST_f$$

With the equation

$$\frac{d[ST_f]}{dt} = k_1[S][T_f] - k_{-1}[ST_f]$$

Where *[S]* is the probability that the binding site *S* in empty, *[S T_f]* is the probability that the binding site is bound with protein, and *[T_f]* is protein concentration

If we assume rapid equilibrium, since the dynamics of DNA binding are faster that the dynamics of transcription, we have

$$[ST_f] = K[S][T_f] \quad K \equiv \frac{k_1}{k_{-1}}$$

Since *[S]* and *[S T_f]* are complementary probabilities, we also have

$$[ST_f] + [T_f] = 1$$

replacing this in the previous equation we have

$$[ST_f] = K(1 - [ST_f])[T_f] \quad [ST_f] = \frac{[T_f]/K}{1 + [T_f]/K}$$

$$1 - [S] = K[S][T_f] \quad [S] = \frac{1}{1 + [T_f]/K}$$

For a repressor, the expression of a gene is assumed to be constitutive when the DNA binding site is empty, while expression is blocked when the site is bound

For an activator, mRNA production happens only when the protein is bound to the DNA site

$$\frac{dm}{dt} = \alpha[ST_f] = \alpha \frac{[T_f]/K}{1 + [T_f]/K}$$

$$\frac{dm}{dt} = \alpha[S] = \frac{\alpha}{1 + [T_f]/K}$$

This way, as both activation and repression rates are proportional to the probability a site is bound. We also can consider the same parameter α for the maximum transcription rate.

In the case of the protein T_f forms a complex of n molecules before binding to DNA, the same expressions for transcription rates can be obtained, with $[T_f]$ replaced with $[T_f]^n$. These are the familiar Hill-type functions, common in models of genetic regulation

LIST OF REFERENCES CITED IN CHAPTER 6

Conrad E, Mayo AE, Ninfa AJ and Forger DB. *Rate constants rather than biochemical mechanism determine behaviour of genetic clocks*. J R Soc Interface. 2008 Aug 6;5 Suppl 1:S9-15

El Samad H, Del Vecchio D and Khammash M, *Repressilators and promotilators: Loop dynamics in synthetic gene networks* Acc: proceedings of the 2005 american control conference, vols 1-7 - 2005, pp. 4405–4410

Englesberg E, Irr J, Power J and Lee N. *Positive control of enzyme synthesis by gene C in the L-arabinose system*. J Bacteriol (1965) 90, 946–957.

Jacob F and Monod J *Genetic regulatory mechanisms in the synthesis of proteins.* Journal of Molecular Biology Volume 3, Issue 3, June 1961, Pages 318-356

Mallet-Paret J and Smith HL *The Poincare-Bendixson theorem for monotone cyclic feedback systems* Journal Of Dynamics And Differential Equations 1990 Volume 2, Number 4, 367-421

Muller-Hill B *The Lac Operon: A Short History of Genetic Paradigm* 1996 (Walter de Gruyter, Berlin; New York), pp. 158-169

Othmer, HG *The qualitative dynamics of a class of biochemical control circuits*. J. math Biol. 1976 3, 53-78.

Schleif R *Regulation of the L-arabinose operon in Escherichia coli* in SGM symposium 61: Signals, switches, regulons and cascades: control of bacterial gene expression. Editors D. A. Hodgson, C. M. Thomas. 2002 Cambridge University Press.

Thron CD, *The secant condition for instability in biochemical feedback control – Parts I and II*, Bull. Math. Biol., vol. 53, pp. 383–424, 1991.

