INVESTIGATING THE STRUCTURAL PROPERTIES OF ACTIVATORS AND REPRESSORS IN EUKARYOTIC CIRCADIAN CLOCK

A Thesis

Presented to the Faculty of the Graduate School

of Cornell University

In Partial Fulfillment of the Requirements for the Degree of

Master of Science

by

Kritika Dusad

May 2017

ABSTRACT

Many organisms have an internal circadian clock that helps them adapt to the external changes in light and temperature. The clock is important for cell cycle regulation and genome integrity. Disrupting the internal circadian clock can cause depression, obesity, cardiovascular disease, and even cancer. Importantly, circadian clocks are highly conserved among eukaryotes. Thus, we decided to focus on the well-characterized model systems- fungus *Neurospora crassa* and fly *Drosophila melanogaster* to study the structural properties of the interacting circadian clock components. These clock components make up the positive and negative arms of the circadian oscillators. The *Neurospora crassa* circadian clock consists of a positive regulator White Collar Complex (WCC) that activates the transcription of clock gene frequency. Frequency (FRQ) acts as a repressor in the negative arm of the oscillator. FRQ acts as a repressor by binding with its partner Casein Kinase 1 (CK1) to phosphorylate and inactivate WCCs. Similar to FRQ, the *Drosophila* circadian clock is composed of a repressor Period (PER) that phosphorylates and inactivates *Drosophila* positive activators. While many biochemical and genetic studies have explored the functional aspects of the circadian clock components, very little is known about the mechanisms of the interactions with each other. Therefore, an analogous study of the components in both model systems would provide better understanding of the respective protein structures and interactions. In this thesis, we show that the *N. crassa* WCC can recognize and tightly bind to *frq* promoter regions in the absence of WC1 photosensor domain. Secondly, *N. crassa* repressor FRQ is stabilized and phosphorylated if co-purified with

binding partner CK1. Additionally, FRQ:CK1 has a more flexible conformation when it is hyperphosphorylated. Lastly, like FRQ, repressor PER in *D. melanogaster*, is stabilized by phosphorylation on specific sites. This shows structural similarities between these two repressors FRQ and PER. These results provide a deeper understanding of the molecular bases of the circadian clocks in *N. crassa* and *D. melanogaster* to pave way for future structural studies of many eukaryotic clock proteins.

BIOGRAPHICAL SKETCH

Kritika Dusad was born in New Delhi, India. She developed an interest in science when her mother brought home a book full of experiments. During high school, Kritika decided to pursue biology at an undergraduate level in the United States. She went to University of Chicago for undergraduate studies, where she graduated with Honors in 2010 with a Bachelor of Science in Biology with a specialization in cancer biology. At University of Chicago, Kritika developed an interest in biochemistry. Thus, she decided to pursue an M.S. in biochemistry at Cornell University. Under the advisement of Dr. Brian Crane, Kritika investigated the structural properties circadian clock proteins. In the future, she plans to use her experience in research to do science outreach and design hands-on activities for kids and adults.

To my grandparents

ACKNOWLEDGMENTS

I would like to thank my advisor Brian Crane for his immense about encouragement and understanding. He taught me to go after the right problems, to never stop searching for answers and to always stay motivated even if nothing is working. I would also like to thank Richard Cerione and Hening Lin who provided invaluable guidance on my research and this thesis.

My lab-mates were extremely helpful throughout my graduate school career. Without their support, this thesis would not have been possible. I would especially like to thank Estella for being a wonderful friend and an amazing mentor. I am also thankful to Changfan for being so positive and helpful.

I am so grateful to my family. Ma, papa and Ritika have always supported me in whatever I've wanted to do. My graduate career would have been extremely difficult and boring without Ayush. Thank you!

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

CHAPTER 1

INTRODUCTION

Circadian Clock

Organisms have an internal oscillator that allows them to adapt to changes in the environment. This oscillator, or circadian clock, can persist in synchrony with the external environment in the absence of stimuli[1]. These endogenous clocks, which are present in many cell types such as liver cells and neurons, generate internal timing. The molecular clocks present in the brain are called "master" clocks because they establish synchrony in the surrounding cells. The master clock and peripheral clocks (in tissues) are composed of transcriptional/translational feedback loops that possess positive and negative elements, which together produce 24-hour rhythms of gene expression[2]. Since circadian clocks are conserved across eukaryotes, they share common mechanisms in the regulation of the circadian clock network. Thus, studying the molecular basis of circadian clock elements in well-behaved and well-studied eukaryotic systems, like those of *Drosophila melanogaster* and *Neurospora crassa*, will be particularly useful[3].

Circadian Clocks and Diseases

Circadian clocks are vital to the health of eukaryotes, especially humans. Circadian clocks allow humans to adapt their internal processes to changes in the external environment. Disrupting the circadian clock by travelling across time zones or working night shifts causes disorder to physiological processes. Some of the affected processes are cell-cycle progression, DNA repair, glucose homeostasis, and body temperature[4]. Thus, problems in the circadian rhythm of humans and other mammals can cause diseases such as insomnia, depression, bipolar disorder, and cancer[5,6]. Since the

circadian clock has a huge impact on health, studying the structure and properties of proteins involved would be fruitful for therapeutic discoveries.

Conservation in Eukaryotic Circadian Clocks

Even though *N. crassa, D. melanogaster*, and other eukaryotes are evolutionarily distant, they have common features in their circadian clocks, such as the presence of autoregulatory feedback loops[7]. These feedback loops are made of positive and negative elements that interact and regulate each other. For example, the positive elements transcriptionally activate the expression of negative elements and in turn the negative elements inhibit the activity of positive elements. This interplay between the positive and negative elements establishes the negative feedback loop. Additionally, the negative elements in circadian clocks are post-translationally regulated by phosphorylation. Specifically, these negative elements are step-wise phosphorylated and are ultimately degraded via ubiquination and proteasome systems. Lastly, the clock components across different eukaryotes are functionally conserved, and often homologous. For example, the negative elements FRQ in *N. crassa* and PER in *D melanogaster* are phosphorylated by analogous casein kinases and dephosphorylated by analogous phosphatases. These negative elements are progressively phosphorylated and are thus degraded by conserved ubiquitin ligases[8]. Thus, analogous studies on the structures of these elements from dissimilar organisms like *N. crassa* and *D. melanogaster* may provide a uniform understanding of key mechanisms in circadian biology.

Neurospora crassa as a Model Organism

The circadian clock of *N. crassa* can be monitored by observing the rhythmic and alternative production of conidia and mycelia throughout the subjective day..

Neurospora is grown in a glass tube in constant light (LL) and subsequent constant darkness (DD) to entrain its circadian clock. Thus, the length of the filamentous growth between spore bands in the tube correlates with the time since light to dark transition. This so-called "race tube" assay has been widely used to identify various central circadian clock proteins and the basic circadian rhythm network[9]. *Neurospora crassa* has been an excellent model organism for studying circadian clocks due to ease of growth, measurable circadian clock patterns, and a well-characterized genome.

Circadian Clock in Neurospora crassa

The core circadian oscillator of *N. crassa* consists of autoregulatory negative and positive feedback loops. The core components of the negative limb of the Neurospora clock are the repressor Frequency (FRQ) that binds to FRQ-interacting RNA helicase (FRH) to form a complex called FFC. The positive limb is made of transcription factors White Collar-1 (WC-1) and White Collar-2 (WC-2)[10]. These proteins make a heterodimeric complex called the White Collar Complex (WCC) that forms around subjective late night[11]. The complex binds to the *frequency* (*frq*) gene promoter and activates the transcription of *frq*. During late subjective day, *frq* mRNA is at its peak transcript level, and after 4 – 6 hours, the FRQ protein reaches its peak expression level. FFC translocates to the nucleus and inhibits the activity of WCC, resulting in a decrease in the level of *frq*-mRNA[12]. Additionally, FRQ levels are also regulated by post-translational modifications such as phosphorylation by kinases and dephosphorylation by phosphatases. Casein Kinases I and II progressively phosphorylate FRQ upon synthesis and, upon hyperphosphorylation, FRQ is degraded by the ubiquitin ligase FWD-1. Consequently, FRQ levels decrease and result in reactivation of WCC around

subjective late night[13]. Thus, the cycle of expression of *frequency* gene begins again. In addition to the post-translational modification of FRQ, there are other negative feedback loops that are central to the functioning of *N. crassa* clock. For example, down-regulation of WCC transcription by photoreceptor like Vivid (VVD). This protein contains a Light-Oxygen-Voltage (LOV) domain that detects light. Upon illumination, VVD changes its conformation and interacts with the WCC to decrease WCC transcription activity[14] (Figure 1).

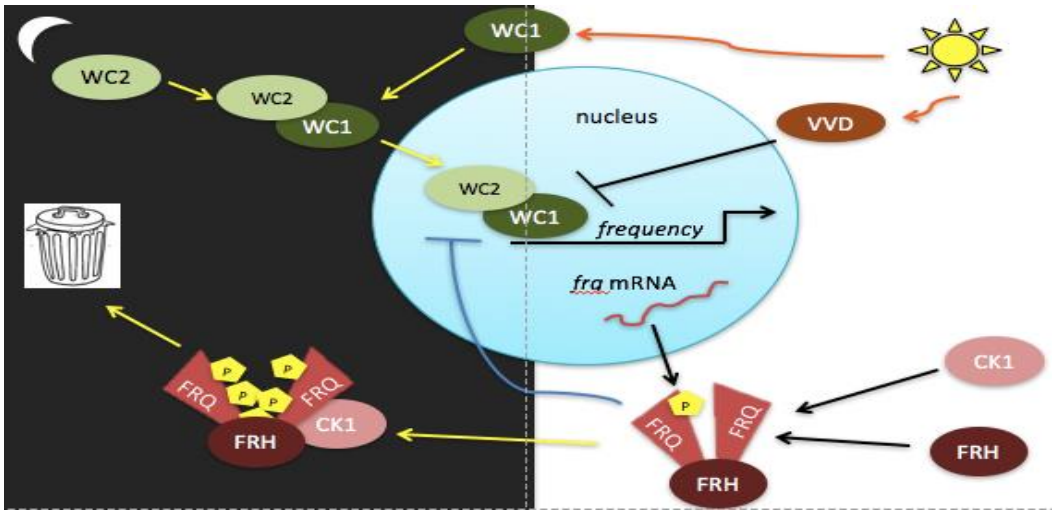

Figure 1: The Neurospora crassa circadian clock

Neurospora White Collar-1 and White Collar-2 Complex.

WC-1 contains three PAS (Per-Arnt-Sim) domains. These domains are composed of five anti-parallel beta-strands and alpha helices that form the cofactor-binding pocket and N-terminal/C-terminal helix[15]. The N-terminal-most PAS domain of WC-1 belongs to a special PAS subclass called a Light-Oxygen-Voltage (LOV) domain. This domain is proposed to bind to flavin adenine dinucleotide and under blue light irradiation and activate gene expression. The C-terminal-most PAS domain binds to the PAS domain of WC-2 to form the WCC. Some studies suggest that WC-1 binds to WC-2 to form a

heterodimer in the dark and a multimeric complex in the presence of light and FAD[16]. Additionally, both WC-1 and WC-2 proteins each contain one Zn finger domain for binding GATA-like DNA sequences. Studies suggest the WCC binds to the *frq* promoter using their Zn finger domains[17]. Specifically, the Zn finger domain and a proximal basic region of WC-2 are required for WCC to bind to DNA in the presence of light. However, in the absence of light, Zn finger domains of both WC-1 and WC-2 are required for WCC to bind to DNA[18]. While genetic studies have found that the WCC binds to specific Light Response DNA elements (LREs), the details of how the WCC recognizes an LRE are currently unknown.

Neurospora Frequency and Casein Kinase Complex.

Frequency (FRQ) is a major repressor in the *Neurospora* circadian clock. It has a coiled-coil region necessary for making a homodimer, a CK1 binding region, FRH binding domain, and two protein degradation domains that contain Proline, Glutamic Acid, Serine and Threonine residues (PEST). FRQ, with its binding partner FRH, make the FFC, a complex that inhibits the activity of WCC. FRQ is progressively phosphorylated by CK1 as soon as it is synthesized. This FRQ-CK1 complex is also responsible for phosphorylating the WCC, inhibiting its activity, and closing the circadian negative feedback loop[19]. CK1 phosphorylates the middle of FRQ that is followed by the two PEST domains. Studies suggest that hypophosphorylated FRQ has a closed conformation versus a hyperphosphorylated open conformation. The open conformation of FRQ may occur due to the repulsive electrostatic forces present in the hyperphosphorylated form[20] (Figure 1). Due to its variable phosphorylated states, the structure of FRQ is challenging to characterize and is still unknown.

