# Molecular mechanisms of temperature sensing in plants

Aditya Nayak

Communauté
UNIVERSITÉ Grenoble Alpes

# THÈSE

Pour obtenir le grade de

## DOCTEUR DE LA COMMUNAUTÉ UNIVERSITÉ GRENOBLE ALPES

Spécialité : Biologie Végétale

Arrêté ministériel : 25 mai 2016

Présentée par


## Aditya NAYAK


Thèse dirigée par **Chloé ZUBIETA**, cnrs

préparée au sein du **Laboratoire Laboratoire de Physiologie Cellulaire Végétale**
dans **l'École Doctorale Chimie et Sciences du Vivant**

# Mécanismes moléculaires de la perception de la température ambiante chez les plantes

# Molecular mechanisms of temperature sensing in plants

Thèse soutenue publiquement le **28 mars 2019**,
devant le jury composé de :

**Madame CHLOE ZUBIETA**
CHARGE DE RECHERCHE, CEA GRENOBLE, Directeur de thèse
**Monsieur GREGORY VERT**
DIRECTEUR DE RECHERCHE, CNRS DELEGATION OCCITANIE OUEST, Rapporteur
**Monsieur RICHARD IMMINK**
PROFESSEUR, UNIVERSITE DE WAGENINGEN - PAYS-BAS, Rapporteur
**Monsieur MARC JAMIN**
PROFESSEUR, UNIVERSITE GRENOBLE ALPES, Président
**Monsieur CYRIL BOYAULT**
CHARGE DE RECHERCHE, CNRS DELEGATION ALPES, Examinateur
**Monsieur PHILIP WIGGE**
DIRECTEUR DE RECHERCHE, UNIVERSITE DE CAMBRIDGE - ROYAUME-UNI, Examinateur

Communauté
UNIVERSITÉ Grenoble Alpes

**THÈSE**

Pour obtenir le grade de

**DOCTEUR DE LA COMMUNAUTE UNIVERSITE GRENOBLE ALPES**

Spécialité : Chimie & Science du Vivant
Arrêté ministériel : 25 mai 2016

Présentée par
**Aditya NAYAK**

Thèse dirigée par **Chloe Zubieta, CR1**
Laboratoire Physiologie Cellulaire & Végétale
Institut de Biosciences et Biotechnologies de Grenoble
CEA-Grenoble

préparée au sein du **LPCV**
dans **l'École Doctoral Chimie et Science du Vivant**

# Mécanismes Moléculaires De La Perception De La Température Ambiante Chez Les Plantes

**Thèse soutenue publiquement le 28th Mars, 2019**
devant le jury composé de :

**Marc JAMIN**
Professeur, IBS/UGA, Président

**Gregory VERT**
Directeur de Recherche, CNRS/ Université de Toulouse III, Rapporteur

**Richard IMMINK**
Professeur, Wageningen University and Research, Rapporter

**Cyril, BOYAULT**
Charge de Recherche, IAB Grenoble, Membre

**Philip, WIGGIE**
Professeur, University of Potsdam, Membre

Communauté
UNIVERSITÉ Grenoble Alpes

# Molecular Mechanisms Of Temperature Sensing In Plants

By
**Aditya P. Nayak**

A DISSERTATION

Submitted to
The Doctoral School of Chemistry and Life Sciences, University of Grenoble
In partial fulfillment of the requirements

Presented to :

**Dr. Chloe, Zubieta**, Thesis Supervisor
LPCV, BIG, CEA-Grenoble
*University of Grenoble Alpes*

**Dr. Gregory Vert**
LRSV, CNRS
*University of Toulouse III*

**Prof. Marc Jamin,**
Institute for Structural Biology
*University of Grenoble Alpes*

**Dr. Cyril Boyault**
Institute for Advanced Bioscienes
*University of Grenoble Alpes*

**Prof. Richard Immink**
Department of Plant Sciences
*Wageningen University & Research*

**Prof. Philip Wiggie**
Department of Plant Adaptation
*University of Potsdam*

To be defended on: 28th March 2019

**Résume**

Le Complex Evening (EC, Evening complex) est composé de trois protéines : EARLY FLOWERING 3 (ELF3), ELF4 et LUX ARRYTHMO (LUX). Ce complexe est un composant clé de l'horloge circadienne de la plante et un important régulateur de gènes impliqués dans la croissance de la plante en réponse à la température, comme *PHYTOCHROME INTERACTING FACTEUR 4 (PIF4) notamment*. Des études ont montré que l'activité de ce complexe dépend de la température, avec une activité répressive qui augmente à des températures plus basses. Pour autant les mécanismes moléculaires impliqués dans la formation du complexe EC, dans sa liaison à l'ADN et son activité thermosensible étaient encore très mal compris. Une série d'expériences structurales in vitro, et in planta ont été réalisées afin de mieux comprendre l'ensemble de ces mécanismes moléculaires.

