The regulatory roles of alternative splicing and intrinsically disordered regions of proteins in circadian rhythms

A thesis submitted for honors in the degree of

BACHELOR OF SCIENCE

in

BIOCHEMISTRY AND MOLECULAR BIOLOGY

by

Ivette Perez

Department of Chemistry and Biochemistry, University of California,
1156 High Street, Santa Cruz, CA, 95064

Table of Contents

Acknowledgments

There are a lot of people who have helped me during my undergraduate journey and I will always be grateful for the help and guidance they provided. First, I will like to thank Professor Carrie L. Partch for allowing me to be part of her lab. She is an amazing scientist, mentor and role model. Thank you to the Partch Lab members who encouraged me and always believed in me. I would like to especially thank the following Partch lab members for their continuous support, for always listening to my endless rambling, and for helping me grow as a scientist: Leslee Nguyen, Gian Carlo Parico, Dr. Jennifer Fribourgh. I am grateful for the support and guidance I received from Yulianna Ortega and Xingci Situ, STEM Diversity Programs Director and STEM Diversity Programs Coordinator, respectively. Thank you to Edmundo Perez and Daniela Marquez, for always being there for me when I needed someone to talk to and for always being awesome friends! Lastly, I will like to thank my family members for their continuous love and support.

Chapter 1- Background

1.1 Problem Statement

The main goal for my independent research project was to gain insight how circadian rhythms are regulated at a molecular level. Most researchers in the field of chronobiology take genetic and cellular approaches, which leads to a deficit in knowledge of the detailed molecular mechanisms of circadian regulation. Although the founders of chronobiology, Colin Pittendrigh and Jurgen Aschoff, made many important findings in the 1950s-1960s, we still do not fully understand how circadian rhythms are regulated at the molecular level (Aschoff, 1965; Pittendrigh and Minis, 1964). While we understand the genetic network that comprises the clock and how the clock genes work with one another, we do not have enough insight into the biochemical mechanisms of the clock to understand why mutation in clock genes elicit particular effects and are linked to certain disorders. In order to understand the regulation of circadian rhythms at the biochemical level, we need to study protein-protein interactions between the core circadian proteins: the heterodimer transcription factor CLOCK:BMAL1, and their repressors (Gustafson and Partch, 2015). The first repressor I studied was the Clock-Interacting Protein Circadian, CIPC. This repressor was discovered using a two-hybrid assay that tested for protein interaction between CLOCK WT and $CLOCK^{\Delta 19}$ (Zhao et al., 2007). $CLOCK^{\Delta 19}$ is a mutant allele through alternative splicing that leads to an internal deletion of 51 residues encoded by Exon 19 in the disordered C-terminus of CLOCK. CIPC does not have any known vertebrate

homologs and its only known role is to inhibit transcriptional activation by CLOCK:BMAL1(Zhao et al., 2007). CIPC loses the ability to inhibit CLOCK:BMAL1 with the deletion of CLOCK Exon 19. In addition to studying the Exon 19-CIPC interaction, I studied the role of dimerization of Exon 19 and how it mediates the binding of two distinct CLOCK:BMAL1 heterodimers to tandem DNA E-box promoter regions through positive cooperativity. Positive cooperativity as a result of binding to tandem sites on DNA has been shown to produce robust circadian rhythms (Shimomura et al., 2013). The second circadian repressor I studied is Cryptochrome 1, (CRY1). For this repressor, I mostly focused on the C-terminal disordered tail domain of CRY1. A single nucleotide mutation of the CRY1 gene, leads to alternative splicing and deletion of Exon 11 in CRY1 (Patke, A. et al., 2017). Patke, A. et al., had shown through cell-based assays that deleting Exon 11, a 24-residue region on the CRY1 disordered C-terminus, increases inhibition of CLOCK:BMAL1 by CRY1. This mutation is common in people with Delayed Sleep Phase Disorder (DSPD), which has an estimated prevalence of up to 1 in 75 people in the general population. We hypothesized that Exon 11 might be interfering with repression by regulating CLOCK:BMAL1 binding. The Partch Lab had previously demonstrated how CRY1, without the disordered C-terminus, interacts with CLOCK. To probe how CRY1 C-terminus might alter the interaction with CLOCK and change affinity, first we needed to determine how the disordered c-terminus interacts with CRY1 PHR and then test how it might affect the binding to CLOCK.

1.2 Discovery of Circadian Rhythms

The first observations that organisms might have an internal clock were made in plants. In 1729, the French astronomer Jean Jacques d'Ortous de Mairan was intrigued by the rhythmic movements that a mimosa plant would follow during a day (de Mairan, J. 1729). He noticed that the mimosa plant would open up its leaves during the day to potentially uptake the sunlight but would close its leaves during the night. To test whether the plant would keep the rhythmicity independent of sunlight, he placed the plant in constant darkness and, to his surprise, the mimosa still opened and closed its leaves at the same time of the day, indicating the possibility of an internal timing mechanism (de Mairan, J. 1729). We now know that many organisms, from bacteria to humans, have internal clocks to help synchronize their behavior and physiology to anticipate changes in a day and changes in their surroundings (Bell-Pedersen et al., 2005). In humans, our circadian rhythms not only control our sleep/wake cycle but many aspects of our physiology like metabolism, hormone production and DNA repair (Figure 1).

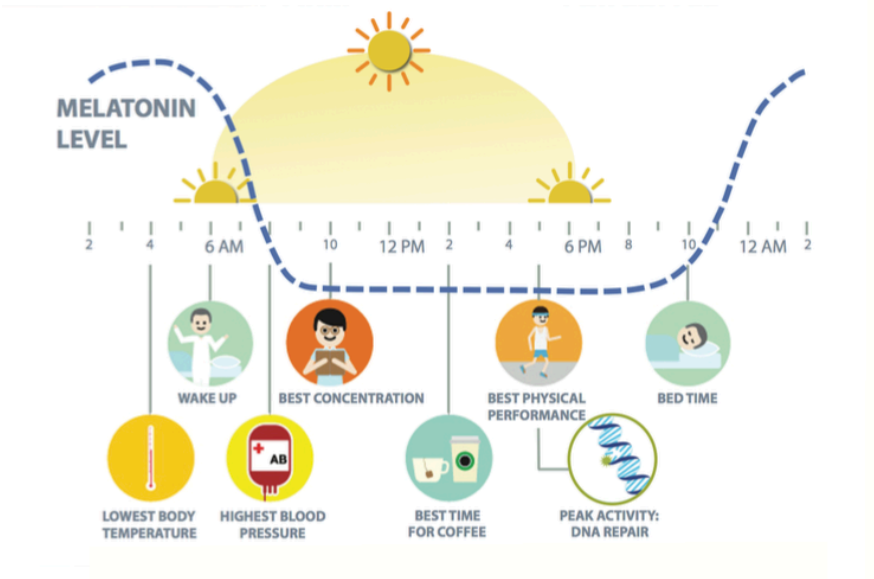


Figure 1: Circadian rhythms control our physiology
Circadian rhythms regulate many physiological processes such as sleep/wake cycle, body temperature, blood pressure, and DNA repair. Melatonin levels (represented by dotted line), which regulate sleep/wake cycles, increase in the late evening but then begin to drop early morning.

1.3 Circadian Rhythms at a Molecular Level

The mammalian circadian clock ensures that our physiology and behavior are synchronized with the 24-hour solar day. The heterodimer transcription factor CLOCK:BMAL1 binds to DNA at E-box promoter regions and promotes transcription of more than 40% of our complete genome including core repressors Period (PER) and Cryptochrome (CRY) (Zhang et al., 2014). After their peak protein expression is reached, complexes of the PER and CRY proteins then translocate to the nucleus where they interact with CLOCK:BMAL1 and inhibit their own transcription, closing the negative feedback loop (Gustafson and Partch, 2015). Based on genome-wide binding studies of these complexes, we know that there are two repressive complexes that form on DNA throughout

the night, the early and late repressive complex (Figure 2) (Koike et al., 2012).The transition from the repressive state to the active state depends on the eventual degradation of the proteins found during the repressive complex due to post-translational modifications like ubiquitination, thus poising CLOCK:BMAL1 to begin another round of circadian gene transcription in the morning.

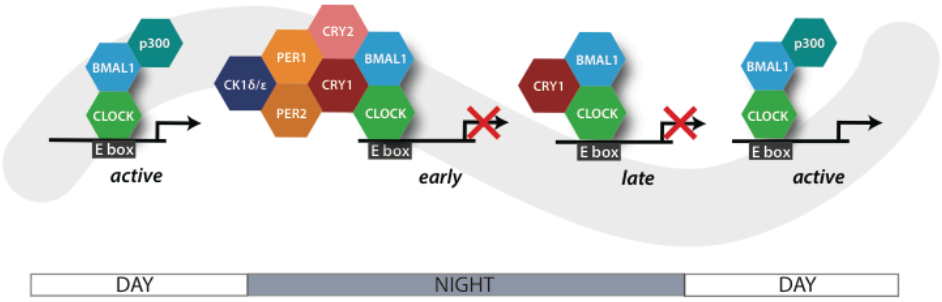


Figure 2: Transcription/Translation Feedback Loop
The CLOCK:BMAL1 complex regulates clock genes in coordination with its PER and CRY repressors. This regulation in gene expression leads to a peak of protein expression once every 24-hours, giving rise to the characteristic oscillation of circadian genes.