Tyson J J and Othmer HG. *The dynamics of feedback control circuits in biochemical pathways*. 1978 Pro 9. theor. Biol. 5, 1-62_

APPENDIX

CONSTRUCTION OF SYNTHETIC MODULES

This appendix details the construction of various synthetic genetic modules. These modules were constructed as part of projects for new genetic clocks that were not included in this thesis, as well as other synthetic modules used for other projects in the Ninfa lab. All of the modules were constructed in plasmids that contain a multiple cloning site (MCS) next to an antibiotic marker flanked by sequences homologous to chromosomal regions, in such a way that when the plasmid is linearized and transformed into a strain with an inhibited recBCD complex it can recombine into the chromosome, at the specific site where the homologous sequences in the chromosome and the plasmid match. This type of plasmid is named a "landing pad", and it has been described previously (Ninfa et al 2007; Ninfa et Al 2009). The schematics of landing pads are shown in figure A-1.

In the construction of the modules described here, I used two landing pads previously constructed in the Ninfa lab by postdoc Don Eun Chang. These landing pads are the Leucine landing pad called pLeuLP, with homologous regions to the LeuCD genes and the setA gene, and the eda landing pad called pEdaLP, with homologous regions to the eda gene and the purT gene. Both landing pads were constructed on top of the plasmid pBR322 (Bolivar et al 1977). Both of these landing pads have a chloramphenicol resistant marker next to the MCS to allow selection. The

chloramphenicol resistant marker in pLeuLP and pEdaLP was later replaced with a kanamycin resistant marker flanked with FTR sequences (cloned from pKD4, see Datsenko and Wanner 2000) in order for the marker to be deleted after the linearized plasmid was integrated into the chromosome. The resulting landing pads were named pLeuLP3 and pEdaLP3.

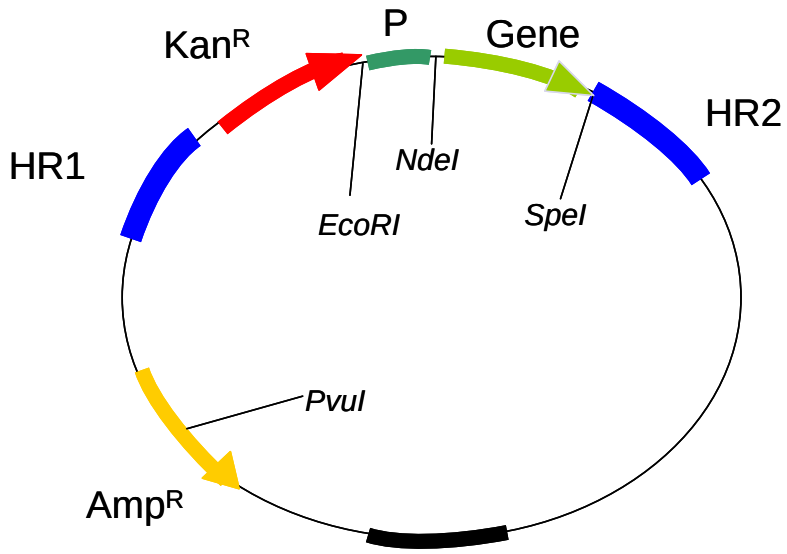


Figure A1 Schematics of landing pads. Homologous regions (blue) flank a module, that consists of an antibiotic marker (red), a promoter (dark green) and a gene (green)

To construct the desired modules, a gene and a promoter region were cloned into the MCS of one of the landing pads. In most cases the promoter was cloned as a EcoRI-NdeI fragment (with the exception of the wild type glnK promoter that was cloned as a EagI-NdeI fragment), and the gene was cloned as a NdeI-SpeI fragment. The genes *LacI* and *AraC* were PCR amplified from existing plasmids in the lab (equivalently they could have been PCR amplified directly from the E.coli

chromosome), while the gene nifA from Klebsiella Oxycota was PCR amplified from the plasmid pTacnifA, also named pJES851 (He et al 1998), and the gene gpfmut3 was PCR amplified from the plasmid pGfpmut3 (Cormack et al 1996), which was a gift from professor Michele Swanson. The promoters used for synthetic modules were the glnK, glnL, glnG, glnAp2 and AraB promoters, PCR amplified from existing lab plasmids, the hybrid promoter composed of

the glnKp promoter region up to the start of transcription fused to the leading sequence of the gene atpE (called here the glnKpE promoter), cloned from the plasmid pglnKpLacI15, and the nifH promoter from Klebsiella Oxycota, PCR amplified from the plasmid pRT22 (Tuli and Merrick 1988), which was a gift from professor Ray Dixon.