Drosophila melanogaster as a Model Organism.

Fruit fly *Drosophila melanogaster* is a widely used model system for genetic and biochemical studies. *Drosophila* exhibits measurable circadian patterns in its locomotor behaviors. Locomotor activity peaks two times a day – after dawn and before dusk. The locomotor activity is measured by putting a single fly in a tube containing food. The tube is crossed by an infrared light beam that when crossed by a fly, sends an electronic signal to the detector[21]. Additionally, other behaviors like eclosion (emergence of adult fly from pupal case), egg laying, courtship, and learning and memory are also used to monitor the fly circadian clock. Furthermore, there are only 150 clock cells in each brain hemisphere in a fly, facilitating the study of single or cluster of clock cells in the fly brain[22]. Since *Drosophila* is an organism that is easy to work with, genetically well-characterized, and well-suited to study neural circadian responses, it is ideal for studying circadian clock proteins.

Circadian Clock in Drosophila melanogaster.

Similar to the core circadian oscillator in *Neurospora crassa*, *Drosophila* has negative and positive feedback loops that make up the endogenous clock. Here, the positive elements of the loop are Clock (CLK) and Cycle (CYC). These proteins make a heterodimer and bind to the promoter region of *period* gene. Additionally, this CLK:CYC complex transcriptionally activates the expression of other clock associated output genes, like *timeless*. Consequently, the negative limb of the *Drosophila* clock is comprised of Period (PER) and its binding partner Timeless (TIM). After sunset, the concentrations of PER and TIM peak. This complex subsequently translocates to the nucleus, where PER:TIM represses the transcription activity of CYK:CLK heterodimer

and negatively regulates further PER:TIM production. Additionally, PER is further regulated by the action of kinases like Doubletime (DBT)[23]. Similar to *Neurospora* kinase CK1, DBT progressively phosphorylates PER. Once PER is extensively phosphorylated, it is recruited by the E3 ubiquitin ligase SLIMB for degradation. The degradation of PER results in reactivation of transcriptional activity of CYC:CLK complex and production of clock proteins PER and TIM again[24] (Figure 2).

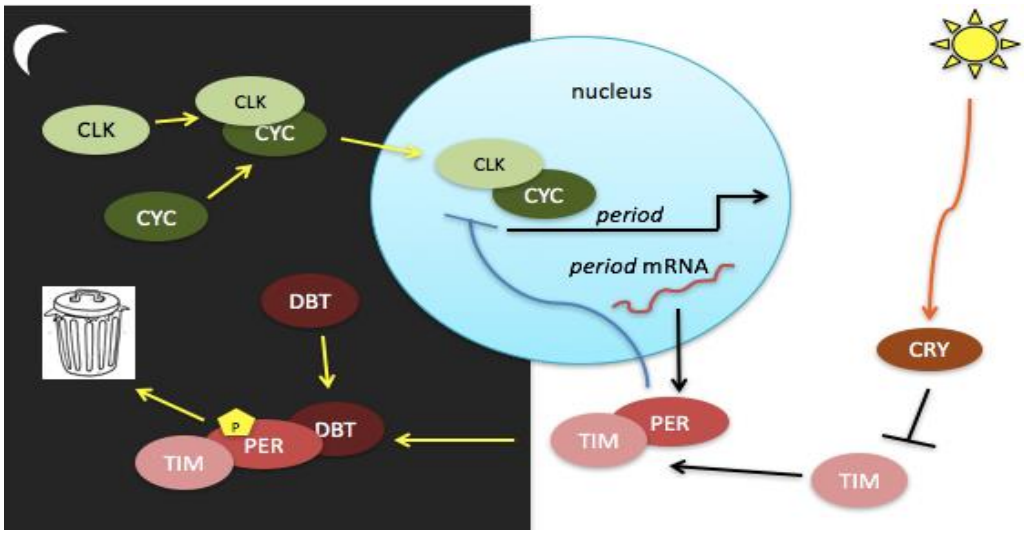

Figure 2: The Drosophila melanogaster circadian clock

Drosophila PERIOD and its Phosphorylation States.

Similar to FRQ, PER is a major repressor in *Drosophila* circadian clock. With its binding partner TIM, PER inactivates the positive regulators CLK:CYC to close the circadian negative feedback loop. PER has two PAS domains that allow PER to form a homodimer, a post translation modification (PTM) domain, DBT binding domain, and TIM binding domain. The two PAS domains have been crystallized and show a homodimer forming with the N-terminal PAS domain interacting with the αF (alpha-helical extension) of the opposite subunit. While the αF is essential for dimer formation, its extension suggests it unlatches from itself in order to associate to TIM[25,26]. Similar

to *Neurospora* FRQ, PER is phosphorylated in a step-wise fashion by kinase DBT. This allows gating of PER degradation by affecting the conformation of the protein. In a hyperphosphorylated state, PER has an open conformation that allows access to SLIMB ubiquitin ligase. Mutations in phosphorylation sites result in short period length in *Drosophila*[27]. Thus, similar to FRQ, PER has a variable structure due to its different phosphorylated states (Figure 2).

In the following chapters, I describe efforts to purify different variants of WCC with and without Zinc finger (ZnF) domains that can bind target *frq* promoter regions, proximal Light Regulatory Element (pLRE) and Clock-box (C-box). Through this work, I show that only WC2 with ZnF domain (WC2/ZnF) is important to bind pLRE tightly. Additionally, if both WC1 and WC2 ZnF domains are present, the affinity of the complex to pLRE is low. On the other hand, both WC1 and WC2 ZnF domains are required to tightly bind to C-box promoter region of *frq*. Absence of either of the ZnF domain results in weak or no binding. These results have important implications for understanding how WCC ZnF domains recognize and bind to different promoter regions of *frq*. I also describe efforts to purify FRQ with CK1. Through this work, I show that FRQ requires a functional CK1 for stability. Additionally, hyperphosphorylation of FRQ changes its conformation. These results have important implications for understanding how phosphorylation of FRQ results in its degradation due to a change in its conformation. Lastly, I also describe efforst to study the changes in conformation of PER and its phosphomimic mutants. These results have important implications in understanding how phosphorylation of PER affects its structural stability.

REFERENCES

1. Dunlap, J. C. Molecular bases for circadian clocks. *Cell* **96,** 271–290 (1999).
2. Mendoza-Viveros, L. *et al.* Molecular modulators of the circadian clock: lessons from flies and mice. *Cell. Mol. Life Sci.* **74,** 1–25 (2016).
3. Tataroglu, O. & Emery, P. Studying circadian rhythms in Drosophila melanogaster. *Methods* **68,** 140–150 (2014).
4. Partch, C. L., Green, C. B. & Takahashi, J. S. Molecular architecture of the mammalian circadian clock. *Trends Cell Biol.* **24,** 90–99 (2014).
5. Li, M.-D., Li, C.-M. & Wang, Z. The Role of Circadian Clocks in Metabolic Disease. *Yale J. Biol. Med.* **85,** 387–401 (2012).
6. Takahashi, J. & Hong, H. The genetics of mammalian circadian order and disorder: implications for physiology and disease. *Nat. Rev. …* **9,** 764–775 (2008).
7. Merbitz-Zahradnik, T. & Wolf, E. How is the inner circadian clock controlled by interactive clock proteins?. Structural analysis of clock proteins elucidates their physiological role. *FEBS Lett.* **589,** 1516–1529 (2015).
8. Liu, Y. & Bell-Pedersen, D. Circadian rhythms in Neurospora crassa and other filamentous fungi. *Eukaryot. Cell* **5,** 1184–1193 (2006).
9. Baker, C. L., Loros, J. J. & Dunlap, J. C. The circadian clock of Neurospora crassa. *FEMS Microbiol. Rev.* **36,** 95–110 (2012).
10. Talora, C., Franchi, L., Linden, H., Ballario, P. & Macino, G. Role of a white collar-1-white collar-2 complex in blue-light signal transduction. *EMBO J.* **18,** 4961–4968 (1999).

11. Froehlich, A. C., Liu, Y., Loros, J. J. & Dunlap, J. C. White Collar – 1 , a Circadian Blue Light Photoreceptor , Binding to the frequency Promoter. **297,** 815–19 (2002).

12. Cha, J., Zhou, M. & Liu, Y. Mechanism of the Neurospora circadian clock, a FREQUENCY-centric view. *Biochemistry* **54,** 150–156 (2015).

13. Görl, M. *et al.* A PEST-like element in FREQUENCY determines the length of the circadian period in Neurospora crassa. *EMBO J.* **20,** 7074–7084 (2002).

14. Schwerdtfeger, C. & Linden, H. Blue light adaptation and desensitization of light signal transduction in *Neurospora crassa*. *Mol. Microbiol.* **39,** 1080–1087 (2001).

15. Makino, Y. *et al.* Inhibitory PAS domain protein is a negative regulator of hypoxia-inducible gene expression. *Nature* **414,** 550–554 (2001).

16. Ballario, P., Talora, C., Galli, D., Linden, H. & Macino, G. Roles in dimerization and blue light photoresponse of the PAS and LOV domains of Neurospora crassa white collar proteins. *Mol. Microbiol.* **29,** 719–729 (1998).

17. Chen, C.-H., Ringelberg, C. S., Gross, R. H., Dunlap, J. C. & Loros, J. J. Genome-wide analysis of light-inducible responses reveals hierarchical light signalling in Neurospora. *EMBO J.* **28,** 1029–1042 (2009).

18. Wang, B., Zhou, X. Y., Loros, J. J. & Dunlap, J. C. Alternative Use of DNA Binding Domains by the Neurospora White Collar Complex Dictates Circadian Regulation and Light Responses. *Mol. Cell. Biol.* **36,** 781–793 (2016).

19. He, Q. *et al.* Light-independent phosphorylation of WHITE COLLAR-I regulates its function in the Neurospora circadian negative feedback loop. *J.*

Biol. Chem. **280,** 17526–17532 (2005).

20. Querfurth, C. *et al.* Circadian Conformational Change of the Neurospora Clock Protein FREQUENCY Triggered by Clustered Hyperphosphorylation of a Basic Domain. *Mol. Cell* **43,** 713–722 (2011).
21. Chiu, J. C., Low, K. H., Pike, D. H., Yildirim, E. & Edery, I. Assaying locomotor activity to study circadian rhythms and sleep parameters in Drosophila. *J. Vis. Exp.* 1–8 (2010). doi:10.3791/2157
22. Peschel, N. & Helfrich-Förster, C. Setting the clock--by nature: circadian rhythm in the fruitfly Drosophila melanogaster. *FEBS Lett.* **585,** 1435–42 (2011).
23. Saez, L., Meyer, P. & Young, M. A PER / TIM / DBT Interval Timer for Drosophila ' s Circadian Clock. *Cold Spring Harb. Symp. Quant. Biol.* **LXXII,** 69–75 (2007).
24. Crane, B. R. & Young, M. W. Interactive features of proteins composing eukaryotic circadian clocks. *Annu. Rev. Biochem.* **83,** 191–219 (2014).
25. King, H. A., Hoelz, A., Crane, B. R. & Young, M. W. Structure of an Enclosed Dimer Formed by The Drosophila Period Protein. *J. Mol. Biol.* **413,** 561–572 (2011).
26. Yildiz, Ö. *et al.* Crystal structure and interactions of the PAS repeat region of the Drosophila clock protein PERIOD. *Mol. Cell* **17,** 69–82 (2005).
27. Baylies, M. K., Vosshall, L. B., Sehgal, a & Young, M. W. New short period mutations of the Drosophila clock gene per. *Neuron* **9,** 575–581 (1992).
28. Harmer, S. L., Panda, S. & Kay, S. A. Molecular Bases of Circadian Rhythms.

Annu. Rev. Cell Dev. Biol. **17,** 215–53 (2001).

29. Froehlich, A. C., Liu, Y., Loros, J. J. & Dunlap, J. C. White Collar – 1 , a Circadian Blue Light Photoreceptor , Binding to the frequency Promoter. *Science (80-.).* **297,** 815–819 (2002).

30. Hurley, J. M., Larrondo, L. F., Loros, J. J. & Dunlap, J. C. Conserved RNA helicase FRH acts nonenzymatically to support the intrinsically disordered neurospora clock protein FRQ. *Mol. Cell* **52,** 832–843 (2013).