1.4 Alternative Splicing and Disordered Regions

Many of the circadian proteins we work with have intrinsically disordered regions (Gustafson and Partch, 2015). It is common for transcription factors and other proteins associated with cellular signaling pathways to have intrinsically disordered regions because it allows protein to have a broad spectrum of functions (Tóth-Petróczy et al., 2008; Uversky, 2013). In the mid-1990s, the role of intrinsically disorder proteins (IDPs) in eukaryotes began to be recognized as

being important for cellular regulation and signaling (Wright and Dyson, 2015). Mutations of IDPs or changes in their cellular concentration are frequently associated with disease. For example, the coactivators, CBP and p300 are able "to interact with more than 400 transcription factors and other regulatory proteins" due to their intrinsically disordered regions (Dyson and Wright, 2016). Recently, there has been increased interest in studying the intrinsically disordered regions (IDR) of structured proteins since there are a high number of eukaryotic proteins that contain both structured domains and IDRs. These IDRs are highly dynamic, which allows them to mediate interactions with different proteins and perform multiple roles in cell signaling pathways to increase their functional range (van der Lee et al., 2014) (Figure 3).

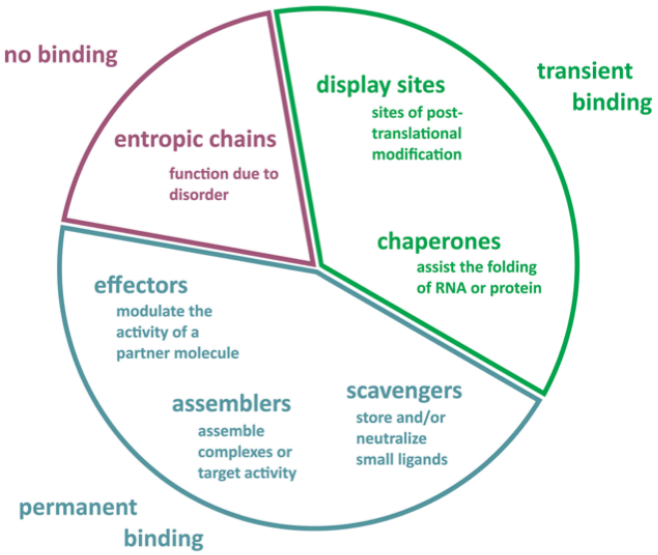


Figure 3: Roles of IDRs
This chart shows the different roles that IDRs have due to their highly dynamic nature in cells. Adapted from van der Lee et al.

1.5 Summary

In mammals, the disruption of circadian rhythms can increase the risk of developing psychiatric disorders, diabetes, cancer and many other diseases (Zhang et al., 2014). Many circadian proteins have disordered regions that potentially mediate many interactions and allow them to interact with a wide range of proteins and perform activities perhaps even outside of the core clock timekeeping mechanism. Circadian proteins have a wide range of interactions due to their complex structure and mutations in these proteins could explain the increased risk of developing certain diseases and disorders. Understanding how circadian rhythms are regulated at the molecular level could give us mechanistic insight to develop therapeutics to manipulate the proteins responsible for giving rise to circadian rhythms and control the mammalian clock.

Chapter 2- Investigating how CLOCK Exon 19 promotes robust circadian rhythms

2.1 Abstract

Most organisms are governed by an internal molecular clock that ensures temporal homeostasis throughout the 24-hour solar day (Zhang et al., 2014). The circadian transcription factor complex known as CLOCK:BMAL1 drives the transcription of the core-clock genes known as *Cry* and *Per* (Gustafson and Partch, 2015). In turn, the transcriptional activity of the CLOCK:BMAL1 is negatively regulated by PER and CRY proteins to complete the central feedback

loop that makes up the circadian clock. Another negative regulator of interest is CLOCK-Interacting Protein Circadian (CIPC). CIPC inhibits CLOCK:BMAL1 transcriptional activity by binding directly to the 51 amino acid region of CLOCK encoded by its nineteenth exon (Zhao et al., 2007). Through a series of *in vitro* binding assays based on size exclusion chromatography, I determined that CIPC and CLOCK Exon 19 form a stable complex together. I then aimed to solve the structure of the CLOCK Exon 19-CIPC complex using x-ray crystallography to gain mechanistic insight into the essential role that CLOCK Exon 19 plays in circadian rhythms. Loss of Exon 19 due to alternate splicing in the mutant mouse, *Clock* $^{\Delta 19}$, it is known to greatly hamper circadian rhythms (King et al., 1997). Understanding how Exon 19 interacts with its regulators would bring us one step close to determining why it is essential for circadian rhythms. While optimizing protein crystals of CLOCK Exon 19 bound to CIPC, a competing lab published the structure (Hou et al. 2017). However, this structure confirmed the biochemical and biophysical data I had already obtained demonstrating that CLOCK Exon 19 is a dimer. We found these data very interesting, because it suggests that Exon 19 could help to bring two molecules of CLOCK together on nearby DNA binding sites, which is required for robust circadian rhythms (Shimomura et al., 2013). I used the structure to design mutations that could potentially disrupt the Exon 19 dimer to further study how the dimerization of Exon 19 influences CIPC binding and circadian rhythms.

2.2 Introduction

CLOCK Exon 19 is found in the C-terminal disordered tail of CLOCK (Figure 4). The importance of CLOCK Exon 19 (residues 514-564) was initially discovered in 1997 using a forward genetic screen in mice with the mutagen N-Ethyl-Nitrosourea (ENU) (King et al., 1997). This mutation screen gave rise to a founder line with an A→T nucleotide transversion, which created a site for alternative splicing and leads to removal of Exon 19. When placed in complete darkness, homozygous *Clock*$^{\Delta 19}$ mice lost rhythmicity, demonstrating that Exon 19 has an essential role in generating circadian rhythms in cells. Therefore, I sought to identify the biochemical role that CLOCK Exon 19 plays in maintaining circadian rhythms.

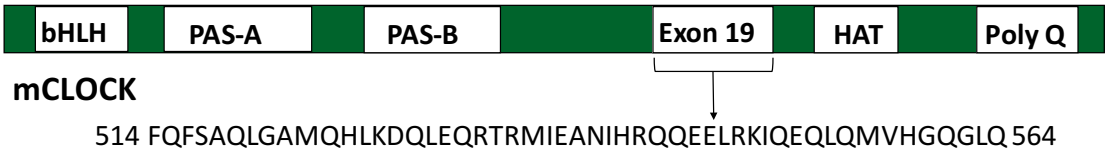


Figure 4: Domain schematic of CLOCK
This figure shows the different domains on the protein CLOCK. Exon 19 (residues 514-564) is located on the disordered C-terminus of CLOCK.

The CLOCK:BMAL1 heterodimer binds to E-box elements in DNA to promote the transcription of core clock genes (CCGs) (Ripperger and Schibler, 2006). Previously, it was shown that CLOCK Exon 19 is essential for CLOCK:BMAL1 to bind to tandem E-Boxes and promote strong circadian rhythms using *in vitro* and cellular approaches (Figure 5) (Shimomura et al., 2013). Shimomura et al.,

demonstrated that the complex CLOCK:BMAL1 has a higher affinity than the *Clock*$^{\Delta 19}$:BMAL1 for DNA, with a binding affinity change of up to 11.8 fold difference depending on the E-box site, suggesting Exon 19 could play a role in the complex recruitment to the DNA though positive cooperativity increasing the affinity of the complex. In addition to playing a role in positive cooperativity an Exon 19 dimer could also mediate interactions with important coactivators and repressors. Currently, we know Exon 19 provides the necessary binding interface for important regulators like the coactivator MLL1 and the repressor CIPC; without Exon 19, the capacity of these cofactors to bind CLOCK is lost (Katada and Sassone-Corsi, 2010; Zhao et al., 2007). However, even though the essential role of Exon 19 was first identified in 1997, we still do not understand how it directly mediates interactions with both negative and positive regulators at the molecular level.

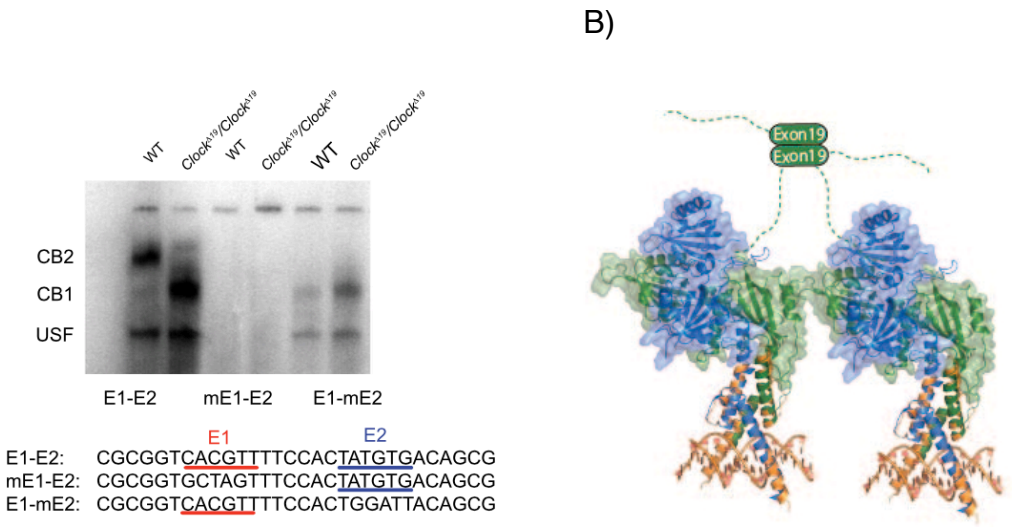


Figure 5: A potential role for CLOCK Exon 19 dimerization in binding to tandem E-box sites on DNA.
A) This native polyacrylamide gel shows an electrophoretic mobility shift assay (EMSA) of CLOCK:BMAL1 complexes from mouse liver nuclear extract binding to a labeled oligonucleotide containing the tandem E-boxes or various mutants. The study focused on a polymorphism in he gene USF1 that allowed mice to suppress the clock disrupting effects of *Clock* Δ19. The assay confirmed the importance of having the idealized E-box sequence of the E1 site for recruitment of a second complex to the non-consensus E2 site. The tandem complex (CB2) forms with WT CLOCK:BMAL1, while only the single complex (CB1) preferably forms with *Clock* $^{\Delta 19}$:BMAL1 or when there is a mutation on E2. (Adapted from Shimomura, et al., 2013) B) Structural model built off of the bHLH-PAS dimer (green, CLOCK; blue, BMAL1) (PDB:4F3L) overlaid with a structure of their bHLH domains bound to an E-box site (orange) (PDB:4H10) is representing possible binding of the CLOCK:BMAL1 tandem complex to E-box elements mediated by dimerization of Exon 19.