Following is the list of plasmids constructed

pLeuLP3 The chloramphenicol marker in pLeuLP (EcoRI-SacI fragment) was replaced with the ftr-Kanamycin marker PCR amplified from pKD4 using the primers EcoRI-ftr-km-US and SacI-ftr-km-DS

pEdaLP3 The chloramphenicol marker in pEdaLP (EcoRI-SacI fragment) was replaced with the ftr-Kanamycin marker PCR amplified from pKD4 using the primers EcoRI-ftr-km-US and SacI-ftr-km-DS

pLeuglnKp the glnKpE promoter was digested as a EcoRI-NdeI fragment from plasmid pglnKpLacI15 and cloned into pLeuLP to produce the plasmid pLeuglnKp. This plasmid was constructed by undergraduate student Amrit Misra.

pLKgfp2 The gene gfpmut3 was PCR amplified from the plasmid pGfpmut3 as a NdeI-SpeI fragment with the primers gfp3-US and gfp3-DS and cloned into NdeI-SpeI digested pLeuglnKp.

pLKgfp3 The chloramphenicol marker (EcoRI-SacI fragment) in pLKgfp2 was replaced by frt-flanked kanamycin marker PCR amplified from pKD4 using the primers EcoRI-ftr-km-US and SacI-ftr-km-DS.

pLKLacI The gene LacI was PCR amplified as a NdeI-SpeI fragment from the plasmid pglnKpLacI2 using the primers LacI-US and lacI-DS and inserted in NdeI-SpeI digested pLKgfp2

pLKLacI2 The chloramphenicol marker in pLKLacI (EcoRI-SacI fragment) was replaced with the ftr-Kanamycin marker PCR amplified from pKD4 using the primers EcoRI-ftr-km-US and SacI-ftr-km-DS

pLHLacI The nifH promoter was PCR amplified from pRT22 as a EcoRI-NdeI fragment with primers nifH-US and nifH-DS and cloned into EcoRI-NdeI digested pLKLacI.

pLeuLp3tet the tetracycline resistant marker was PCR amplified as a EcoRI-SpeI fragment from pBR322 using the primers tetR-US2 and tetR-DS and cloned into EcoRI-SpeI digested pLeuLP3. This was done to introduce XmaI and SalI sites next to the EcoRI site

pEHLacI The SacI-NdeI fragment from pLHLacI (containing the chloramphenicol marker and the nifH promoter) was digested and cloned into SacI-NdeI digested pEdaLP3

pEHLacI2 The chloramphenicol marker in pEHLacI (EcoRI-SacI fragment) was replaced with the ftr-Kanamycin marker PCR amplified from pKD4 using the primers EcoRI-ftr-km-US and SacI-ftr-km-DS

pLKLacI3 The glnKp promoter was PCR amplified from plasmid pglnKpLacI2 as a EagI-NdeI fragment using the primers glnKp-US1 and glnKp-DS and cloned into EagI-NdeI digested pLHLacI

pLKLacI4 The XmaI-SpeI fragment in pLKLacI3 (containing the glnKp promoter fused to *LacI*) was digested and cloned into XmaI-SpeI digested pLeuLP3tet

pLKnifAx The gene *nifA* was PCR amplified from pTacnifA as a NdeI-SpeI fragment using the primers nifA-KX-US and nifA-KX-DS and cloned into NdeI-SpeI digested pLKLacI