31. Anderson, P. W. *et al.* White Collar – 1 , a Circadian Blue Light Photoreceptor , Binding to the frequency Promoter. **1605,** (2001).

32. Cheng, P., Yang, Y., Wang, L., He, Q. & Liu, Y. WHITE COLLAR-1, a multifunctional Neurospora protein involved in the circadian feedback loops, light sensing, and transcription repression of wc-2. *J. Biol. Chem.* **278,** 3801–3808 (2003).

33. Smith, K. M. *et al.* Transcription factors in light and circadian clock signaling networks revealed by genomewide mapping of direct targets for neurospora white collar complex. *Eukaryot. Cell* **9,** 1549–1556 (2010).

34. Hurley, J. M., Larrondo, L. F., Loros, J. J. & Dunlap, J. C. Conserved RNA helicase FRH acts nonenzymatically to support the intrinsically disordered neurospora clock protein FRQ. *Mol. Cell* (2013). doi:10.1016/j.molcel.2013.11.005

35. Cheng, P., Yang, Y., Heintzen, C. & Liu, Y. Coiled-coil domain-mediated FRQ-FRQ interaction is essential for its circadian clock function in Neurospora. *EMBO J.* **20,** 101–108 (2001).

36. Garceau, N. Y., Liu, Y., Loros, J. J. & Dunlap, J. C. Alternative initiation of translation and time-specific phosphorylation yield multiple forms of the essential clock protein FREQUENCY. *Cell* **89,** 469–476 (1997).

37. Colot, Hildur V., Loros, Jennifer J., Dunlap, J. C. Temperature-modulated Alternative Splicing and Promoter Use in the Circadian Clock Gene frequency. *Mol. Biol. Cell* **16,** 5563–5571 (2005).

38. Baker, C. L., Kettenbach, A. N., Loros, J. J., Gerber, S. A. & Dunlap, J. C. Quantitative Proteomics Reveals a Dynamic Interactome and Phase-Specific Phosphorylation in the Neurospora Circadian Clock. *Mol. Cell* **34,** 354–363 (2009).

39. Querfurth, C. *et al.* Circadian Conformational Change of the Neurospora Clock Protein FREQUENCY Triggered by Clustered Hyperphosphorylation of a Basic Domain. *Mol. Cell* (2011). doi:10.1016/j.molcel.2011.06.033

40. Cheng, P., He, Q., He, Q., Wang, L. & Liu, Y. Regulation of the Neurospora circadian clock by an RNA helicase. *Genes Dev.* 234–241 (2005). doi:10.1101/gad.1266805.234

41. Lauinger, L., Diernfellner, A., Falk, S. & Brunner, M. The RNA helicase FRH is an ATP-dependent regulator of CK1a in the circadian clock of Neurospora crassa. *Nat. Commun.* **5,** 3598 (2014).

42. Peschel, N. & Helfrich-Förster, C. Setting the clock - By nature: Circadian rhythm in the fruitfly Drosophila melanogaster. *FEBS Lett.* **585,** 1435–1442 (2011).

43. Dubowy, C. & Sehgal, A. Circadian Rhythms and Sleep in Drosophila

melanogaster. *Genetics* **205,** 1373–1397 (2017).

44. Chiu, J. C., Vanselow, J. T., Kramer, A. & Edery, I. The phospho-occupancy of an atypical SLIMB-binding site on PERIOD that is phosphorylated by DOUBLETIME controls the pace of the clock. *Genes Dev.* **22,** 1758–1772 (2008).

45. Saez, L. & Young, M. W. Regulation of nuclear entry of the Drosophila clock proteins period and timeless. *Neuron* **17,** 911–920 (1996).

46. Garbe, D. S. *et al.* Cooperative Interaction between Phosphorylation Sites on PERIOD Maintains Circadian Period in Drosophila. *PLoS Genet.* **9,** (2013).

CHAPTER 2

RECOGNITION OF TARGET DNA SEQUENCES BY THE WHITE COLLAR COMPLEX

Introduction

Circadian clocks are important for the behavior and physiology of organisms. These clocks help organisms adapt to daily changes in light and temperature. Circadian clocks from distantly related eukaryotes are similar at the molecular level[1]. The clocks are composed of positive and negative feedback loops and similar regulatory components such as phosphatases and kinases that regulate key clock proteins. Among the eukaryotes, fungus *Neurospora crassa* is a widely-used model system to study circadian clock[2].

The *Neurospora crassa* circadian clock is composed of activators White Collar 1 (WC1)[3] and White Collar 2 (WC2). These activators form a complex (WCC) and act as transcription factors for clock controlled genes such as *frequency* (*frq*)[4]. The WCC binds to the Clock box (C-box) region of *frq* promoter in the dark and proximal Light Regulatory Element (pLRE) region of the *frq* promoter in the light[5]. While the WCC acts as a transcriptional activator, the *frq* gene product, the protein Frequency (FRQ) acts as a repressor. FRQ, an otherwise highly disordered protein, immediately interacts with Frequency-interacting RNA Helicase (FRH) to form a FRQ-FRH complex (FFC)[6]. This complex interacts with Casein Kinase 1 (CK1) and represses the transcriptional activity of the WCC through phosphorylation of the complex (refer to Figure 1).

Although WC1 functions as a transcription factor owing to its Zinc-finger (ZnF) DNA binding domain, it also has a Light-, Oxygen-, Voltage-sensing (LOV) domain that allows it to act as a

photosensor. This LOV domain is one of the three PER-ARNT-SIM (PAS) domains present in WC1[7]. On the other hand, WC-2 only has two PAS domains; they function to bind and stabilize WC1. Like WC1, WC2 possesses one ZnF domain that is required for DNA binding. The WCC changes its structure in the presence of light. While, in the dark state, WCC is a heterodimer, in the light, it is a heterotrimer (two WC1 and one WC2)[8,9].

Additionally, either one or two ZnFs are used to bind to different elements of the *frequency* promoter. In the dark state, both the ZnF domains of WC1 and WC2 are required for the heterodimer of WCC to bind to C-box element of *frq* promoter. However, in the light, the ZnF domain of only WC2 is required for binding to pLRE of *frq* promoter[9,10]. The different conformations of WCC quaternary structure in light versus dark suggest that there is cross-talk between the light sensing LOV domain of WC1 and ZnF domains of WC1 and WC2. This flexibility in structure and function of the WCC in response to light is not well understood at the molecular level. This makes the WCC an interesting complex to study. Interestingly, the dual role of WC1 is unique to the fungal clock as other animals such as *Drosophila melanogaster* have separate photosensors and transcription factors.

In the following section, I describe efforts to study how the WCC recognizes different regions of *frq* promoter region in the absence of WC1 LOV domain. In particular, I compare the binding affinity of WC1 ZnF and WC2 ZnF domains for C-box and pLRE promoter regions of *frq*. Through this work, I show that both WC1 and WC2 ZnF domains are required to tightly bind to C-box promoter region of *frq*. Absence of either of the ZnF domain results in weak or no binding. On the other hand, only WC2 with ZnF domain (WC2/ZnF) is important to bind pLRE tightly. Moreover, if both WC1 and WC2 ZnF domains are present, the affinity of the complex to pLRE is lowered.

These results have important implications for understanding how WCC ZnF domains recognize and bind to different promoter regions of *frq* in the absence of WC1 light sensing LOV domain.

Results

Studies suggest that only WC2 ZnF is required to bind to the *frq* promoter element, pLRE in the light. However, both WC1 and WC2 ZnF domains are required to bind *frq* promoter element C-box in the dark[9]. This variability in ZnF binding and *frq* promoter element in the dark and light led us to investigate the molecular basis of WCC binding to *frq* promoter. To study the structure of WCC binding to DNA, we aim to use X-ray crystallography. Before attempting to crystallize the WCC with DNA, we studied the binding of WCC with one or both ZnF domains to *frq* promoter regions pLRE and C-box using electrophoretic mobility shift assays (EMSAs) (Figure 3).

Figure 3: DNA sequences of target DNA sequences sub-1 ELRE and frq promoter regions C-box and pLRE. The DNA sequences used in EMSAs or SAXS are in green boxes.

Recognition of C-box and pLRE by WC1/Zn+WC2/Zn

In previous studies, both WC1 and WC2 ZnF domains were required to activate the transcription of *frequency* through the C-box[9]. To study the specific binding of WC1 and WC2 Zn finger domains, I used *frq* promoter element C-box with WC1/Zn+WC2/Zn complex in DNA binding studies. Fixed amount of C-box DNA was incubated with increasing concentration of WC1/Zn+WC2/Zn. The samples were then run on an acrylamide native gel. As the WCC concentration increased in the sample, the mobility of DNA decreased. This suggests that the WC1/Zn+WC2/Zn complex is binding to C-box tightly (Figure 4). Next, we studied the affinity of WC1/Zn+WC2/Zn complex for pLRE *frq* promoter region. Fixed amount of pLRE DNA was incubated with increasing concentration of WC1/Zn+WC2/Zn for an hour. The samples were then run on an acrylamide native gel. As the WCC concentration increased in the sample, the change in mobility of DNA was very little. This result suggests that the WC1/Zn+WC2/Zn complex is binding to pLRE with very low affinity. Since, previous studies suggest that only WC2 ZnF domain is required to bind to pLRE, in this in vitro setting the ZnF domains may be interfering with each other (Figure 4).

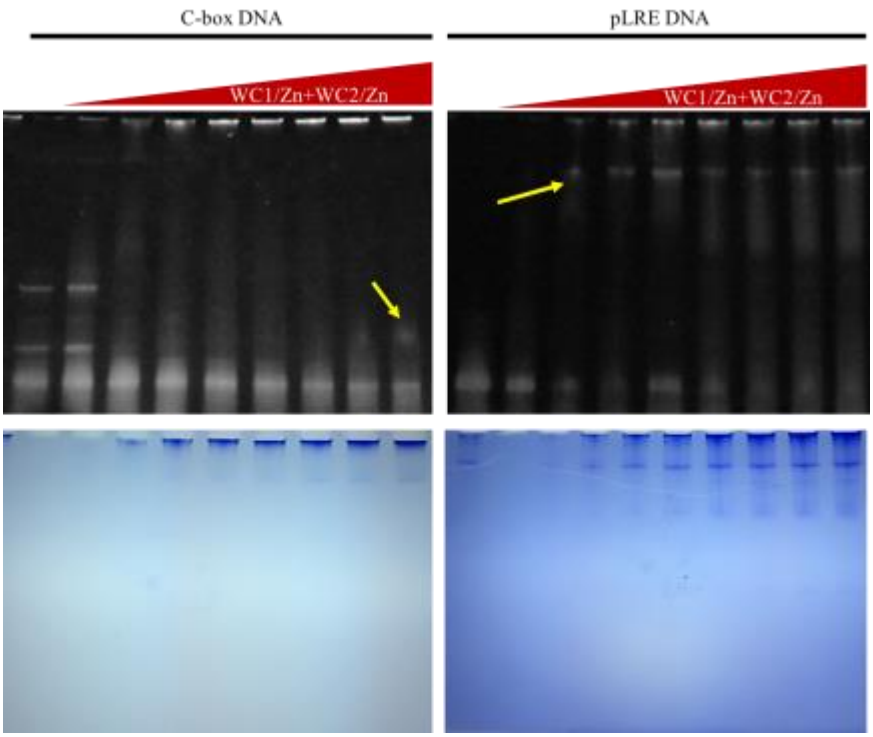

Figure 4: Gel of WC1/Zn+WC2/Zn with C-box and pLRE stained by Ethidium Bromide (top) and Coomassie Blue (bottom). Yellow arrows point to shift in DNA mobility. Concentration of DNA-100uM and protein-1.7 mg/mL.

As controls, I used WC1+WC2 complex without ZnF domains. In this case, I would not expect the WC1+WC2 to bind either pLRE or C-box *frq* promoter elements. This is because WC1 and WC2 do not have their respective ZnF domains. When WC1+WC2 was incubated with fixed amount of either C-box or pLRE, no change in mobility of the DNA was observed. This suggests that ZnF is indeed required for binding of the WCC to *frq* promoter elements.