2.3 Materials and Methods

The DNA encoding the isolated CLOCK Exon 19 (residues 514-564) and the minimal CLOCK-binding region of CIPC (residues 318-367)(Zhao et al., 2007) were cloned into the GST Parallel vector (based on the vector, pGEX 4T1) and the pHis-NusA-XL Parallel vector (based on the pET22b vector backbone), respectively. These plasmids contain a fusion protein tag (GST or His-NusA) that can be leveraged during expression for increased solubility in *E. coli*, and for

affinity chromatography. Both my genes of interest had been previously cloned and sequenced by Dr. Carrie Partch.

2.3.1 Protein Expression

The plasmids for GST CLOCK Exon 19 and HisNusA-XL CIPC were both expressed in *E. coli* Rosetta (DE3) cells, using an IPTG-inducible T7 RNA polymerase to control expression of the proteins of interest. I used Rosetta strain of *E. coli* because it exogenously expresses codons that are rarely used in *E. coli* but are commonly found in mammalian proteins; this can enhance the overall level of protein production. I mixed 1 μL of plasmid DNA with 50 μL of chemically competent Rosetta (DE3) cells and incubated the mixture on ice for 30 minutes in order for the DNA to anneal to the cell membrane. After incubation, I heat shocked the cells for 45 sec in a 42°C water bath in order to allow the cells to uptake the DNA, allowing the cells to recover on ice for a minute before adding 300 μL of Super Optimal broth with Catabolite repression (SOC) growth medium. Afterward, I incubated the transformation in a 37°C shaker for 45 minutes to grow before plating 120 μL of the cells on an LB agar plate with ampicillin and chloramphenicol because my plasmid contains an ampicillin resistance gene and the Rosetta cells are resistant to chloramphenicol. The antibiotics on the plate therefore selected only the Rosetta cells that contain my plasmid to grow on the plates while incubated overnight at 37°C.

2.3.2 Protein Purification

The following morning after the transformation (16 hours incubation at 37°C), I inoculated a 5 mL Luria Broth (LB) starter that contained 100 μg/mL ampicillin and 35 μg/mL chloramphenicol to allow the cells to get accustomed to growing in liquid media and get past the lag phase of growth. After a few hours, I then inoculated a 75 mL starter that contained both antibiotics and let it grow overnight at 37°C. The following day (16 hours incubation at 37°C), I inoculated 6 x 1 L LB cultures with 10 mL each from the overnight culture. For both CLOCK Exon 19 and CIPC, protein expression was induced with 0.5 mM IPTG when the culture reached an optical density (OD_{600}) of 0.6-0.8 and were then grown overnight at 18°C. The next day, cultures were spun down at 4,000 RPM for 15 min (at 4°C). Since each protein has a different tag, different affinity chromatography resins and buffers were used to isolate the soluble protein of interest from cellular extracts. For CLOCK Exon 19, I used Gluthatione Sephorose 4B (GE Healthcare) because my Exon 19 was fused to the GST tag. I used Buffer A (50 mM Tris, 150 mM NaCl, 2 mM DTT, 5% Glycerol, pH 7.5) to re-suspend the cell pellet. I used an Avestin cell disruptor to lyse the cells, which uses high-pressure to mechanically disrupt the cell membrane. The lysed cells were clarified by centrifugation at 19,000 RPM for 45 min at 4°C. The soluble protein was then purified by GST affinity chromatography on Glutathione Sephorose 4B resin. Briefly, I added the clarified supernatant directly to resin previously equilibrated with buffer A in 50 mL conical tubes and let it incubate rotating end-over-end for

30 min to allow protein binding to the resin. Following this incubation, the mixture was centrifuged at 1,000 RPM for 2 minutes at 4°C, subsequently discarding the supernatant (i.e., the flow-through). I then added more buffer A and gently resuspended the resin, spinning down one more time for the first wash of the resin. Once again, I poured off the supernatant and added more buffer A to the resin, but this time I added the resuspended resin to a gravity column for a slow, high volume wash. I monitored for loss of protein in the flow-through using Bradford reagent to know when to stop the final wash. I then transferred the resin-bound protein to a new 50 mL conical tube for tobacco etch virus (TEV) protease cleavage. I added 1 mg of GST-tagged TEV protease for every ~20 mg of GST-tagged Exon 19 bound to resin to cleave the GST tag off overnight at 4°C. This process allows for both the GST tag and the GST-TEV protease to stay bound to the resin, while my cleaved protein of interest will be found in the flow through. For purification of His-NusA-XL CIPC, the soluble protein was purified the same way as described above with minor changes. The buffer used to resuspend my cells was 50 mM Tris, 150 mM NaCl, 20 mM Imidazole, 5 mM Beta-Mercaptoethanol (BME), pH 7.5. The soluble protein was then purified similarly using Ni-NTA Agarose affinity chromatography (Qiagen). However, here I used a His_6-TEV to cleave the tag off.

To further purify my proteins, I used sized exclusion chromatography. I used the Superdex 75 prep HiLoad 16/60 column to separate my proteins based on size with the buffer 20 mM HEPES, 125 mM NaCl, 2 mM DTT, 5% (v/v) Glycerol, pH

7.5. This technique allowed me to separate my protein from other non-specific proteins that might have been found in the flow through after TEV cleavage. Both CLOCK Exon 19 and the minimal CLOCK-binding region of CIPC are relatively small peptides of 5-6 kDa, and both eluted at a volume of about 80 mL. I performed SDS-PAGE electrophoresis of fractions from the column to visualize the proteins. After concentrating the purified proteins, I used liquid nitrogen to flash freeze small aliquots of the proteins for long-term storage at -80°C.

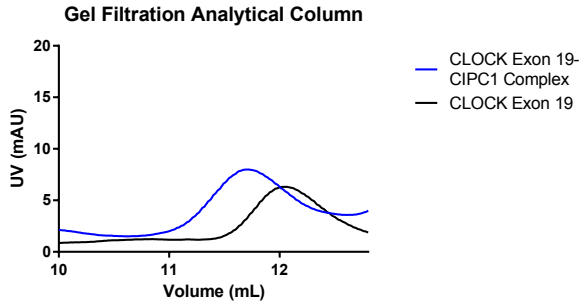
2.4 Analysis of CLOCK Exon 19 and CIPC Interaction

2.4.1 Confirmation of CLOCK Exon 19-CIPC complex formation

We know a short region of CIPC is required to interaction with CLOCK in an Exon 19-dependent manner through *in vitro* transcribed proteins used in binding assays, but it has not been explicitly shown that these two proteins interact directly (Zhao et al., 2007).In order to confirm that Exon 19-CIPC complex formation had occurred, we used size exclusion chromatography. I combined purified CLOCK Exon 19 and CIPC proteins in equimolar amounts and allowed a brief time on ice (~10 minutes) for a complex to form. I injected CLOCK Exon 19 alone or the protein mixture into a Superdex 75 Analytical GL 10/300 column to confirm the complex formation. After chromatographic separation and subsequent SDS-PAGE analysis, I found that both proteins co-eluted at a retention volume corresponding to a larger molecular weight than either protein

(Figure 6). This suggests that both proteins form a larger complex that can be distinguished from the isolated proteins through size exclusion chromatography.

A)



B)

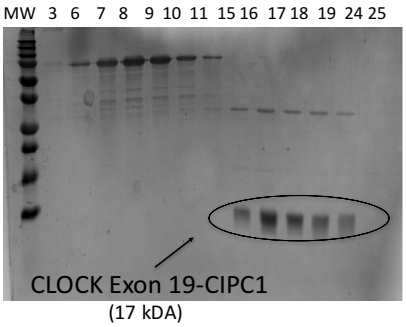


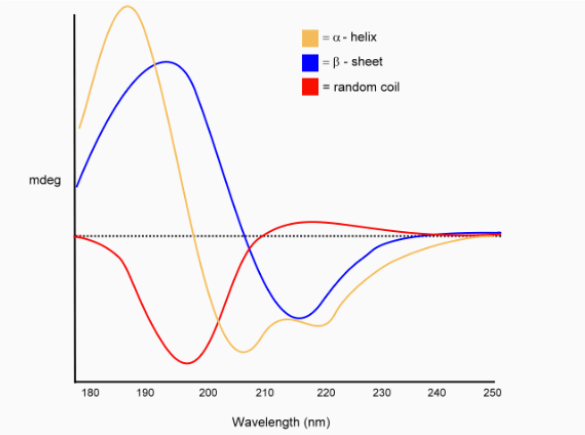
Figure 6: CLOCK Exon 19-CIPC Complex formation
(A) The CLOCK Exon 19-CIPC complex (blue) eluted before CLOCK Exon 19 alone (black) indicating that the Exon 19-CIPC complex has a higher molecular weight than Exon 19 alone. (B) An 18% SDS-PAGE gel was run with the fractions collected from the analytical column to confirm complex formation. Proteins are stained with Coomassie Brilliant Blue dye.