pLArap The *AraC-AraB* promoter region was PCR amplified from pKD46 as a EcoRI-NdeI fragment using the primers Arap-Us and Arap-DS and cloned into EcoRI-NdeI digested pLeuLP3. This plasmid was constructed by undergraduate student David Shouder

pLAraLacI The gene *lacI* was cloned as a NdeI-SpeI fragment from pLKLacI and inserted into NdeI-SpeI digested pLArap

pLAp2LacI The promoter *glnAp2* was PCR amplified from pGlnApLacI as a EcoRI-NdeI fragment using the primers glnApUS and glnApDS and cloned into EcoRI-NdeI digested pLAraLacI

pLLLacI The promoter *glnLp* was PCR amplified from pRSglnLpZYA as a EcoRI-NdeI fragment using the primers glnLp-US and glnLp-DS and cloned into EcoRI-NdeI digested pLAraLacI

pELLacI The SacI-NdeI fragment from pLLLacI (containing the kanamycin marker and the glnL promoter) was digested and cloned into SacI-NdeI digested pEdaLP

pEBLacI The promoter *AraBp* was PCR amplified from pKD46 as a EcoRI-NdeI fragment using the primers AraBp-US and Arap-DS and cloned into EcoRI-NdeI digested pELLacI

pLKAraC The gene *AraC* was PCR amplified as a NdeI-SpeI fragment from pKD46 using the primers AraC-US and AraC-DS and inserted into NdeI-SpeI digested pLKLacI3

pLAraGlnG The gene *glnG* was PCR amplified as a NdeI-SpeI fragment from pGlnG6 using the primers glnG-US2 and glnG-T-DS and inserted into NdeI-SpeI digested pLAraLacI

Primer	Sequence (5'>3')
EcoRI-ftr-km-US	CCCCGAATTCCATTACACGTCTTGAGCGATT
SacI-ftr-km-DS	CCCCGAGCTCTGAATATCCTCCTTAGTTCCT
gfp3-US	AAAAATTAATTTTAAGAAGGAGATATACATATGAGTAAAGGAGAAGAACTTTTC
gfp3-DS	TTTTATTAATGTCGACATGAGTAAAGGAGAAGAACTTTTC
LacI-US	GGGGGAATTCTCTAGACATATGAAACCAGTAACGTTATACGATGTCGC
lacI-DS	GGGGCTCGAGCTGCAGACTAGTCCAAATAAAAAGGCCTGCGATTACCAGCAGGCCT GTTATTTCACTGCCCGCTTTCCAGTCGGGA
nifH-US	GGGGAATTCTCTAGACGGCCGGAATAACAGGCCTGCTGGTAATCGCAGGCCTTTTT ATTTGGAACGCGTTATGAAGAGAG
nifH-DS	GGGGAAGCTTAATTAACATATGGACTTCTCCTGTTGTTGAGTGTTTG
tetR-US2	AAAGAATTCCCCGGGGTCGACTAATGCGGTAGTTTATCACAGTTA
tetR-DS	TTTTCTCGAGCTGCAGACTAGTAGCGCTCCAAATAAAAAGGCCTGCGATTACCAGC AGGCCTGTTATTTCAGGTCGAGGTGGCCCGGCT
glnKp-US1	GGGGCGGCCGCCCGGGGTCGACAATAACAGGCCTGCTGGTAATCGCAGGCCTTTTT ATTTGGCATTGAGCGCCTGAATAGCGCAATA
glnKp-DS	AAAACATATGTCCCCTCCGGTCAGAATTCGGTAATGGTTCTGCTACACGAAGA
nifA-KX-US	GGGGGAATTCTCTAGACATATGATCCATAAATCCGATTCG
nifA-KX-DS	GGGGCTCGAGCTGCAGACTAGTCCAAATAAAAAGGCCTGCGATTACCAGCAGGCCT GTTATTTCACAGTCGCGGCATGGTGA
Arap-US	GGGGAATTCTCTAGACCAAATAAAAAGGCCTGCGATTACCAGCAGGCCTGTTATTT TATGACAACTTGACGGCTACAT
Arap-DS	GGGGACTAGTCATATGTTCACTCCATCCAAAAAAACGGGTATGGAGAAAC
glnApUS	GGGGAATTCTCTAGACGGCCGCCCGGGAATAACAGGCCTGCTGGTAATCGCAGGCC TTTTTATTTGGCAAAGGTCATTGCACCAACATGG
glnApDS	GGGGAAGCTTAATTAACATATGTTAACTCTCCTGGATTGGTCATGG
glnLp-US	GGGGAATTCCCCGGGGTCGACCGGCCGAATAACAGGCCTGCTGGTAATCGCAGGCC TTTTTATTTGGGCGACGCAAATGCGTCTTATC
glnLp-DS	AAAACATATGGCAGTCTCCTGAACAGGTTGCACC
AraBp-US	AAAAGAATTCCCCGGGGTCGACCGGCCGCCAAATAAAAAGGCCTGCGATTACCAGC AGGCCTGTTATTCATACTTTTCATACTCCCGC
AraC-US	TTTTCATATGGCTGAAGCGCAAAATG
AraC-DS	GGGGCTCGAGCTGCAGACTAGTCCAAATAAAAAGGCCTGCGATTACCAGCAGGCCT GTTATTTTATGACAACTTGACGGCTACAT
glnG-US2	GGAATTCCATATGCAACGAGGGATAGTCTGGGTAGTC
glnG-T-DS	CCCCCTCGAGCTGCAGACTAGTCCAAATAAAAAGGCCTGCGATTACCAGCAGGCCT GTTATTTCACTCCATCCCCAGCTCTTTT