Recognition of C-box and pLRE by WC1/Zn+WC2

To study the function of either WC1 ZnF or WC2 ZnF domain alone, I used variants of the WC complex that had only one Zn finger domain-WC1/Zn+WC2 and WC1+WC2/Zn. Thus, in the WC1/Zn+WC2 complex, I used WC1/Zn (705-1000, PAS C domain with Zn finger domain) and truncated form of WC2 (174-445, PAS A and PAS B domains). On the other hand, for the

WC1+WC2/Zn complex, I used a truncated form of WC1 (705-918, only PAS C domain) and WC2/Zn (174-500, PAS A, PAS B and Zn finger domains)

To study how WC1 Zn finger domain recognizes *frq* promoter regions, I incubated WC complex variant WC1/Zn+WC2 with C-box for an hour. Similar to the previous experiments, increasing concentration of WC1/Zn+WC2 was incubated with fixed amount of C-box DNA. The samples were then run on an acrylamide native gel. As the WCC concentration increased in the sample, the mobility of DNA decreased again. However, the C-box DNA mobility doesn't decrease significantly. This suggests that the WC1/Zn+WC2 complex is binding to C-box with less affinity compared to WC1/Zn+WC2/Zn complex. This is consistent with previous studies that suggest that both WC1 and WC2 ZnF domains are required to bind to C-box. Thus, WC1 ZnF domain alone is not enough to bind C-box DNA tightly.

To study the affinity of WC1/Zn for pLRE *frq* promoter region, I incubated WC1/Zn+WC2 with pLRE for an hour. The samples were then run on an acrylamide native gel. As the WCC concentration increased in the sample, the mobility of DNA decreased. This suggests that the WC1/Zn+WC2 complex is binding to pLRE with low affinity. This is consistent with previous studies that suggest that only WC2 ZnF domain is required to bind to pLRE. Since WC2 ZnF domain is missing, WC1/Zn+WC2 complex is weakly binding to pLRE (Figure 5).

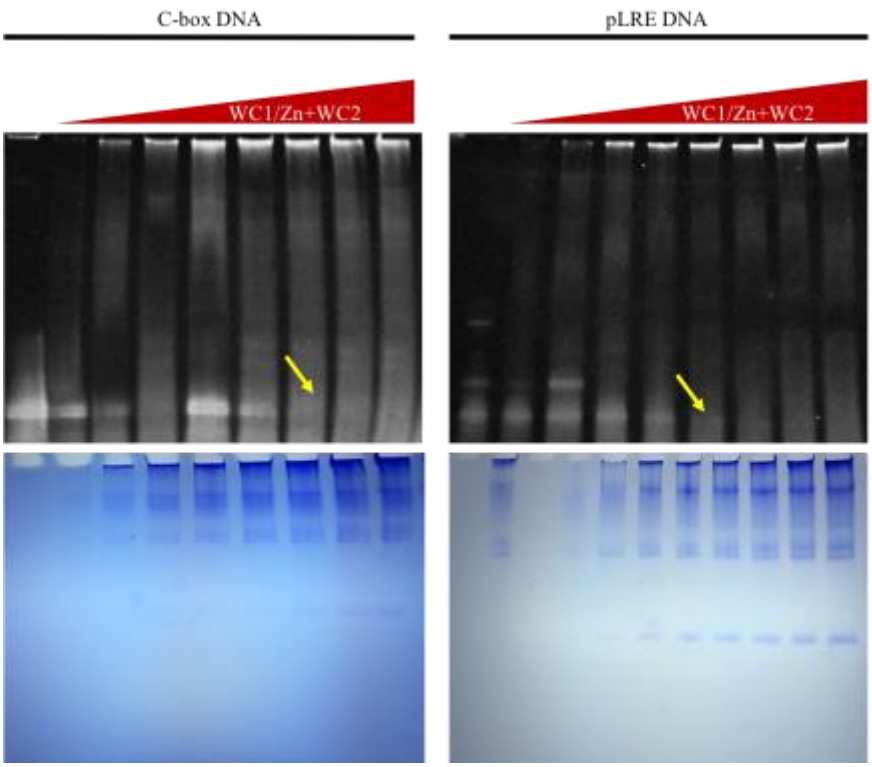

Figure 5: Gel of WC1/Zn+WC2 with C-box and pLRE stained by Ethidium Bromide (top) and Coomassie Blue (bottom). Yellow arrows point to decrease in free DNA. Concentration of DNA-100uM and protein-1.7 mg/mL.

Recognition of C-box and pLRE by WC1 +WC2/Zn

To study how WC2 Zn finger domain recognizes *frq* promoter regions, I incubated WC1+WC2/Zn with C-box. Similar to the previous experiment, increasing concentration of WC1+WC2/Zn was incubated with fixed amount of C-box DNA. The samples were then run on an acrylamide native gel. As the WCC concentration increased in the sample, the mobility of DNA decreased again. However, the mobility of the C-box DNA did not significantly decrease. This suggests that the WC1+WC2/Zn complex is binding to C-box with less affinity. This is consistent with previous studies that suggest that both WC1 and WC2 ZnF domains are required to bind to C-box. Thus, WC2 ZnF domain is not enough to bind C-box DNA tightly.

I incubated WC complex variant WC1 +WC2/Zn with pLRE for an hour. The samples were then run on an acrylamide native gel. As the WCC concentration increased in the sample, the mobility of DNA decreased significantly. This suggests that the WC1+WC2/Zn complex is binding to pLRE. This is consistent with previous studies that suggest that only WC2 ZnF domain is required to bind to pLRE. Since the LOV domain of WC1 is missing in the purified protein, we can assume that the complex is in "dark" state. This suggests that the WC2 ZnF domain can bind to a light regulatory element pLRE even in the absence of light. Thus, perhaps the presence of LOV domain is not required for the WCC to bind to pLRE, but that the LOV domain rather acts to coordinate transcriptional responses, once WC1 has been recruited to DNA (Figure 6).

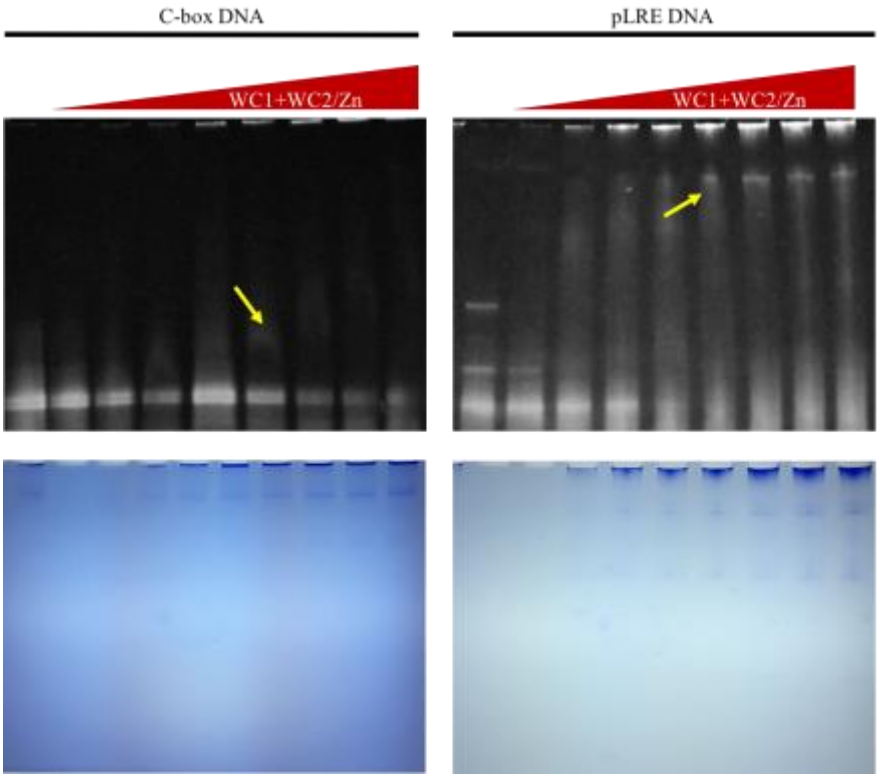

Figure 6: Gel of WC1 +WC2/Zn with C-box and pLRE stained by Ethidium Bromide (top) and Coomassie Blue (bottom). Yellow arrows point to shift in DNA mobility. Concentration of DNA-100uM and protein-1.7 mg/mL.

Recognition of mutated C-box by WC1/Zn +WC2/Zn

To study whether the number of bases between the GATG sequences are important for tight binding of both WC1 and WC2 Zn fingers, I mutated the C-box sequence such that the number of nucleotides is five (Figure 7). This is similar to the number of sequences between the GATC sequence in pLRE. Since both WC1 and WC2 ZnF domains were required to bind to C-box, reducing the number of nucleotides between GATG sequences may result in obstruction between the two sequences. However, this is not observed in the DNA binding assay. To study the specific binding of WC1 and WC2 Zn finger domains, I used mutated *frq* promoter element C-box with WC1/Zn+WC2/Zn complex in DNA binding studies. Fixed amount of mutated C-box DNA was incubated with increasing concentration of WC1/Zn+WC2/Zn. The samples were then run on an acrylamide native gel. As the WCC concentration increased in the sample, the mobility of DNA decreased. This suggests that the WC1/Zn+WC2/Zn complex is binding to mutated C-box tightly similar to that observed with normal C-box. Thus, the number of nucleotides between the GATG sequence does not dictate the WC1 and WC2 Zn finger binding to C-box (Figure 8, 1st panel from the left).

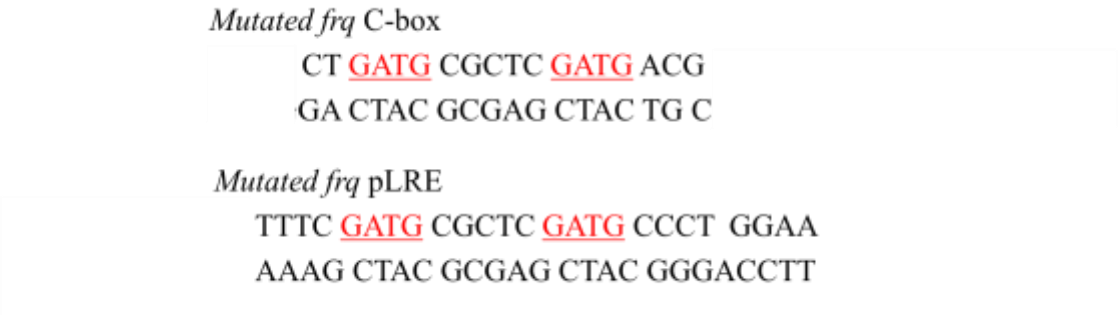

Figure 7: DNA sequences of mutated C-box and mutated pLRE.

Recognition of mutated pLRE by WC1/Zn+WC2 and WC1+WC2/Zn

In previous study, I showed that only WC2 Zn finger is required to tightly bind to pLRE. To study whether WC2 Zn finger has specificity for GATC sequence in pLRE, I mutated the GATG sequence in pLRE to GATC sequence. To study if WC2 Zn finger can still bind to this mutated pLRE, I used this mutated pLRE with WC1+WC2/Zn complex in a DNA binding assay. Fixed amount of mutated pLRE DNA was incubated with increasing concentration of WC1 +WC2/Zn. The samples were then run on an acrylamide native gel. As the WCC concentration increased in the sample, there was no change in DNA mobility or amount of free DNA. This suggests that the WC1+WC2/Zn complex is not binding to mutated pLRE. Thus, WC2 Zn finger is specific to the GATG sequence of pLRE. However, when I used WC1/Zn+WC2 to study the binding of WC1 Zn finger to mutated pLRE, mobility of the mutated pLRE decreased. This suggests that the WC1/Zn+WC2 complex was binding to the mutated pLRE. These results suggest that because WC2 Zn finger has a specificity for GATG sequence in pLRE, it is required and necessary for tight binding to pLRE. (Figure 8, 2nd and 3rd panels from the left).