2.4.2 Investigating the structure of Exon 19 by Circular Dichroism spectroscopy

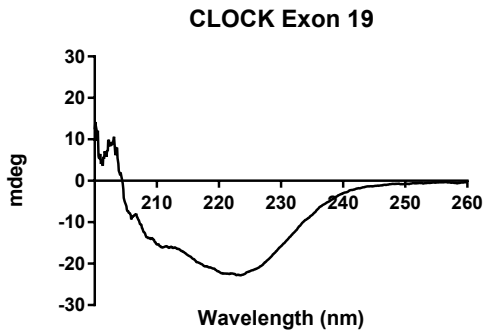
Even though we know that Exon 19 is on the disordered C-terminus tail of CLOCK, secondary structure predictions suggest that Exon 19 is helical. PSI-PRED secondary structure predictions suggested that Exon 19 is a coiled coil and is the only structured region in the CLOCK C-terminus (Figure 7). To investigate the secondary structure of isolated Exon 19, I performed circular dichroism (CD) spectroscopy. CD is a common technique used in biochemistry because it allows for the detection of protein secondary structure based on spectra collected at a low concentration of an unlabeled sample. CD spectra

were collected on a JASCO J-1500 CD spectropolarimeter using 1 mm pathlength quartz cuvette. The proteins were in a low salt buffer ideal for CD, containing 20 mM HEPES, 125 mM NaCl, 2 mM DTT, pH 7.5 at a concentration of 26 μM. As seen in Figure 8, CLOCK Exon 19 has an alpha helical secondary structure and Exon 19 retains the secondary structure when is bound to CIPC.

A)



B)



C)

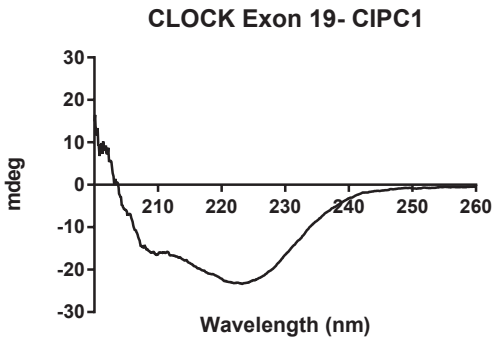


Figure 7: CD Spectra of CLOCK and CIPC complexes
The CD spectra collected from the samples tested: (A) Examples of CD spectra based on the 3 possible secondary structures: Alpha helix (yellow), beta sheet (blue), or random coil (red). Proteins exhibit a linear combination of these pure spectra depending on their secondary

structure **(B)** The spectra show CLOCK Exon 19 as an alpha helical secondary structure, (C) CLOCK Exon19-CIPC also has an alpha helical structure.

2.4.3 SEC MALS

Bioinformatic predictions suggests that Exon19 is a dimer, based on this I expect to see the masses equivalent to a dimer for Exon 19 (12 kDa) and for the complex it would be ~17 kDa. To determine the stoichiometry of isolated Exon 19 and the Exon 19-CIPC complex, I utilized size exclusion chromatography (SEC) coupled with multiangle light scattering (MALS). SEC-MALS is a technique that determines the molar mass of molecules independently of shape after separation by their hydrodynamic radii in the size exclusion column. The mass of a molecule, and its associated polydispersity, will be determined the light scattering produced by the multi-angle laser. The calculated molecular weight of Exon 19 is 6,116.9 g/mol and CIPC is 5,539 g/mol. Based on SEC-MALS of Exon 19 bound to CIPC, the calculated molecular weight was 17 kDa and Exon 19 by itself was 12 kDa, we believe that the Exon19-CIPC complex stoichiometry contains 2 molecules of Exon 19 (a dimer) bound to one molecule of CIPC. This suggests a model in which dimerization of Exon 19 is important for CIPC binding and repression of the CLOCK:BMAL1 complex, and also suggests that the Exon 19 may play a role in transcription factor binding to tandem E-boxes.

2.4.4 CLOCK Exon 19-CIPC

The x-ray crystal structure of the CLOCK Exon19-CIPC complex was published in Summer 2017 by the Zhang group at the University of Texas Southwestern

Medical Center (Hou et al.). The published structure (PDB ID 5VJI) confirmed all my initial solution measurements by showing a dimer of Exon 19 bound to one helical peptide of CIPC (Figure 8).

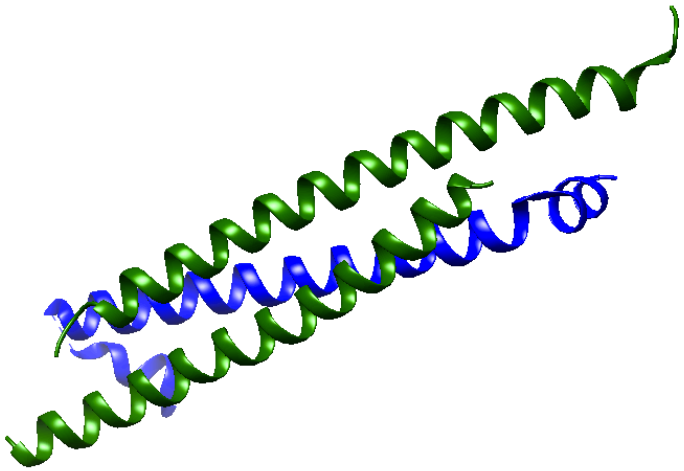


Figure 8: The CLOCK Exon 19-CIPC crystal structure
The CLOCK Exon 19-CIPC structure shows the predicted stoichiometry of the complex: two subunits of Exon 19 (green) bound to one subunit of CIPC (blue) (PDB: 5VJI).

2.5 CLOCK Exon 19 mutations

To date, the exact role of the dimerization of Exon 19 is not understood. Data from the Partch lab and others have shed light at the importance of Exon 19. The structure of the Exon19:CIPC complex confirmed the initial stoichiometry we estimated from the SEC-MALS data, but we have yet to independently confirm whether Exon19 dimerization is the required binding interface for important transcriptional coactivators and repressors. To further investigate if the dimerization of Exon 19 provides the necessary interface for binding to key

transcriptional factors, I designed single point mutations to disrupt the Exon 19 dimer based on the structure from the Zhang lab (Figure 9). The structure allowed me to see the residues that make up the dimerization interface and design mutations that may be more disruptive than sequence predictions alone would have allowed. The first step in mutating the coding sequence of Exon 19 was to design specific primers that would target the sequence of interest with the mutation encoded to amplify the plasmid using polymerase chain reaction (PCR). PCR Site Directed Mutagenesis is a common tool used in molecular biology that allows researchers to introduce amplify particular regions of DNA or encode mutations in a desired DNA sequence. After sequencing confirmation was received, we moved on to transforming the DNA into Rosetta (DE3) cells using the exact protocol described in sections 2.4.1. I then purified the Exon 19 mutants using methods described above (Section 2.4.2).

A) B)

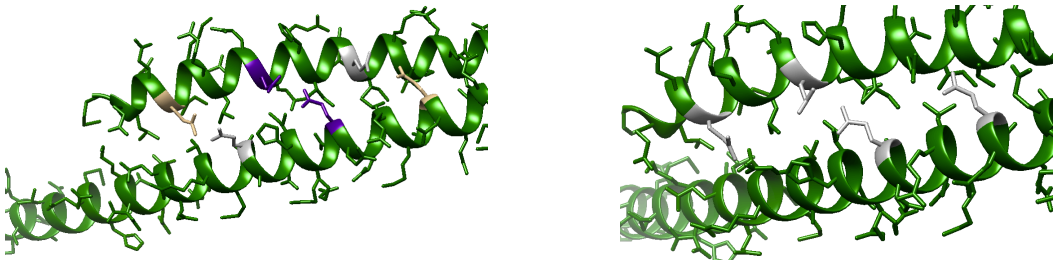


Figure 9: CLOCK Exon 19 Mutations
A) Using the published structure of CLOCK Exon 19 (PDB: 5VJI), I designed point mutations that could potentially disrupt Exon 19 dimerization. B) After SEC-MALS analysis, it showed that Exon 19 dimerization was not completely disrupted, so I designed a double mutant (Q545A/Q552A).

In order to examine for dimerization of Exon 19, I performed SEC-MALS on the mutants alongside the wild-type protein with the assistance of Jeffrey Swan, a graduate student in the Partch Lab. The SEC-MALS data allowed us to determine whether the single point mutations disrupted the Exon 19 dimer. Based on the data and the polydispersity, a low polydispersity means that the complex is more uniform, we know that none of the single mutants completely disrupted the dimer. Based on the mutants that change the SEC-MALS profile compared to the wild-type, I decided to design a double mutant. I performed site directed mutagenesis to clone the double mutant (Q545AQ552A) using the same protocol as described above, using GST Exon 19 Q545A as the DNA template for the PCR reaction.