Table A1. Primers used for construction of synthetic Modules

LIST OF REFERENCES CITED IN APPENDIX

Bolivar F, Rodriguez RL, Greene PJ, Betlach MC, Heynecker HL and Boyer HW *Construction and characterization of new cloning vehicles. II. A multipurpose cloning system*. 1977 Gene, 2, 95-113.

Cormack BP, Valdivia RH and Falkow S. *FACS-optimized mutants of the green fluorescent protein (GFP)* 1996 Gene Volume 173, Issue 1, Pages 33-38

Datsenko A and Wanner BL *One-step inactivation of chromosomal genes in Escherichia coli K-12 using PCR products* PNAS 2000 vol. 97 no. 12 6640-6645

He L, Soupene E, Ninfa AJ and KustuS. *Physiological Role for the GlnK Protein of Enteric Bacteria: Relief of NifL Inhibition under Nitrogen-Limiting Conditions* Journal of Bacteriology 1998, p. 6661-6667, Vol. 180, No. 24

Note on this reference: The strain studied in this paper is presented as Klebsiella Pneumoniae, however it was later revealed to be actually Klebsiella Oxycota

Ninfa AJ, Selinsky S, Perry N, Atkins S, Xiu Song Q, Mayo A, Arps D, Woolf P and Atkinson MR *Using two-component systems and other bacterial regulatory factors for the fabrication of synthetic genetic devices*. 2007 Methods Enzymol 422, 488-512.

Ninfa AJ, Atkinson MR, Forger DB, Atkins S, Arps D, Selensky S, Court D, Perry N and Mayo A. *A synthetic biology approach to understanding biological oscillations: developing a genetic oscillator for Escherichia coli*. In: Bacterial Circadian Programs (Ditty JL, Mackey SR, Johnson CH, eds) 2009 Springer-Verlag, Berlin pp301-329

Tuli R and Merrick MJ. *Over-production and characterization of the nifA gene product of Klebsiella pneumoniae--the transcriptional activator of nif gene expression*. J Gen Microbiol. 1988 Feb;134(2):425-32.

Note on this reference: The strain studied in this paper is presented as Klebsiella Pneumoniae, however it was later revealed to be actually Klebsiella Oxycota