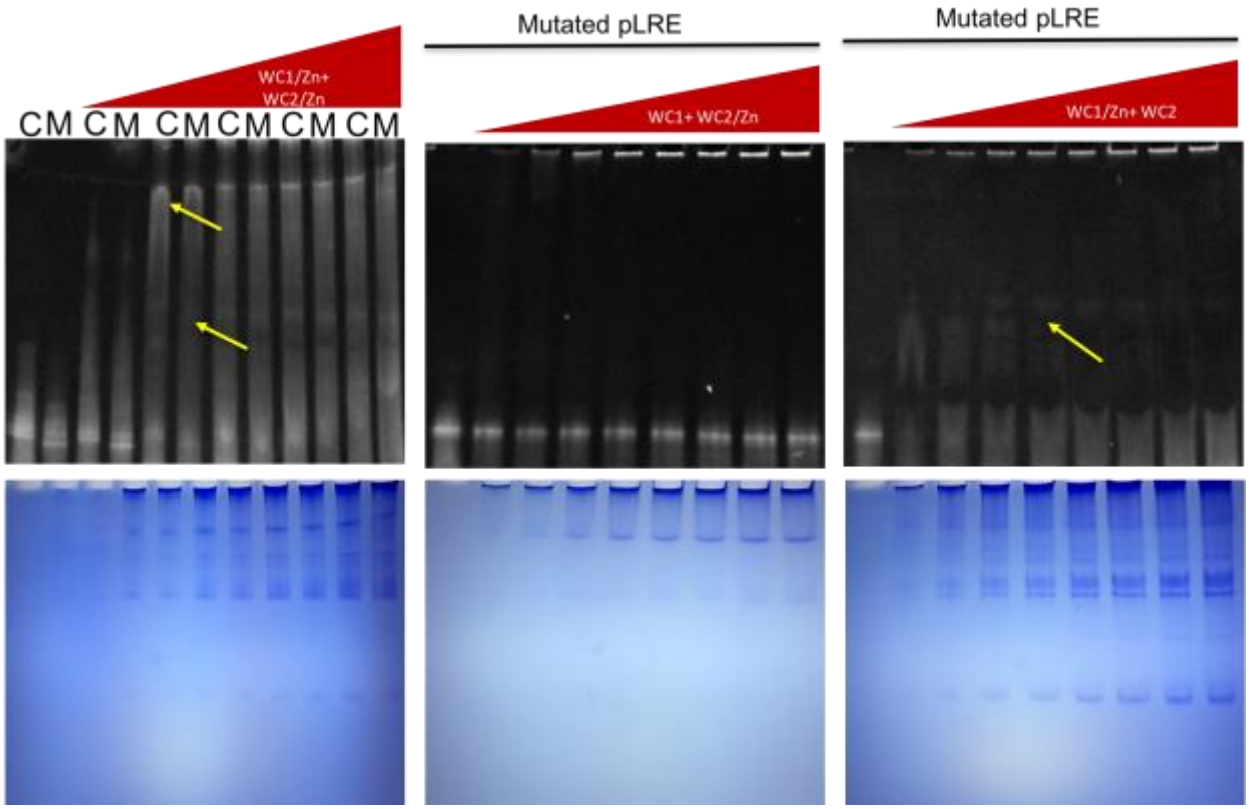

Figure 8: Gels of WC1/Zn +WC2/Zn with mutated C-box, WC1+WC2/Zn and WC1/Zn+WC2 with mutated pLRE (middle and right) stained by Ethidium Bromide (top) and Coomassie Blue (bottom). Yellow arrows point to shift in DNA mobility. Concentration of DNA-100uM and protein-1.7 mg/mL

SAXS of the WC complexes with and without target DNA

To study if there is any change in the overall shape of WCC when both ZnF domains are present with or without target DNA, SEC-SAXS was applied to study the WCC shape. In these experiments, I used the *sub-1* early light regulatory element (ELRE) DNA sequence (Figure 3). In the presence of WCC, the mobility of DNA decreases suggesting WCC-DNA complex is formed. While the WCC does form a complex with *sub-1* ELRE, similar to SAXS results of WC1+WC2/Zn, the overall shape of WC1/Zn+WC2/Zn does not change significantly. However, the shape of the Kratky curve for WC1/Zn + WC2/Zn with no DNA shifts slightly to the right

suggesting there is a change in the flexibility of the complex. This suggests that the WCC is less flexible when it binds to DNA (**Error! Reference source not found.**).

To study if there is any change in the overall shape of WC1+WC2/Zn in the presence or absence of target DNA, I used SEC-SAXS to study the overall shape and flexibility properties of the WCC. In these experiments, I used *sub-1* ELRE DNA sequence. In the presence of WCC, the mobility of DNA decreases suggesting WCC-DNA complex is formed. Whereas the WCC does form a complex with *sub-1* ELRE, the overall shape of WCC does not change significantly. However, the shape of the Kratky curve for WCC+ no DNA slightly shifts to the right suggesting there is a change in the flexibility of the complex (Figure 9 and Table 1).

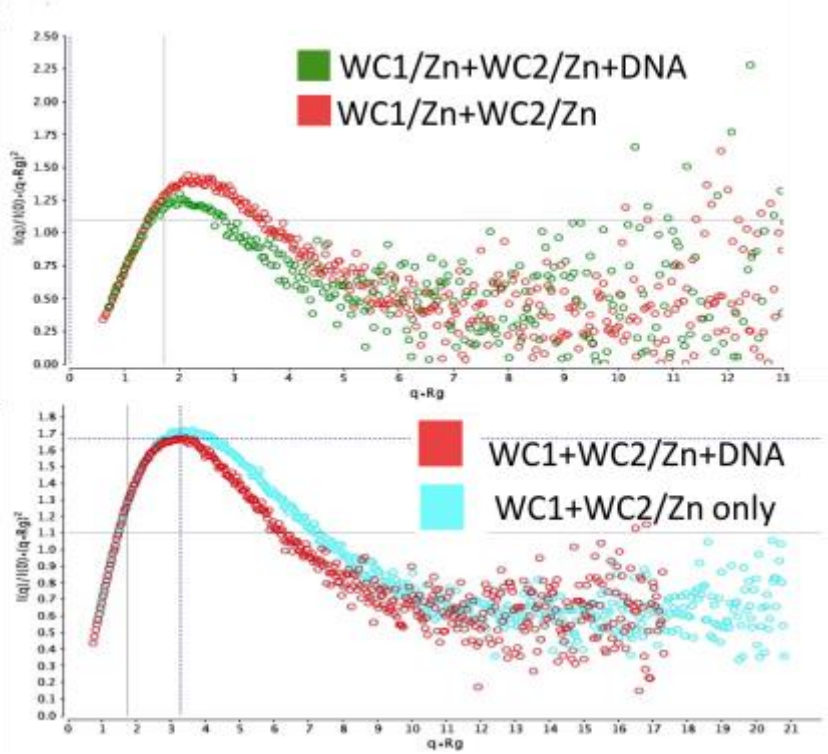

Figure 9: Kratky plots of WC1/Zn+WC2/Zn (top) and WC1+WC2/Zn (bottom) with and without sub-1 ELRE DNA.

Sample	Porod Volume ($Å^3$)	Radius of Gyration(Å)	Molecular Weight (kDa)
WC1/Zn+WC2/Zn +DNA	**430111.4**	**65.3**	**357**
WC1/Zn+WC2/Zn	**247378.3**	**55.9**	**205.3**
WC1+WC2/Zn +DNA	**341561.9**	**83.5**	**283.5**
WC1+WC2/Zn	**547300.2**	**85.3**	**298.9**

Table 1: SAXS data tables of WC1/Zn+WC2/Zn (top) and WC1+WC2/Zn (bottom) with and without sub-1 ELRE DNA.

Discussion

Transcription factors play a key role in the circadian clock by activating clock-controlled genes, which include the key repressor proteins for the negative feedback loop. In most systems the transcription factors are distinct from the photoreceptors. However, *Neurospora crassa* clock consists of a photoreceptor WC1 that also acts as a transcription factor. This property of WC1 makes it a central protein in the *Neurospora* circadian clock. In the presence of light, WC1 detects the light signal and induces transcription of certain genes that result in a response by *Neurospora crassa*[4]. In the dark, WC1 acts as a general transcription factor and induces the expression of clock-controlled genes. WC1 interacts with WC2 to form a heterodimer to perform its transcription factor activity.

In this study, I have studied the DNA binding of both WC1 and WC2 in the absence of WC1 photoreceptor ability. Thus, the LOV domain of WC1 is absent in the truncated WC1 used in the assays. The LOV domain of WC1 is important for its photoreceptor activity. The LOV domain binds to FAD and may result in oligomerization of the WCC. This multimeric complex acts as an activator for transcription in light. Previous studies suggest that in the presence of light, WC1 ZnF

domains are not required to bind to pLRE. Only WC2 ZnF domain is required to bind pLRE. In this study, I show that in the absence of WC2 ZnF domain, WC1 ZnF can also bind to pLRE but with very little affinity. Additionally, in the absence of WC1 ZnF and presence of WC2 ZnF, there is very tight binding of the complex to pLRE. This is consistent with the previous studies suggesting that only WC2 ZnF is required to bind to pLRE[9].

Interestingly, in the presence of both WC1 and WC2 ZnF and absence of WC1 LOV domain, the binding of the complex to pLRE is very low. This suggests that the LOV domain plays a role in changing the conformation of WC1 such that there is no steric hindrance when the WCC binds to pLRE such that WC2 ZnF can easily bind to pLRE. In this study, I also show that LOV domain is not important for WCC binding to C-box. This is because, in the presence of both WC1 and WC2 ZnF domains, the binding of WCC to C-box is very strong. A crystal structure of WCC without a LOV domain in the presence of pLRE or C-box would provide a better understanding of how WCC interacts with DNA in the absence of LOV domain. Study of mutated pLRE and C-box DNA sequences in the presence of WCC variants would also provide information about how the target DNA sequence affects WCC binding.

Methods

Purification of WCCs

For the purification of *Neuropsora* WCCs, four constructs of the proteins were used: WC1 (705-918, PAS C domain), WC2 (174-445, PAS A and PAS B domains), WC1/Zn (705-1000, PAS C domain with Zn Finger Domain) and WC2/Zn (139-500, PAS A, PAS B and Zn finger domains). The WC1 construct was designed with only PAS C domain because this domain is necessary for interacting with WC2 to form the WCC heterodimer (reference). All the constructs are cloned in pET28a with a poly-His tag and kanamycin resistance marker. The constructs were used for BL21

competent *E. coli* transformations and then plated onto agar media containing kanamycin antibiotic. A colony was selected to grow 8 L of bacterial cell culture at 37° C in the presence of kanamycin. The cell culture was induced at 17° C using 100 μM IPTG at OD_{600} 0.6 – 0.8 and grown overnight. The cells were harvested the following day. Four complexes were purified: WC1+WC2, WC1/Zn+WC2, WC1+WC2/Zn, and WC1/Zn+WC2/Zn. Approximately 12 g of each WC1 and WC2 pellet were used to purify the complex by Ni^{2+} affinity chromatography using the His-tag. Before doing the Ni^{2+} affinity chromatography, the WCC lysate was incubated with 10 uL of benzonase nuclease (Santa Cruz) at 4°C on a rocker. After Ni^{2+} affinity chromatography, the WCC was further purified using size-exclusion chromatography (SEC). After concentrating down the appropriate fractions, the complexes were frozen and stored at -80°C.

EMSA of WCC variants

Concentration of purified WCC was measured. For all assays, the concentration of WC1/Zn+WC2/Zn, WC1+WC2, WC1/Zn+WC2 WC, WC1 +WC2/Zn was kept the same-1.7 mg/mL. Single stranded DNA (ssDNA) oligos for pLRE and C-box (25 nmol) were bought from Integrated DNA Technologies (IDT). The ssDNA oligos were annealed at 94°C for 2 minutes and subsequently kept at room temperature for 45 minutes to let the oligos cool down. 3 uL of 100 uM DNA (pLRE or C-box) was incubated with increasing amount of WC complex. The WC complex volume in each tube was increased by 1 uL increments starting with1 uL and ending with 8 uL of WC complex. 1 uL of 10 uM zincacetate was added to the tube. The samples were incubated on ice for 1 hour. After 1 hour, appropriate amount of filtered nanopure water was added to all tubes to make the final volume 12 uL for each tube. 8 uL of the native loading buffer was added. The samples were run on a native gel for 25-30 minutes at 200 V using Tris-Glycine buffer. The gel was transferred to Ethidium Bromide staining buffer that contained Tris-Glycine buffer+10uL of

Ethidium Bromide dye (1 ug/mL). The gel was stained with Ethidium Bromide buffer for 15-20 minutes and imaged. After imaging the gel, the gel was stained with Coomassie Blue stain.

SAXS of WCC variants

Variants of WCC (5 mg/ml) with and without equal concentration of *sub-1* ELRE DNA were analyzed by BioSepTM 5 μm SEC-s3000 LC Column for SEC-MALS with buffer containing 150 mM NaCl, 50 mM Tris pH 7.5, 5 mM DTT and 10% glycerol. "q" is calculated as $q = 4\pi\sin(\theta)/\lambda$, where θ is half of the angle between the incident X-ray beam and the scattered beam, and λ is the wavelength of X-rays.