To determine whether the mutations disrupted the native alpha helical structure of Exon 19, I collected CD spectra of the mutants. All the mutants retained the expected alpha helical structure as seen in Figure 10**.** I changed the buffer in order to minimize the background noise obtained in the CD spectra from CLOCK Exon 19 and CLOCK Exon 19-CIPC complex. After talking to Dr. Eefie Chen, I decide to use the following buffer: 20 mM Sodium Phosphate, 100 mM NaCl, pH 7.5.

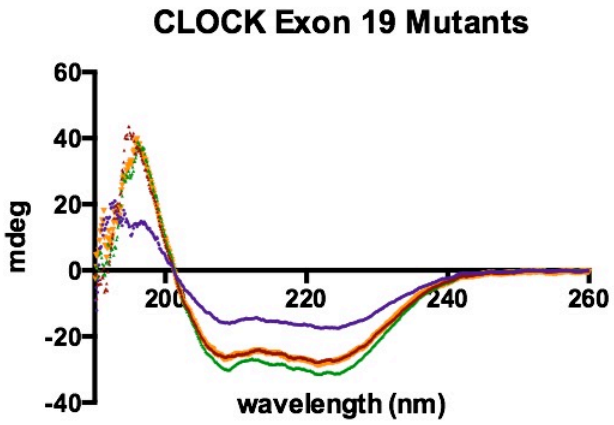


Figure 10: CD spectra of Exon 19 Mutants.
The single point mutants used in this study maintained alpha helical structure. The spectra of Exon 19 mutants Q552A (green), T534A (yellow), Q545K (red), E538A (purple) have distinct negative ellipticities at 208 nm and 222 nm, indicating alpha helical secondary structure.

2.6 Discussion

Based on my data and the Exon19-CIPC crystal structure, it appears that CLOCK Exon 19 dimerization provides the surface necessary for binding CIPC and possibly other repressors or activators. Although I have done initial work in trying to understand the role of Exon 19 in regulating circadian rhythms, there is still a lot of work to be done before we can have a more accurate model of how it contributes. It is important to understand how CLOCK:BMAL1 is recruited to DNA and regulated to promote transcription of clock controlled genes. Currently, we are in favor of a model where Exon 19 mediates the dimerization of CLOCK to facilitate binding of CLOCK:BMAL1 to tandem E-box elements through positive cooperativity. Both our data and the published structure confirmed that the dimerization of Exon 19 forms a binding interface for CIPC. Once I identify a mutation that disrupts Exon 19 dimerization, I will first probe the role of Exon 19

in vivo by analyzing CLOCK:BMAL1 activity on a luciferase reporter assay driven by a *Per1* promoter followed by a complementation test using our *Clock* -/- mouse embryonic fibroblasts using a lentivirus to study defects in circadian rhythms due to lack of Exon 19 dimerization. Secondly, I will test whether deletion of Exon 19 and/or dimer-disrupting mutations influence CLOCK dimerization of the full-length protein in cells by performing Co-immunoprecipitation assays. The importance of Exon 19 will further be studied by analyzing the effect of dimer disruption and its ability to bind to other transcription factors. One coactivator of interest is the histone methyltransferase MLL1, which loses the ability to bind to CLOCK:BMAL1 with the deletion of Exon 19 (Katada and Sassone-Corsi, 2010). MLL1 is required to create a relaxed, open chromatin state which allows CLOCK:BMAL1 to bind to DNA and promote transcription of clock controlled genes (Katada and Sassone-Corsi, 2010). Our lab has previously shown that both repressors and coactivators bind a mutually exclusive binding site on the BMAL1 transactivation domain (Xu et al., 2015). We believe that CLOCK Exon 19 may also bind to either CIPC (a repressor) or MLL1 (a coactivator) in a mutually exclusive manner. Therefore, elucidating the interplay of MLL1, CIPC, and dimerization of Exon 19 will give us a better understanding how circadian rhythms are generated at the molecular level.

Chapter 3-Determining how an intramolecular interaction in regulates clock timing and sleep phase

3.1 Abstract

The deregulation of circadian rhythms can lead to either advanced or delayed sleep phase disorders that influence the timing of sleep onset, due to clock control of the hormone melatonin. Delayed sleep phase syndrome (DSPD) is particularly disruptive, as the afflicted are “night owls” that are only able to fall asleep hours after midnight and have difficulty waking up in the morning (Shi et al., 2017). Circadian rhythms are driven by a negative feedback loop where the heterodimer CLOCK:BMAL1 promotes transcription of clock controlled genes (CCGs), including its own repressors, the Period (PER) and Cryptochrome (CRY) proteins. CRY proteins contain a photolyase homology region (PHR) and an intrinsically disordered C-terminal tail. Our lab recently identified how the CRY1 PHR binds to parts of both CLOCK and BMAL1 to repress transcriptional activation, thereby closing the negative feedback loop. However, it is not known if the CRY1 PHR and tail interact to control this process. A highly prevalent human DSPD mutation was just identified that leads to alternate splicing and deletion of CRY1 Exon 11, a region that encodes for 24 amino acids in the disordered tail. The CRY1Δ11 variant exhibits enhanced binding to CLOCK:BMAL1 and a longer circadian period, leading to delays in melatonin expression and late/fragmented sleep at night. Given our background, we hypothesized that the CRY1 tail might bind the PHR transiently to regulate its interactions with, and repression of, CLOCK:BMAL1. Through a series of

fluorescence polarization binding assays, I show here that the CRY1 C-terminal tail does binds directly to the PHR, and that the 25 amino acids encoded by Exon 11 are required for high affinity binding. Therefore is poised to directly regulate how CRY1 binds to CLOCK:BMAL1. This work provides initial insight into the molecular basis of the DSPD and the role of the CRY1 tail in circadian timekeeping.

3.2 Introduction

Based on pioneering chromatin immunoprecipitation-sequencing studies done by Joe Takahashi's lab, we now know that there are two distinct circadian repressive complexes that can be found bound to CLOCK:BMAL1 depending of the time of night (Koike et al., 2012). Figure 11 shows how these complexes evolve over time, with the initial PER-CRY "early" complex binding around dusk and a "late" complex of CRY1 alone bound to CLOCK:BMAL1 right before dawn. CRY1 is therefore a potent inhibitor in the early morning and is able to bind and inhibit CLOCK:BMAL1 alone, in the absence of PER proteins. Previously, the Partch lab demonstrated how CRY1 uses two distinct sites to bind to the CLOCK:BMAL1 complex (Michael et al., 2017; Xu, H. et al. 2015), showing how the CRY1 PHR coordinates both of these interactions. The first interaction is mediated by the so-called secondary pocket of CRY1, into which docks the PAS-B domain of CLOCK (Michael et al., 2017). The second interaction with BMAL1 occurs on a highly conserved helical motif right before the disordered C-terminal

tail begins (Xu et al., 2015). Despite understanding the roles for these two sites, we do not understand the functional role of the disordered C-terminus tail and how it might affect CRY1 interactions with CLOCK or BMAL1. None of the mammalian CRY crystal structures include the disordered C-terminus—we simply do not know if it binds to the PHR domain, and if so, where it docks on the PHR. In other organisms like *Drosophila melanogaster* and *Arabidopsis thaliana*, CRYs have a regulatory role that is mediated by light, where the CRY PHR and its C-terminal tail are bound in an inactive state during the dark, but become activated with light, leading to release of the tail from the PHR. This sensitivity to light has been lost in mammalian cryptochromes, but To probe whether the mammalian C-terminal tail might have a similar regulatory role, I have been focusing on the late repressive complex by studying the interaction between CRY1 PHR and its C-terminal tail to understand how it could regulate binding to CLOCK or BMAL1 and play a role in DSPD.

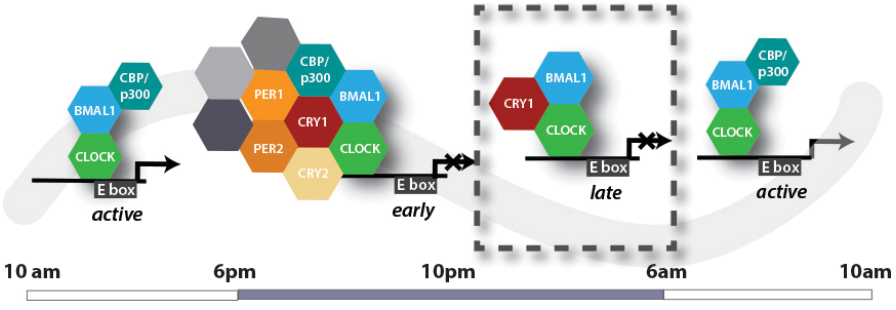


Figure 11: The core transcription/translation feedback loop of the mammalian clock
The transcription factor CLOCK:BMAL1 oscillates through cycles of activation and repression throughout the day. There are two distinct repressive complexes: the early repressive complex and late repressive complex. Even through both repressive complexes are needed for

timekeeping, the late repressive complex is required for maintaining healthy circadian rhythms. The only repressor bound to CLOCK:BMAL1 in the late repressive complex is CRY1.