REFERENCES

1. Harmer, S. L., Panda, S. & Kay, S. A. Molecular Bases of Circadian Rhythms. *Annu. Rev. Cell Dev. Biol.* **17,** 215–53 (2001).
2. Baker, C. L., Loros, J. J. & Dunlap, J. C. The circadian clock of Neurospora crassa. *FEMS Microbiol. Rev.* **36,** 95–110 (2012).
3. Froehlich, A. C., Liu, Y., Loros, J. J. & Dunlap, J. C. White Collar – 1 , a Circadian Blue Light Photoreceptor , Binding to the frequency Promoter. **297,** 815–19 (2002).
4. Ballario, P., Talora, C., Galli, D., Linden, H. & Macino, G. Roles in dimerization and blue light photoresponse of the PAS and LOV domains of Neurospora crassa white collar proteins. *Mol. Microbiol.* **29,** 719–729 (1998).
5. Froehlich, A. C., Liu, Y., Loros, J. J. & Dunlap, J. C. White Collar – 1 , a Circadian Blue Light Photoreceptor , Binding to the frequency Promoter. *Science (80-.).* **297,** 815–819 (2002).
6. Hurley, J. M., Larrondo, L. F., Loros, J. J. & Dunlap, J. C. Conserved RNA helicase FRH acts nonenzymatically to support the intrinsically disordered neurospora clock protein FRQ. *Mol. Cell* **52,** 832–843 (2013).
7. Anderson, P. W. *et al.* White Collar – 1 , a Circadian Blue Light Photoreceptor , Binding to the frequency Promoter. **1605,** (2001).
8. Cheng, P., Yang, Y., Wang, L., He, Q. & Liu, Y. WHITE COLLAR-1, a multifunctional Neurospora protein involved in the circadian feedback loops, light sensing, and transcription repression of wc-2. *J. Biol. Chem.* **278,** 3801–3808 (2003).

9. Wang, B., Zhou, X. Y., Loros, J. J. & Dunlap, J. C. Alternative Use of DNA Binding Domains by the Neurospora White Collar Complex Dictates Circadian Regulation and Light Responses. *Mol. Cell. Biol.* **36,** 781–793 (2016).

10. Smith, K. M. *et al.* Transcription factors in light and circadian clock signaling networks revealed by genomewide mapping of direct targets for neurospora white collar complex. *Eukaryot. Cell* **9,** 1549–1556 (2010).

CHAPTER 3

CONFORMATION OF FREQUENCY CHANGES WHEN IT IS PHOSPHORYLATED

Introduction

Circadian clocks are important for the behavior and physiology of eukaryotic organisms. All eukaryotic clocks or oscillators are similar at the molecular level. They are composed of positive and negative feedback loops. Additionally, similar clock protein regulators such as phosphatases and kinases are present. Among the eukaryotes, fungus *Neurospora crassa* is a widely-used model system to study the circadian clock. The *Neurospora crassa* circadian clock is readily monitored through the diurnal production of asexual spores. Additionally, the fact that the genome of *Neurospora crassa* is highly studied and characterized also that makes the fungus a good model organism for circadian clock studies.

The *Neurospora crassa* circadian clock is composed of a key repressor protein Frequency (FRQ). FRQ is a highly-disordered protein that is stabilized by Frequency-interacting RNA Helicase (FRH)[1]. FRQ bound to FRH makes a stable repressor complex (FFC). This repressor complex completes the negative limb of the circadian clock by inactivating the White Collar Complex (WCC). The WCC is composed of activators White Collar 1 (WC1) and White Collar 2 (WC2). These activators act as transcription factors for clock controlled genes such as *frequency* (*frq*). The WCC binds to the Clock box (C-box) region of *frq* promoter in the dark and proximal Light Regulatory Element (pLRE) region of the *frq* promoter in the light. After being

synthesized, FRQ, interacts with FRH and subsequently interacts with Casein Kinase 1 (CK1). This large complex represses the transcriptional activity of the WCC through phosphorylation of the complex.

FRQ is a highly-disordered protein that contains many structural domains. It is composed of a coiled-coil (CC) domain that has been proposed to generate FRQ homodimers[2]. FRQ also contains CK1 and FRH binding domains. Additionally, the repressor FRQ also contains proline, glutamic acid, serine and threonine rich sequences called the PEST-1 and PEST-2 domains. These domains are highly phosphorylated by casein kinases and result in degradation of FRQ[3] (Figure 1). Interestingly, FRQ has two mRNA spliced versions, full length FRQ (lFRQ) and 100 amino acids shorter FRQ (sFRQ) at the N- terminus[4]. These two versions exist so that the *N. crassa* clock can compensate for the change in temperature. Thus, an increase in temperature results in an increase in levels of lFRQ[5].

While change in temperature results in different versions of FRQ, the structure of FRQ is directly affected by phosphorylation. During a circadian cycle, FRQ is progressively phosphorylated by kinases like Casein Kinase 1a. Casein Kinases phosphorylate the middle of FRQ immediately after its synthesis. This action results in hypophosphorylated form of FRQ. As time passes, the PEST sequences are phosphorylated by the end of the day[3]. This results in a hyperphosphorylated form of FRQ. Finally, the N-terminus of lFRQ is phosphorylated at the end of the circadian cycle[6]. This phosphorylation of FRQ changes its conformation and results in its degradation (Figure 1). However, there is also evidence that the ability of FRQ to repress the WC complex also depends upon its phosphorylation state[7]. Hypophosphorylated

FRQ is in a closed conformation and thus it takes longer for trypsin to cleave it. However, hyperphosphorylated FRQ is in a more open conformation due to electrostatic repulsion between phosphorylated residues. This form of FRQ is quickly cleaved by trypsin[8]. Thus, some structural features of FRQ are known. However, the structure of FRQ is largely unknown even though the function FRQ as a repressor is widely studied.

Results

FREQUENCY stabilized by CASEIN KINASE 1 phosphorylation

FRQ is a highly disordered protein. Immediately after it is synthesized, FRH, an RNA helicase binds to FRQ and stabilizes it[9]. This complex of FRQ-FRH acts as a repressor complex when CK1 binds to FRQ. Since FRQ is such a disordered protein, purifying FRQ on its own is very challenging. Truncated FRQ (amino acids 100-540) purified alone aggregated. However, it is unknown if FRQ needs phosphorylation for its stability. Thus, I co-purified FRQ with its binding partner and partner kinase CK1. When truncated FRQ (with His-tag) was co-purified with CK1 (without His-tag), a stable FRQ-CK1 complex was formed. Mass spectrometry of this truncated FRQ bound to CK1 confirmed that FRQ is phosphorylated at 76 different sites at amino acids Serine, Threonine and Tyrosine. These results suggest that the truncated FRQ bound to CK1 was hyperphosphorylated. Additionally, the molecular weight of hyperphosphorylated truncated FRQ was higher compared to that of dephosphorylated truncated FRQ when observed on an SDS PAGE. Additionally, truncated FRQ:CK1 gets degraded when dephosphorylated. This suggests that phosphorylation of this complex is important for

its stability. Similar results were obtained when full length FRQ (with His-tag) was co-purified with CK1(with His-tag). However, no degradation was observed (Figure 10).

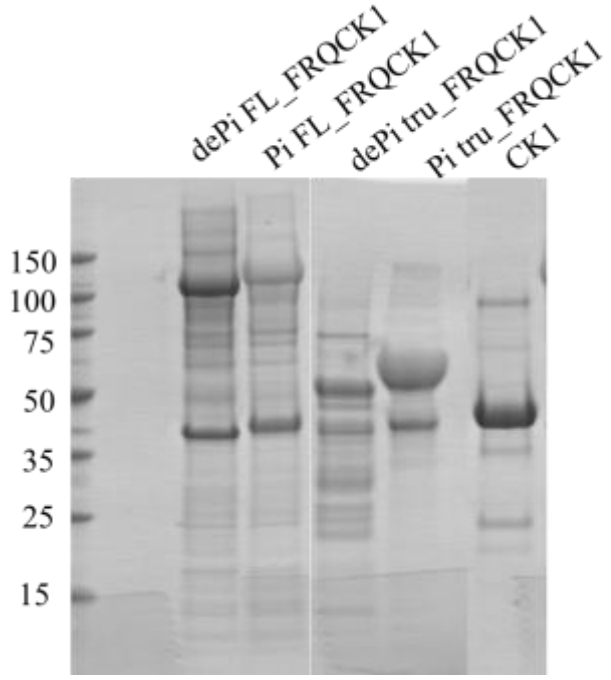

Figure 10: SDS-PAGE of phosphorylated and dephosphorylated full length FRQ bound to CK1 with molecular weights 145 kDA and 110 kDa respectively. Truncated phosphorylated (65 kDa) and dephosphorylated FRQ:CK1 (51 kDa) are also present.

During purification, CK1 (without His-tag) co-elutes with truncated FRQ after CK1 is co-expressed with truncated FRQ in *E. coli* cells (Figure 11). This suggests that CK1 is bound to FRQ during purification. Since the FRQ:CK1 complex that is purified is phosphorylated by CK1 as shown above, CK1 can stabilize FRQ by both binding to it and phosphorylating it.

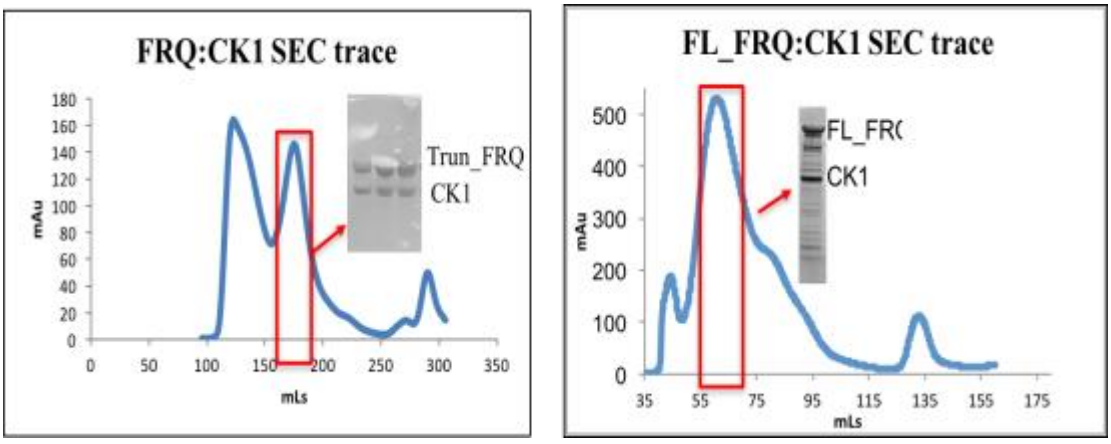

Figure 11: SEC traces and corresponding SDS-PAGE gel of truncated FRQ:CK1(left panel) and full length FRQ:CK1 (right panel).

To study phosphorylation effects on the stability of FRQ, I co-purified truncated FRQ with CK1 kinase null mutants (D131N and K41R)[10]. In both cases, CK1 mutants eluted in the wash buffer and thus did not form stable complexes with FRQ (Figure 12). This resulted in aggregation of FRQ in the size exclusion column. This suggests that FRQ is unstable when it is not bound by CK1. Additionally, CK1 needs an active kinase domain to interact with FRQ and stabilize it.

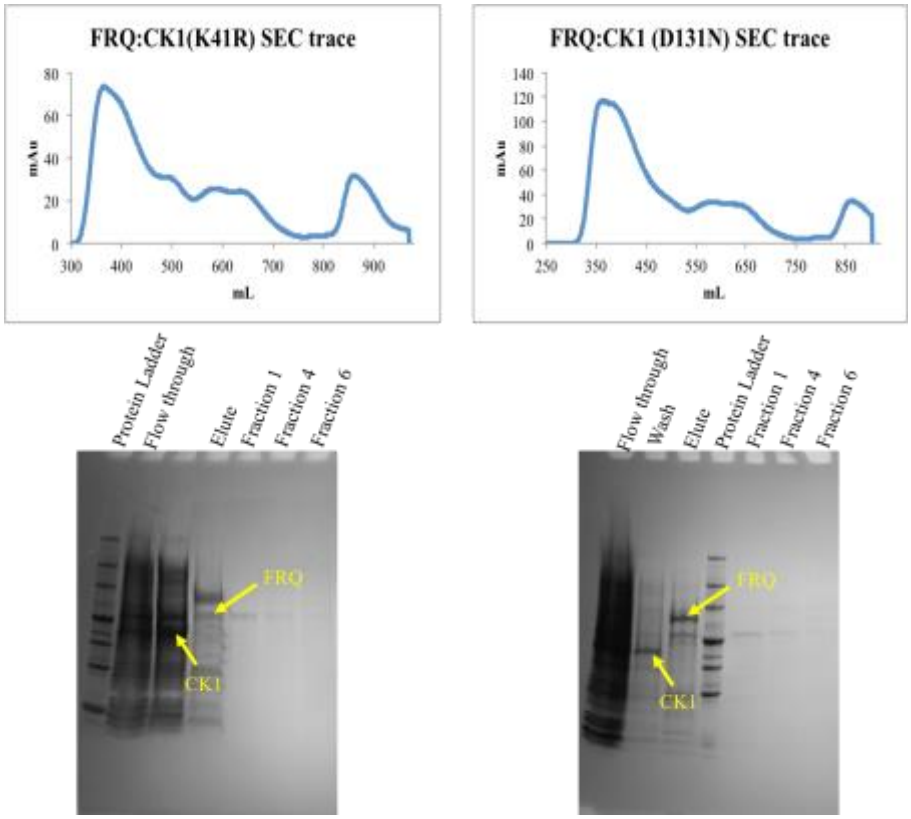

Figure 12: : SEC traces and corresponding SDS-PAGE gels of truncated FRQ purified with CK1 mutants K41R and D131N.

SAXS of phosphorylated and dephosphorylated FRQ:CK1

Studies suggest that hyperphosphorylated FRQ is an open conformation. This was confirmed by a trypsin digest of phosphorylated and dephosphorylated truncated FRQ:CK1 complexes. The trypsin digests showed that phosphorylated FRQ:CK1 complex is in a more open conformation compared to a closed conformation of dephosphorylated FRQ:CK1 complex. These results were consistent with previous published trypsin digests[8]. To learn more about the overall change in conformation of FRQ:CK1 complex, I used small angle X-ray scattering (SAXS). SAXS of phosphorylated and dephosphorylated full length FRQ:CK1 was performed to determine shape of the complex. The Kratky plots showed that the peak of phosphorylated FRQ:CK1 complex is shifted to the right compared to that of

dephosphorylated FRQ:CK1 peak. This suggests that phosphorylated FRQ:CK1 is more flexible compared to dephosphorylated FRQ:CK1. These results are consistent with the trypsin digest results suggesting that phosphorylation of FRQ causes it to form a more open and flexible conformation (Figure 13 and Table 2).

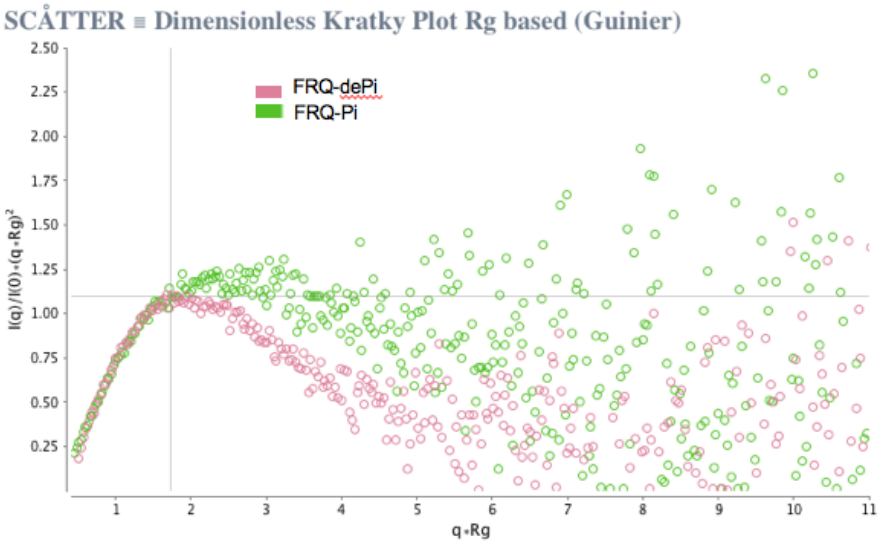

Figure 13: : The Kratky plots of phosphorylated FRQ (FRQ-Pi) and dephosphorylated FRQ (FRQ-dePi).

Sample	Porod Volume (Å^3)	Radius of Gyration (Å)	Molecular Weight (kDa)
Dephosphorylated FRQ:CK1	**470137.2**	**63.6**	**390.2**
Phosphorylated FRQ:CK1	**201081.3**	**59.5**	**166.9**

Table 2: SAXS data of phosphorylated and dephosphorylated full length FRQ:CK1

Discussion

FRQ is a highly disordered protein and wheras it is a major repressor in the *N. crassa* circadian clock, very little is known about the structure of the protein. While studies show that FRH binds to FRQ and stabilize it, there is little known about CK1 binding. In this study, I show that CK1 can also stabilize FRQ by phosphorylating it.

Additionally, phosphorylation of CK1 results in stabilizing the protein. In the absence of CK1 phosphorylation and binding, FRQ is unstable. This suggests that hyperphosphorylated form of FRQ is more stable in the presence of CK1.

Studies have also shown that hyperphosphorylated FRQ has a more open conformation because of electrostatic repulsion present due to negative charged phosphate groups. This was confirmed by the SAXS studies. Phosphorylated FRQ bound to CK1 had a more flexible conformation while dephosphorylated FRQ:CK1 was in less flexible/globular conformation. This is consistent with the studies that suggest that when FRQ is hyperphosphorylated, it is an open conformation so that ligases can recognize specific sequences for FRQ degradation. A crystal structure of FRQ:CK1 alone or in complex with FRH would shed light on how FRQ is stabilized by CK1 or FRH binding.

Methods

FRQ:CK1 Purification

In order to maintain protein stability, truncated form of *Neuropsora* FRQ (CK1 binding domain and coiled coil region) is co-purified with the CK1. The FRQ construct (100 – 540) is cloned in pET28a with a poly-His tag and kanamycin resistance marker. For purification of full length CK1, two constructs were used. The first construct was in pCDFDuet-1 plasmid without a poly-his tag and streptomycin resistance marker. This construct was used for purifying full length FRQ and CK1 together. The other construct was in pET28a plasmid with a poly-His tag and kanamycin resistance marker. This construct was used for purification of CK1 only.

To purify FRQ and CK1 together, FRQ with poly-His tag and CK1 without His tag were

co-expressed in BL21 cells and then plated onto agar media containing kanamycin and streptomycin antibiotic. A colony was selected to grow 8 L cell culture 37° C in the presence of kanamycin and streptomycin. The cell culture was induced overnight at room temperature at OD_{600} 0.6. After harvesting the cells the following day, the complex was purified by using Ni^{2+} affinity chromatography using the His-tag of FRQ. The complex was further purified by using SEC. Mass spectrometry analysis confirmed that the complex FRQ and CK1 isolated from SEC was pure.

To purify CK1 by itself, a similar procedure was taken as mentioned above. CK1 with poly-His tag was expressed in BL21 cells and plated onto agar media containing kanamycin antibiotic. A colony was selected to grow 8 L cell culture 37° C in the presence of kanamycin. The cell culture was induced overnight at room temperature at OD_{600} 0.6. After harvesting the cells, the complex was purified by using Ni^{2+} affinity chromatography and SEC.

Dephosphosphorylation of FRQ:CK1

FRQ:CK1 complex was dephosphorylated by using 1 μL of λ Phosphatase (NEB), 5μL of λ Phosphatase Buffer (10X), FRQ:CK1 and water to make up a 50 μL of reaction mixture. This reaction mixture was incubated at 4°C for 30 minutes.

FRQ:CK1 SAXS

Phosphorylated and dephosphorylated full length FRQ:CK1 (10 mg/ml) were analyzed by BioSepTM HEPES pH 7.5, 5 mM DTT. "q" is calculated as $q = 4\pi\sin(\theta)/\lambda$, where θ is half of the angle between the incident X-ray beam and the scattered beam, and λ is the wavelength of X-rays.

REFERENCES

1. Hurley, J. M., Larrondo, L. F., Loros, J. J. & Dunlap, J. C. Conserved RNA helicase FRH acts nonenzymatically to support the intrinsically disordered neurospora clock protein FRQ. *Mol. Cell* (2013). doi:10.1016/j.molcel.2013.11.005
2. Cheng, P., Yang, Y., Heintzen, C. & Liu, Y. Coiled-coil domain-mediated FRQ-FRQ interaction is essential for its circadian clock function in Neurospora. *EMBO J.* **20,** 101–108 (2001).
3. Görl, M. *et al.* A PEST-like element in FREQUENCY determines the length of the circadian period in Neurospora crassa. *EMBO J.* **20,** 7074–7084 (2002).
4. Garceau, N. Y., Liu, Y., Loros, J. J. & Dunlap, J. C. Alternative initiation of translation and time-specific phosphorylation yield multiple forms of the essential clock protein FREQUENCY. *Cell* **89,** 469–476 (1997).
5. Colot, Hildur V., Loros, Jennifer J., Dunlap, J. C. Temperature-modulated Alternative Splicing and Promoter Use in the Circadian Clock Gene frequency. *Mol. Biol. Cell* **16,** 5563–5571 (2005).
6. Baker, C. L., Kettenbach, A. N., Loros, J. J., Gerber, S. A. & Dunlap, J. C. Quantitative Proteomics Reveals a Dynamic Interactome and Phase-Specific Phosphorylation in the Neurospora Circadian Clock. *Mol. Cell* **34,** 354–363 (2009).
7. Larrondo, L. F., Olivares-Yañez, C., Baker, C. L., Loros, J. J. & Dunlap, J. C. Circadian rhythms. Decoupling circadian clock protein turnover from circadian period determination. *Science* **347,** 1257277 (2015).

8. Querfurth, C. *et al.* Circadian Conformational Change of the Neurospora Clock Protein FREQUENCY Triggered by Clustered Hyperphosphorylation of a Basic Domain. *Mol. Cell* (2011). doi:10.1016/j.molcel.2011.06.033

9. Cheng, P., He, Q., He, Q., Wang, L. & Liu, Y. Regulation of the Neurospora circadian clock by an RNA helicase. *Genes Dev.* 234–241 (2005). doi:10.1101/gad.1266805.234

10. Lauinger, L., Diernfellner, A., Falk, S. & Brunner, M. The RNA helicase FRH is an ATP-dependent regulator of CK1a in the circadian clock of Neurospora crassa. *Nat. Commun.* **5,** 3598 (2014).

CHAPTER 4

STRUCTURAL PROPERTIES OF PERIOD AND ITS PHOSPHOMIMIC MUTANTS

Introduction

The earth rotates about its axis with a period of twenty-four hours. This causes changes in light and temperature. To adapt to these changes, many eukaryotic organisms have circadian clocks that allow respective changes in their behavior and physiology. The eukaryotic clock also has a period of approximately twenty-four hours. To maintain this period or daily circadian rhythm in eukaryotes, many positive and negative feedback loops are present at the molecular level.

Model organism *Drosophila melanogaster* has been used to study circadian clocks at a molecular level for many decades now. *Drosophila's* genome is well studied and characterized that allows large-scale forward genetic screens. Additionally, the circadian clock manifests in *Drosophila* in a variety of behaviors such as eclosion, egg laying, olfactory sensitivity and most importantly locomotor activity. These behaviors are easily monitored. For example, locomotor activity in flies peaks rhythmically once in the morning and once in the evening.[1] This can be monitored by entrapping a fly in a glass tube with food on one end and crossing an infra-red beam that emits an electronic signal each time the fly cross the beam.[2] *Drosophila* is a good circadian clock model organism because it has a small neuronal network of which only 150 clock cells per brain hemisphere are present[3].