Some patients with delayed sleep phase disorder (DSPD) have a variation in the *Cry1* gene that is estimated to be present in up to 1 out of 75 people depending on the genetic background, making it the most prevalent circadian variation found in the general human population to date (Patke et al., 2017). The *Cry1* variation arises from an A→C nucleotide transversion creates an alternative splicing site that leads to deletion of Exon 11, a 24 residue region in the disordered C-terminus of CRY1. The authors of this paper showed that deletion of Exon 11 leads to enhanced inhibition of CLOCK:BMAL1 by the CRY1Δ11 variant, which likely arises from an apparent higher affinity for the CLOCK:BMAL1 complex. Even though this indicates the importance of Exon 11 in regulating circadian rhythms, this worked lacked the molecular mechanism demonstrating how Exon 11 modulates CRY1 function. Elucidating how circadian proteins interact at the molecular level is vital for understanding how we can manipulate our clock in order to develop therapeutic strategies for circadian-related disorders.

An ortholog to the DNA repair enzyme photolyase, mammalian cryptochromes were first discovered in 1995 using the human genome database (Sancar, 2004). Orthologs are genes that share common ancestral DNA and that have diverged after a speciation event. Although structurally related to the photolyase enzyme, cryptochromes lack the ability to repair DNA damage (Michael et al., 2017). The CRY photolyase homology region (PHR) is structurally similar to the photolyase

enzyme, but cryptochomes have a unique disordered C-terminal tail that is not present in photlyase (Figure 12). Since deleting the CRY C-terminal tail changes the period, but is not required for timekeeping, most of the initial work focused on the CRY PHR. Khan, K.S, et al. demonstrated that the CRY1 PHR is all that is required to reconstitute the clock in Cry1-/-Cry2-/- cells, but now we know that the C-terminal tail plays an important role in regulating CRY1 interactions with other proteins (Khan et al., 2012).

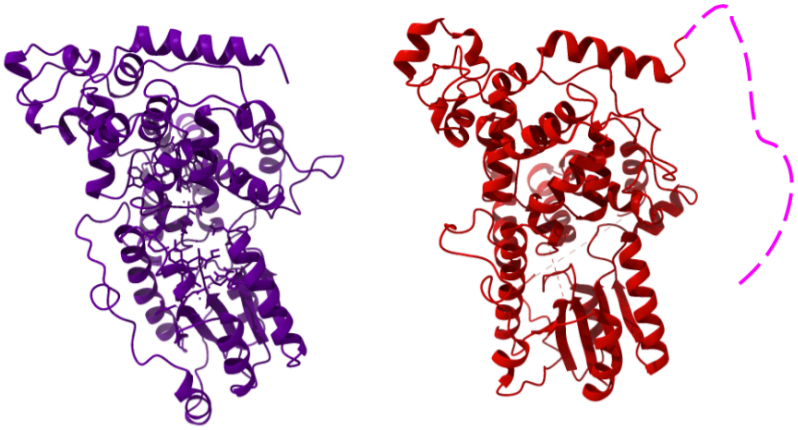


Figure 12: Cryptochromes are evolutionarily related to the DNA repair enzyme photolyase
The photolyase enzyme (PDB: 1QNF) and the CRY1 PHR (PDB: 5T5X) are largely structurally conserved. One major difference is the presence of a C-terminal disordered region in cryptochromes (pink dashed line).

3.3 Materials and Methods

3.3.1 Cloning the human CRY1 tail and the Δ11 variant for expression of recombinant protein in bacteria

To study the interaction of the CRY1 PHR and its C-terminus, I cloned the CRY1 C-terminus as an independent expression construct. The CRY1 PHR is a 56 kDa

protein whose expression and purification had been previously optimized by the Partch lab (Michael et al., 2017) . The human CRY1 C-terminal tail ("CT", residues 496-586) was cloned into a His-GST Parallel vector. After sequencing confirmation was received, I transformed the DNA into Rosetta (DE3) *E. coli* cells using the protocol described in Chapter 2. To further probe the importance of Exon 11 in this interaction, I used His-GST hCRY1 CT as the DNA template to delete Exon 11 using two-step PCR (Figure 13). In the first PCR reaction, I obtained two fragments: one spanning the DNA before the section to be deleted and the other fragment following this region. After purifying these two fragments, I did a second PCR that allowed for the deletion due to the engineered DNA overlap in these fragments. This new tail sequence, with a deletion of Exon 11 (CT Δ11) was then cloned into the same His-GST vector and verified by sequencing.

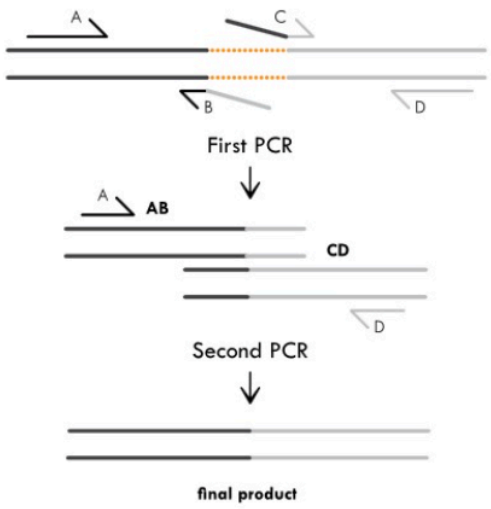


Figure 13: Schematic representation of a two-step PCR deletion
PCR deletion allows the removal of a fragment in the DNA sequence (dashed yellow line). In the first round of PCR, two fragments are created. The second PCR allows for the two fragments to combine and removal of the desired DNA sequence.

3.3.2 Protein Expression and Purification

The same protocol from Chapter 2 for expression and purification was used to express and purify both hCRY1 CT and hCRY1 CT Δ11. The buffer used for the final polishing step in protein purification on the Superdex 75 prep HiLoad 16/60 column was 50 mM Tris, 150 mM NaCl, pH 7.5.

3.4 Sortase A Peptide Ligation Reaction

3.4.1 His-Sortase A Protein Expression

The main objective here was to use the *S. aureus* Sortase A enzyme to chemically ligate a fluorophore on the N-terminus of my CRY1 tail peptides to generate a labeled probe for fluorescence polarization binding assays (Theile et al., 2013). The *E. coli* codon-optimized His-Sortase A construct was obtained from the B. Zoltowski lab. The His-Sortase was expressed in *E. coli* BL21 (DE3) cells following standard procedures outlined in Chapter 2. Protein expression was induced with 0.5 mM IPTG at an optical density (OD_{600}) of 0.6 and grown overnight at 18 °C degrees at 200 RPM.

3.4.2 His-Sortase A Protein Purification

I used buffer A (50 mM Tris, 150 mM NaCl, 20 mM Imidazole, pH 7.5) to resuspend cell pellets and mechanically lyse the cells. I used the Ni-NTA Agarose affinity chromatography (Qiagen) to bind my protein as described in Chapter 2. After washing the resin extensively with buffer A and checking by Bradford for protein during the final wash, I eluted the purified His-Sortase A by increasing imidazole with the buffer 50 mM Tris, 150 mM NaCl, 300 mM Imidazole, pH 7.5. After running an SDS-PAGE gel and staining with Coomassie Brilliant Blue Dye, I concentrated my elutions that contained the purified protein and loaded them onto the Superdex 75 prep HiLoad 16/60 column to separate His-Sortase A (~20 kDa) from contaminating proteins using the Sortase Reaction buffer (50 mM Tris, 150 mM NaCl, pH 7.5). Following the size exclusion run, I ran a 10% SDS-PAGE gel to visualize the purification and determine fractions where His-Sortase A is present. I then proceeded to concentrate the protein to about ~100 μM because that is the desired concentration for Sortase A labeling reaction. Next, I made 100 μL aliquots and used liquid nitrogen to quick freeze before storing small aliquots of the protein at -80 °C.

3.4.3 Sortase A-mediated protein ligation

I used this reaction to add a fluorophore to the purified hCRY1 CT so I could monitor its interaction with CRY1 PHR using a fluorescence polarization-based binding assay. The Sortase A-mediated reaction requires a glycine residue at the

N-terminus of the target protein (Theile et al., 2013). Since our recombinant proteins have a glycine on the N-terminus after TEV cleavage, we can easily label any protein we purify with a fluorophore via a Sortase A labeling reaction. The Sortase A enzyme also requires an LPXTG peptide recognition sequence on the labeling probe for recognition and ligation to the N-terminal glycine residue (Figure 14). It is important that modifications made to a protein do not disrupt it's structure or function; we find that targeting the free N-terminus is often the best option (Theile et al., 2013). After some optimization, my graduate student mentor Gian Carlo Parico and I determined that the ideal concentrations for the Sortase A reaction to label the CRY C-terminal tail are: 50 μM protein, 5 μM His-Sortase A enzyme, 150 μM FAM probe peptide, 10 mM $CaCl_2$ at 4°C overnight. We then proceeded to separate our FAM labeled CRY1 C-terminus from Sortase by doing a nickel clean up because Sortase has affinity to nickel resin and allow us to separate the protein from Sortase.To visualize the protein that was fluorescently labeled, we ran our samples on an 18% SDS-PAGE gel and scanned the gels on a Typhoon to imager. After confirmation by the gel, we quantified the percentage of actual protein labeled by Sortase using UV-VIS spectrometry.

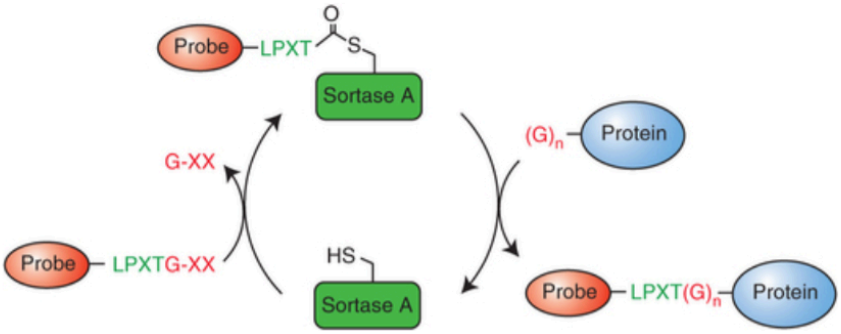


Figure 14: An overview of the Sortase-A peptide ligation reaction
Schematic of the enzymatic reaction to chemically ligate a fluorophore on the N-terminus of a protein. Adapted from Theile et al., 2013.

3.5 Confirmation of CRY1 interaction with the disordered C-Terminus

Fluorescence polarization (FP) assays allow us to study protein-protein interactions at the molecular level and obtain the K_d (equilibrium dissociation constant) from the two interacting molecules. The K_d gives us insight into the binding affinity of the proteins; the lower the K_d, the higher the affinity is between the two molecules. The fluorescence polarization signal increases as the molecular weight increases, so we are able to monitor the interaction of the labeled FAM-CRY1 CT (10 kDa) with the CY1 PHR (50 kDa).

The CRY1 PHR concentration was varied in onto an assay mixture containing 20 nM FAM-CRY1 CT using an opaque 384-well plate. The 384-well plate was then place on the plate EnVision, Perkin Elmer plate reader. In order to keep conditions constant, the same buffer was used for all the FP assays performed (50 mM Bis-Tris Propane, 100 mM NaCl, 2 mM TCEP, 0.05% Tween 20, pH 7.5).

The concentration- dependent change in FP signal observed upon titration of CRY1 PHR protein demonstrates that there is a direct interaction between the CRY1 PHR and its disordered C-terminal tail (CT). The moderate affinity of the PHR-tail interaction observed here is reasonable given that these two protein fragments are normally part of one protein. Based on the biochemical analysis of the CRY1 Δ11 allele in Patke et al., it appears as though Exon 11 may act as an autoinhibitory module that controls how tightly CRY1 binds to CLOCK:BMAL1. Therefore, I expected that deletion of Exon 11 would decrease affinity of the tail for the CRY1 PHR. As expected, I found that a tail lacking Exon 11 (Δ11 CT) has a four-fold higher K_d than the wild-type, denoting a weaker interaction with the PHR and highlighting the importance of Exon 11 for the interaction between the CRY1 PHR and its disordered C-terminal tail (Figure 15). Previous studies in our lab showed that a change of CRY1 affinity of this magnitude for CLOCK:BMAL1 was enough to change the period of clock by 1-2 hours (Xu et al., 2015), the same magnitude observed in people with the CRY1 Δ11 variant (Patke et al., 2017). Therefore, my data are beginning to shed light on a prevalent variation in humans that encodes circadian timing and sleep onset.

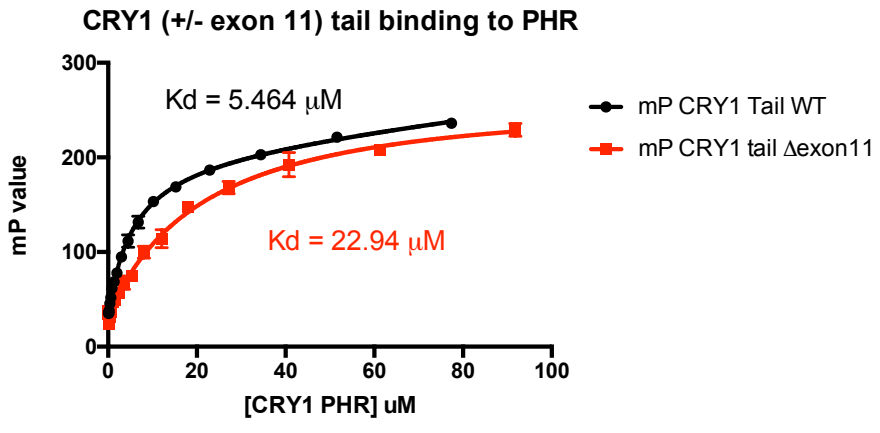


Figure 15: Fluorescent Polarization Assay
The fluorescent polarization assay allowed us to determine the binding affinities of hCRY1 WT and hCRY1 CT Δ11 to the CRY1 PHR. hCRY1 Tail WT binds with a lower affinity (K_d: 5.464 μM) than hCRY1 CT Δ11 (K_d: 22.94 μM) indicating Exon 11 might be important for tail binding to CRY1 PHR.

3.6 Discussion

Through a series of biochemical assays with purified proteins, I have demonstrated the direct interaction between the CRY1 PHR and its disordered C-terminal tail. These data also suggest that Exon 11 is needed to fully stabilize the interaction between the two regions of the protein. Based on these data from my thesis project, we hypothesize that Exon 11 might compete for binding at one or both of the two CLOCK:BMAL1 binding sites on CRY1. The deletion of CRY1 Exon 11 in patients diagnosed with DSPS might cause CRY1 to bind tighter to CLOCK:BMAL1 since Exon 11 is not present to compete for binding. To further probe this interaction, I am preparing set up competition assays to demonstrate whether the normal tail or the Δ11 variant tail are able to displace CLOCK or BMAL1 from CRY1 PHR. My graduate student mentor Gian Carlo Parico has

been optimizing these binding assays and we are now ready to examine how the presence of Exon 11 in the tail regulates CLOCK:BMAL1 binding to CRY1.

It is interesting that the Δ11 tail is still able to bind to the PHR, albeit with a weaker affinity. This implies that at least one other region of the tail can bind to the CRY1 PHR, presumably to mediate some other function of CRY1. Recently, I worked with Carlo to clone different regions of the CRY1 C-terminus to identify this other minimal binding region needed for the C-terminus to bind to the CRY1 PHR. These proteins will be purified and fluorescently labeled as described above to perform fluorescent polarization binding assays with the CRY1 PHR. Mapping the general architecture of the CRY1 PHR/tail interaction will provide insight into the functionality of this intrinsically disordered region and help us understand how the protein.

Chapter 4-Conclusion

The main goal for my independent senior thesis research project was to develop a deeper insight how circadian rhythms are regulated at the molecular level. Our lab leverages the genetic and cellular data obtained by others in the field to probe the detailed molecular mechanisms of circadian regulation. My first goal in the Partch Lab was to explore the essential role of CLOCK Exon 19 in circadian rhythms. Originally isolated from a forward genetic screen in mice, we know that this 51 amino acid region in the disordered C-terminus of CLOCK is required to

generate circadian rhythms, but after two decades of study, we don't understand its biochemical role yet. I set out to crystallize CLOCK Exon 19 bound to its repressor CIPC. Even though a different lab published the structure before me (Xu et al., 2015), I've obtained meaningful solution biophysical data on Exon 19, which was confirmed by the published structure. I used this structure to design mutations to disrupt the dimer and performed site directed mutagenesis, SEC-MALS, and CD on a first round of Exon 19 mutants. Identifying a mutation that disrupts the dimer is only the first step to probe the positive cooperative regulation of DNA binding that might be mediated by Exon 19. The future goal of this project is to find a set of mutations that disrupt the dimer and test for defects in DNA binding and circadian rhythms due to the lack of dimerization *in vivo*. The second goal of my thesis project was to understand the role of a naturally occurring human genetic variation that led to alternate splicing of the CRY1 C-terminus, thereby regulating the interaction between CRY1 and CLOCK:BMAL1. I have determined Exon 11 plays an important role in maintaining a high affinity interaction between the CRY1 PHR and its disordered C-Terminal tail. Interestingly, my work also identified another potential PHR binding site. In the future, we will further probe this by truncating the C-terminus tail into smaller pieces to find the minimal region needed to bind the CRY1 PHR.

Overall, these studies have helped to address the molecular basis for some genetic variants (*Clock*$^{\Delta 19}$ and *Cry1*$^{\Delta 11}$) that have a powerful effect on the clock. The field of chronobiology only recently entered its molecular era as the core

clock genes were identified in mammals just under two decades ago. Therefore, we are still working to understand how circadian rhythms are regulated at a molecular level. Understanding how ~24-hour timekeeping is generated by circadian rhythms to control human behavior and physiology has the potential to open new therapeutic strategies to control not our sleep/wake cycle but many aspects of our physiology like metabolism, hormone production and DNA repair.

5 Bibliography

Bell-Pedersen, D., Cassone, V.M., Earnest, D.J., Golden, S.S., Hardin, P.E., Thomas, T.L., and Zoran, M.J. (2005). Circadian rhythms from multiple oscillators: lessons from diverse organisms. Nat. Rev. Genet. *6*, 544–556.

Dyson, H.J., and Wright, P.E. (2016). Role of Intrinsic Protein Disorder in the Function and Interactions of the Transcriptional Coactivators CREB-Binding Protein (CBP) and p300. J. Biol. Chem. jbc.R115.692020.

Gustafson, C.L., and Partch, C.L. (2015). Emerging Models for the Molecular Basis of Mammalian Circadian Timing. Biochemistry (Mosc.) *54*, 134–149.

Hou, Z., Su, L., Pei, J., Grishin, N.V., and Zhang, H. Crystal Structure of the CLOCK Transactivation Domain Exon19 in Complex with a Repressor. Structure *25*, 1187–1194.e3.