Drosophila is used as a model system to study circadian clocks in eukaryotic organisms because the molecular mechanisms of the *Drosophila* clock are conserved.[4] The *Drosophila* circadian clock consists of positive and negative feedback loops. These loops are composed of positive regulators such as transcription factors Clock (CLK) and Cycle (CYC) and negative regulators such as Period (PER) and Timeless (TIM). CLK and CYC activate transcription of clock –controlled genes such as *per* and *tim*.[5] In the subjective day, *per* and *tim* genes are transcribed such that the respective mRNA levels begin to increase. This leads to an accumulation of PER and TIM in the cytoplasm. TIM binds to PER and stabilizes it. The PER-TIM complex translocates to the nucleus and represses the activity of the transcription factors CLK:CYC. This results in decrease in transcription of *per* and *tim* and closing of the feedback loop.[6]

The PER-TIM repressor complex regulated by kinases such as Shaggy (SGG), Casein Kinase 2 (CK2) and Double-Time (DBT). DBT progressively phosphorylates PER and this results in change in conformation of PER. The F-box component of an SCF ligase SLIMB binds to phosphorylated PER. This results in degradation of PER[7](Figure 2). While functional aspects of PER are known, very little is known about the structure of PER. Additionally, structural properties of PER when it is phosphorylated are also unknown.

Results

Repressor protein PER is a highly-disordered protein. However, the protein does contain two PAS (Per-Arnt-Sim) domains near the N-terminal region.[8] These PAS domains consist of canonical five-strand beta sheet and a cluster of 4-5 alpha helices. These PAS domains are called PAS A and PAS B and are responsible for making a PER

homodimer.[8,9] The PER PAS domains are also important for TIM binding.[10] While the crystal structure of the PAS domains of PER has been solved, not much is known about the remaining domains of PER repressor.

Stability of PER (1-700) and its phosphomimic mutants

To study the PER protein, we purified a truncated form of the protein from 1-700 amino acids. PER (1-700) contains the structurally stable PAS domains in addition to DBT binding domain. This construct is more stable than full length PER protein during purification. This construct is considered WT PER from here on.

Kinase DBT phosphorylates PER at various Serine/Threonine sites. For example, mutation of amino acid S589 or neighboring amino acids can result in a short circadian rhythm.[11] To understand how the conformation of PER is affected by its phosphorylation, variants of PER were purified. Phosphomimic mutants had one or more serines/threonines of the phosphorylation cascade substituted to aspartate. On the other hand, phosphonull mutants had an alanine substituted. To study the stability of the PER variants, WT PER and the variants were digested by trypsin.

When S629 was mutated to aspartate (XXD), the expression level of the protein was medium compared to low expression levels of phosphonull variant, S629A (XXA). Both the mutants aggregated upon purification. This suggests that these mutants are unstable compared to WT PER. Similarly, variants S589D (DXX) and S589A (AXX) had similar results where DXX had medium expression levels while AXX had low expression levels. However, both mutants aggregated upon purification. This suggests that these mutants are also highly unstable compared to WT PER.

Interestingly, when S589 and the amino acid group S604/S607/T610/S613 were mutated to aspartate (DDX), the protein was stable upon purification. Additionally, mutating amino acids S589, S604/S607/T610/S613 and S629 to aspartate (DDD) resulted in low aggregation upon purification. This suggests that DDD is also a stable protein variant. These results suggest that progressive phosphorylation of PER results in a more structurally stable protein (Table 3).

PER mutant	Expression level	Aggregate	Degradation upon purification
XXX	High	Low	None
DDD	High	Low	None
DDX	High	Low	None
XXD	Medium	High	None
XXA	Low	High	Yes
AXX	Low	High	Yes
DXX	Medium	High	Yes
XAX	Low	Low	N/A
XDX	Low	Low	N/A

Table 3: This table depicts PER variants levels of expression, aggregation and whether the variant degrades upon purification.

MALS of PER (1-700) and its phosphomimic mutants

Since WT PER and its variants have different stability, we wanted to study whether the structure is also different by using Multi Angle Light Scattering (MALS). We found that there is no difference in the overall structure of WT PER, DDX and DDD. All proteins formed dimers (~155 kDa) where a monomer is 75 kDa (Figure 14).

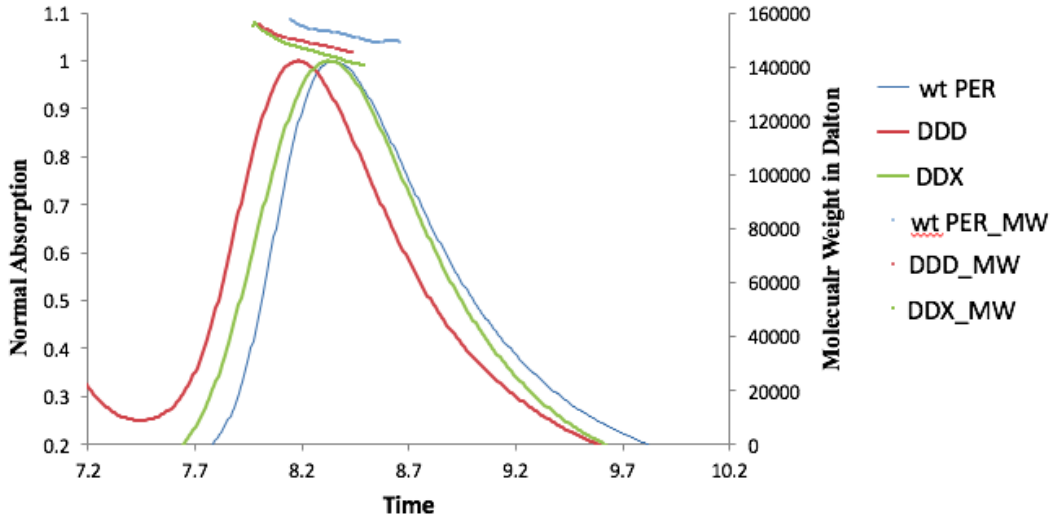

Figure 14: : MALS of PER and its phosphomimic mutants shows that the protein makes a dimer (~155kDa).

Trypsin Digests of PER (1-700) and its phosphomimic mutants

While MALS showed that WT PER, DDX and DDD make dimers, we performed trypsin digests to study local structural changes between WT PER and its variants. Trypsin cleaves at the carboxyl side of lysine and arginine in peptide chains. Thus, a protein that gets cleaved quicker has a more open conformation and vice versa. WT PER and phosphomimic mutants were partially digested by trypsin for 30, 45 and 60 seconds. WT PER was digested faster than mutants DDD and DDX. On the other hand, PER XXD was digested by trypsin faster than WT PER. Stability of AXX and DXX was comparable to WT PER. These results suggest that phosphorylation at the sites: S589, S604/S607/T610/S613 and S629 results in a more stable PER mutant (Figure 15).

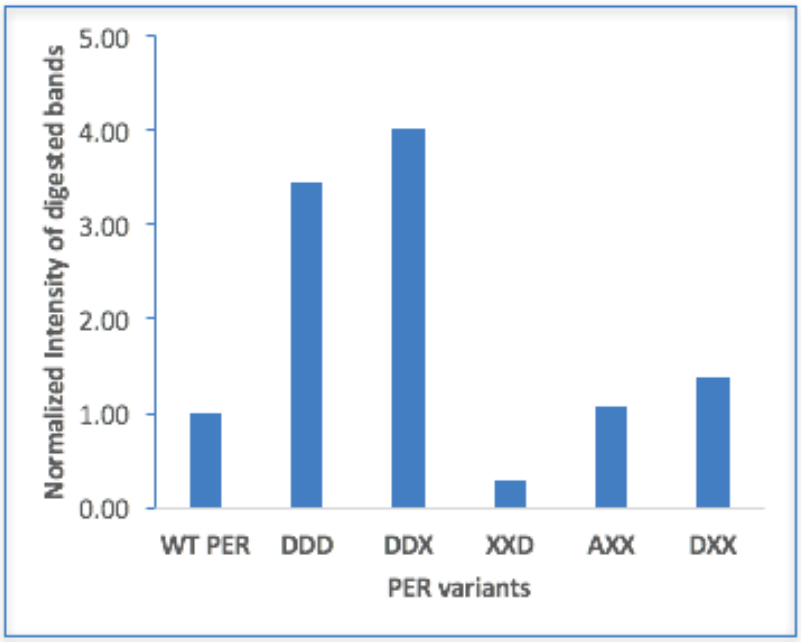

Figure 15: Bar chart shows the stability of PER and its mutants in the presence of trypsin. Higher the intensity of the digested bands, more stable the PER variant.

Discussion

In this study, we present a model in which phosphorylation of certain residues in PER positively affects its stability and conformation. While the PER phosphomimic mutant S589D (DXX) is unstable when expressed, DDX and DDD mutants are more stable compared to WT PER (XXX). This suggests that as PER gets progressively phosphorylated at sites S589, S604/S607/T610/S613 and S629, it becomes more stable and compact. Thus, later stages in the phosphorylation cascade are probably responsible for destabilizing PER.

Methods

Purification of PER and its mutants

The PER construct (1-700) is in pGEX-6-1p plasmid with a GST-tag and ampicillin resistance marker. To purify WT PER or its mutants, the PER construct was expressed in chloramphenicol resistant pLys S *E. coli* cells and then plated onto agar media containing ampicillin and chloramphenicol antibiotic. A colony was selected to grow 8 L cell culture 37° C in the presence of ampicillin and chloramphenicol. The cell culture was induced overnight at 17° C OD_{600} 0.6. After harvesting the cells, PER 7 were lysed using lysis buffer (150 mM NaCl, 50 mM Tris (pH 7.5), and 5 mM DTT). Bug Buster and nucleases RNase and DNase were added to the lysate. Protease inhibitor tablet was added to the solution and incubated for more than 90 minutes at room temperature. The protein was purified by using glutathione beads affinity chromatography using the GST-tag of PER. The protein elute was incubated with nuclease RNase before further purifying it by SEC. Mass spectrometry analysis confirmed that Period was purified. MALS confirmed that PER (monomer – 75kDa) forms a dimer of size 150 kDa in solution.

MALS of PER and its mutants

The molecular weights of PER and its phosphomimic mutants (5 mg/ml) were analyzed by SEC-MALS by using BioSepTM 5 μm SEC-s3000 LC Column. Buffer containing 150 mM NaCl, 50 mM Tris pH 7.5, and 5 mM DTT.

Trypsin Digests of PER and its mutants

PER and its mutants (75 μM) were incubated with 2 μL of trypsin (75 μM) for 15, 30 and 60 seconds. Digestion reactions were stopped using 2μL of trypsin inhibitor (150

μM). Digested products were analyzed by 4-12% NuPAGE BisTris gel, Coomassie stain, and quantified using ImageJ. The trypsin digest products intensities were normalized using undigested PER/mutant band intensities.

REFERENCES

1. Merbitz-Zahradnik, T. & Wolf, E. How is the inner circadian clock controlled by interactive clock proteins?. Structural analysis of clock proteins elucidates their physiological role. *FEBS Lett.* **589,** 1516–1529 (2015).
2. Chiu, J. C., Low, K. H., Pike, D. H., Yildirim, E. & Edery, I. Assaying locomotor activity to study circadian rhythms and sleep parameters in Drosophila. *J. Vis. Exp.* 1–8 (2010). doi:10.3791/2157
3. Peschel, N. & Helfrich-Förster, C. Setting the clock - By nature: Circadian rhythm in the fruitfly Drosophila melanogaster. *FEBS Lett.* **585,** 1435–1442 (2011).
4. Dunlap, J. C. Molecular bases for circadian clocks. *Cell* **96,** 271–290 (1999).
5. Mendoza-Viveros, L. *et al.* Molecular modulators of the circadian clock: lessons from flies and mice. *Cell. Mol. Life Sci.* **74,** 1–25 (2016).
6. Dubowy, C. & Sehgal, A. Circadian Rhythms and Sleep in Drosophila melanogaster. *Genetics* **205,** 1373–1397 (2017).
7. Chiu, J. C., Vanselow, J. T., Kramer, A. & Edery, I. The phospho-occupancy of an atypical SLIMB-binding site on PERIOD that is phosphorylated by DOUBLETIME controls the pace of the clock. *Genes Dev.* **22,** 1758–1772 (2008).
8. Yildiz, Ö. *et al.* Crystal structure and interactions of the PAS repeat region of the Drosophila clock protein PERIOD. *Mol. Cell* **17,** 69–82 (2005).
9. King, H. A., Hoelz, A., Crane, B. R. & Young, M. W. Structure of an Enclosed Dimer Formed by The Drosophila Period Protein. *J. Mol. Biol.* **413,** 561–572

(2011).

10. Saez, L. & Young, M. W. Regulation of nuclear entry of the Drosophila clock proteins period and timeless. *Neuron* **17,** 911–920 (1996).

11. Garbe, D. S. *et al.* Cooperative Interaction between Phosphorylation Sites on PERIOD Maintains Circadian Period in Drosophila. *PLoS Genet.* **9,** (2013).