Katada, S., and Sassone-Corsi, P. (2010). The histone methyltransferase MLL1 permits the oscillation of circadian gene expression. *17*, 1414.

Khan, S.K., Xu, H., Ukai-Tadenuma, M., Burton, B., Wang, Y., Ueda, H.R., and Liu, A.C. (2012). Identification of a Novel Cryptochrome Differentiating Domain Required for Feedback Repression in Circadian Clock Function. J. Biol. Chem. *287*, 25917–25926.

King, D.P., Zhao, Y., Sangoram, A.M., Wilsbacher, L.D., Tanaka, M., Antoch, M.P., Steeves, T.D., Vitaterna, M.H., Kornhauser, J.M., Lowrey, P.L., et al. (1997). Positional cloning of the mouse circadian clock gene. Cell *89*, 641–653.

Koike, N., Yoo, S.-H., Huang, H.-C., Kumar, V., Lee, C., Kim, T.-K., and Takahashi, J.S. (2012). Transcriptional architecture and chromatin landscape of the core circadian clock in mammals. Science *338*, 349–354.

van der Lee, R., Buljan, M., Lang, B., Weatheritt, R.J., Daughdrill, G.W., Dunker, A.K., Fuxreiter, M., Gough, J., Gsponer, J., Jones, D.T., et al. (2014). Classification of Intrinsically Disordered Regions and Proteins. Chem. Rev. *114*, 6589–6631.

Michael, A.K., Fribourgh, J.L., Chelliah, Y., Sandate, C.R., Hura, G.L., Schneidman-Duhovny, D., Tripathi, S.M., Takahashi, J.S., and Partch, C.L. (2017). Formation of a repressive complex in the mammalian circadian clock is mediated by the secondary pocket of CRY1. Proc. Natl. Acad. Sci. *114*, 1560–1565.

Patke, A., Murphy, P.J., Onat, O.E., Krieger, A.C., Özçelik, T., Campbell, S.S., and Young, M.W. (2017). Mutation of the Human Circadian Clock Gene CRY1 in Familial Delayed Sleep Phase Disorder. Cell *169*, 203–215.e13.

Pittendrigh, C.S., and Minis, D.H. (1964). The Entrainment of Circadian Oscillations by Light and Their Role as Photoperiodic Clocks. Am. Nat. *98*, 261–294.

Ripperger, J.A., and Schibler, U. (2006). Rhythmic CLOCK-BMAL1 binding to multiple E-box motifs drives circadian *Dbp* transcription and chromatin transitions. Nat. Genet. *38*, 369–374.

Sancar, A. (2004). Regulation of the mammalian circadian clock by cryptochrome. J. Biol. Chem. *279*, 34079–34082.

Shi, G., Wu, D., Ptáček, L.J., and Fu, Y.-H. (2017). Human genetics and sleep behavior. Curr. Opin. Neurobiol. *44*, 43–49.

Shimomura, K., Kumar, V., Koike, N., Kim, T.-K., Chong, J., Buhr, E.D., Whiteley, A.R., Low, S.S., Omura, C., Fenner, D., et al. (2013). Usf1, a suppressor of the circadian Clock mutant, reveals the nature of the DNA-binding of the CLOCK:BMAL1 complex in mice. eLife *2*, e00426.

Theile, C.S., Witte, M.D., Blom, A.E.M., Kundrat, L., Ploegh, H.L., and Guimaraes, C.P. (2013). Site-specific N-terminal labeling of proteins using sortase-mediated reactions. Nat Protoc. *8*, 1800–1807.

Tóth-Petróczy, Á., Oldfield, C.J., Simon, I., Takagi, Y., Dunker, A.K., Uversky, V.N., and Fuxreiter, M. (2008). Malleable Machines in Transcription Regulation: The Mediator Complex. PLOS Comput. Biol. *4*, e1000243.

Wright, P.E., and Dyson, H.J. (2015). Intrinsically disordered proteins in cellular signalling and regulation. Nat. Rev. Mol. Cell Biol. *16*, 18–29.

Xu, H., Gustafson, C.L., Sammons, P.J., Khan, S.K., Parsley, N.C., Ramanathan, C., Lee, H.-W., Liu, A.C., and Partch, C.L. (2015). Cryptochrome 1 regulates the circadian clock through dynamic interactions with the BMAL1 C-terminus. Nat. Struct. Mol. Biol. *22*, 476–484.

Zhang, R., Lahens, N.F., Ballance, H.I., Hughes, M.E., and Hogenesch, J.B. (2014). A circadian gene expression atlas in mammals: implications for biology and medicine. Proc. Natl. Acad. Sci. U. S. A. *111*, 16219–16224.

Zhao, W.-N., Malinin, N., Yang, F.-C., Staknis, D., Gekakis, N., Maier, B., Reischl, S., Kramer, A., and Weitz, C.J. (2007). CIPC is a mammalian circadian clock protein without invertebrate homologues. Nat Cell Biol *9*, 268–275.

6 Appendices

GST Buffer A	50 mM Tris, 150 mM NaCl, 2 mM DTT, 5% (v/v) Glycerol, pH 7.5
His Buffer A	50 mM Tris, 150 mM NaCl, 20 mM Imidazole, 5 mM BME, pH 7.5
His Buffer B	50 mM Tris, 150 mM NaCl, 300 mM Imidazole, 5 mM BME, pH 7.5
S75 Buffer	20 mM HEPES, 12 5mM NaCl, 2 mM DTT, pH 7.5
Fluorescence Polarization Buffer	50 mM Bis-Tris Propane, 100 mM NaCl, 2 mM TCEP, 0.05% (v/v) Tween 20, pH 7.5
Circular Dichroism Buffer	20 mM Sodium Phosphate, 100 mM NaCl, pH 7.5
Sortase Reaction Buffer	50 mM Tris, 150 mM NaCl, pH 7.5
CLOCK Exon 19 Q545A Primers DNA Template: CLOCK Exon 19 WT	Forward Primer: AATATTCATCGGCAGGCAGAAGAA Tm: 62 °C Reverse Primer: TAGTTCTTCTGCCTGCCGATG Tm: 61.2 °C
CLOCK Exon 19 Q552A Primers DNA Template: CLOCK Exon 19 WT	Forward Primer: AACTAAGGAAAATTGCAGAGCAACTT Tm: 61.7 °C Reverse Primer: CTGAAGTTGCTCTGCAATTTTCCT Tm: 62.0 °C
CLOCK Exon 19 Q545A/Q552A Primers DNA Template: CLOCK Exon 19 Q545A	Forward Primer: AACTAAGGAAAATTGCAGAGCAACTT Tm: 61.7 °C Reverse Primer: CTGAAGTTGCTCTGCAATTTTCCT Tm: 62.0 °C
hCRY1 CT WT Primers DNA Template:	Forward Primer: Linker A: CGCGAATTCGGACTAGGTCTTCTGGCATCAGTA Tm: 75.4 °C

	Reverse Primer: CCAGGTTGTAGCAGCAGTGGAAGTGGAAGAAGAAGCT CC ATGGGCACT Tm: 84 °C
hCRY1CT Δ11 Primers DNA Template: HisGST hCRY1 CT WT	1st Step: Forward: Linker A: CGCGAATTCGGACTAGGTCTTCTGGCATCAGTA Tm: 75.4°C Linker B: CCAGGTTGTAGCAGCAGTGGAAGTGGAAGAAGAAGCT C CATGGGCACT Tm: 84 °C Reverse: Linker A: AGTGCCCATGGAGCTTCTTCCACTTCCACTGCTGCTAC AACCTGG Tm:83.6 °C Linker B: ATAAGAATGCGGCCGCTTAATTAGTGCTCTGTCTCTGG A CT TTAGG Tm: 79.9 °C 2nd Step: Forward: CGCGAATTCGGACTAGGTCTTCTGGCATCAGTA Tm: 75.4 °C Reverse: ATAAGAATGCGGCCGCTTAATTAGTGCTCTGTCTCTGG A CTT TAGG Tm: 79.9 °C
hCRY1 CT_N_11 Primers	Forward: CGCGAATTCGGACTAGGTCTTCTGGCATCAGTA

DNA Template: HisGST hCRY1 CT WT	Tm: 75.4 °C Reverse: ATAAGAATGCGGCCGCTTATTGCTTCAACAGGTGAGTT TGCTG ACT Tm: 79.9 °C
hCRY1 CT_11_C Primers DNA Template: HisGST hCRY1 CT WT	Forward: CGCGAATTCTGCTCTCAAGGGAGTGGTATTTTACAC Tm: 75.4 °C Reverse: ATAAGAATGCGGCCGCTTAATTAGTGCTCTGTCTCTGG ACTT TAGG Tm: 79.9 °C
hCRY1 CT_C Primers DNA Template: HisGST hCRY1 CT WT	Forward: CGCGAATTCGGAAGAAGCTCCATGGGCACTGGT Tm: 77.8 °C Reverse: ATAAGAATGCGGCCGCTTAATTAGTGCTCTGTCTCTGG AC TTTAGG Tm: 79.9 °C
hCRY1 CT_N Primers DNA Template: HisGST hCRY1 CT WT	Forward: CGCGAATTCGGACTAGGTCTTCTGGCATCAGTA Tm: 75.4 °C Reverse: ATAAGAATGCGGCCGCTTAACTTCCACTGCTGCTACAA CCT GG Tm: 80.7 °C