Pour cela, les trois protéines recombinantes du « EC » ont été produites et purifiées jusqu'à homogénéité. Ces protéines ont été utilisées pour reconstituer le complexe in vitro et pour étudier son activité de liaison à l'ADN. Le rôle des trois protéines dans la formation et l'activité du complexe ont été déterminé. LUX est nécessaire à la liaison à l'ADN via son domaine MYB , et cible le EC vers ses sites de liaison. ELF3 agit comme un échafaudage pour la formation du EC en liant à la fois LUX et ELF3. Cependant, le complexe LUX-ELF3 ne se lie pas avec une haute affinité à l'ADN. ELF4 est nécessaire pour rétablir la liaison de l'ADN au complexe. Pour explorer davantage les déterminants de la spécificité de liaison à l'ADN, le domaine de liaison à l'ADN (DBD) de LUX a été exprimé, purifié et cristallisé. La structure cristalline de LUX DBD en complexe avec deux oligonucléotides d'ADN différents, révèle les acides aminés critiques pour la liaison à l'ADN. La majorité de ces résidus entrent en contact avec le sillon principal de l'ADN et font partie d'un motif spécifique à la plante SH (A / L) QK (F / Y). De plus, un résidu d'arginine (Arg146) dans la région flexible N-terminale de la protéine joue un rôle important avec des contacts dans le sillon mineur de l'ADN. Sur la base de ces études structurales, une mutation de l'arginine en alanine (R146A) a été réalisée. Cette mutation diminue l'affinité de liaison à l'ADN mais conserve la spécificité déterminée in vitro par des expériences de gels de retard en comparaison avec la protéine de type sauvage. Des expériences transgéniques ont été utilisées pour déterminer l'effet de la mutation R146A chez la plante. Comme prévu, cette mutation a abouti à un phénotype intermédiaire entre le type sauvage et un mutant « *knock out* » du EC. Ceci suggère qu'en modifiant l'affinité de LUX pour sa liaison à l'ADN, l'activité de la totalité du EC peut être ajustée dans la plante. L'expression de *PIF4* chez le mutant R146A est plus élevée que chez le sauvage mais moins affectés que chez le mutant *lux*, ce qui confirme la diminution de l'activité répressive du EC par la mutation R146A de LUX.

Pour explorer plus en profondeur la régulation du gène *PIF4*, la technologie CRISPR-Cas9 a été utilisée pour cibler différents éléments *cis* de son promoteur, notamment le site de liaison de LUX (LBS) et un élément qui s'appelle une « G-box » . Ces mutations ont entraîné des effets opposés sur la croissance des plantes et leur réponse à la température. Le mutant dans le LBS présente des hypocotyles allongés et un phénotype de floraison précoce à 22 ° C par rapport au sauvage, alors que le mutant de « G-box » entraine un raccourcissement des hypocotyles et une floraison tardive à 27 ° C par rapport au sauvage. Cela suggère qu'une modification des éléments *cis* dans le promoteur de *PIF4* pourrait être un moyen de reprogrammer la croissance des plantes et leur réponse à la température.

Ces résultats fournissent différentes stratégies pour influer sur la croissance des plantes sous différents régimes de température ambiante sans contrainte de stress, que ce soit par l'ingénierie des protéines basée sur la structure, comme indiqué pour la mutation LUX R146A ou par l'édition génomique d'éléments de régulation spécifiques connus pour affecter la croissance et la réponse à la température tels que LBS et G-box dans le promoteur de *PIF4*. Avec l’augmentation des températures due au changement climatique et à ses effets néfastes sur la productivité des plantes, la capacité de modifier de manière prévisible la croissance des plantes et leur réponse à la température constituent un moyen intéressant de relever ce défi mondial. Ces résultats constituent une base potentielle pour de futures applications en bio-ingénierie d'espèces cultivées.

**Abstract**

The Evening Complex (EC), a three protein complex comprising EARLY FLOWERING 3 (ELF3), ELF4 and LUX ARRYTHMO (LUX), is a key component of the plant circadian clock and an important regulator of genes important for thermosensitive growth, including *PHYTOCHROME INTERACTING FACTOR 4 (PIF4)*. Studies have shown that EC activity is temperature dependent, with increased repressive activity at lower temperatures. However, the molecular mechanisms for EC complex formation, DNA-binding and thermosensitive activity were not known. In order to address this, a series of in vitro, structural and in planta experiments were performed.

All three proteins of the EC were recombinantly produced and purified to homogeneity. These proteins were used to reconstitute the EC in vitro and to study its DNA-binding activity. The role of all three proteins in complex formation and activity were determined. LUX acts as the driver of DNA-binding via its MYB DNA-binding domain and targets the EC to its cognate sites. ELF3 acts as a scaffold for EC formation by binding both LUX and ELF3. However, the LUX-ELF3 complex poorly binds to DNA. ELF4 is required to restore DNA-binding of the complex. To further explore the DNA-binding specificity determinants, the DNA-binding domain (DBD) of LUX was expressed, purified and crystallized. The crystal structure of the LUX DBD in complex with two different DNA oligonucleotides reveals the residues critical for base read-out. The majority of these residues contact the major groove and are part of a plant-specific signature motif SH(A/L)QK(F/Y). In addition, an arginine residue (Arg 146) in the flexible N-terminal region of the protein acts as a clamp with contacts in the minor groove. Based on these structural studies, an arginine to alanine mutation (R146A) was made which had decreased DNA-binding affinity but retained specificity as determined in vitro via band shift assays as compared to the wild type protein. Transgenic experiments were used to determine the effect of the R146A mutation in planta. As predicted, this mutation resulted in a phenotype between wild type and an EC knock out mutant. This suggests that by altering the DNA binding affinity of LUX, the activity of the entire EC can be tuned in the plant. PIF4 expression levels were measured in the mutant and were shown to be elevated with respect to wild type but less affected than in a *lux* mutant, further supporting the decreased repressive activity of the EC due to the R146A mutation in LUX.

To further explore *PIF4* regulation, CRISPR-Cas9 was used to target different cis-elements in the *PIF4* promoter including the LUX Binding Site (LBS) and a G-box element. These mutations had opposite effects on plant growth and thermoresponse with the LBS mutant exhibiting elongated hypocotyls and an early flowering phenotype at 22°C as compared to wt. The G-box mutant on the contrary exhibited shortened hypocotyls and a late flowering phenotype at 27°C as compared to wt. This suggests that altering cis-elements in the *PIF4* promoter may be a way to reprogram plant growth and thermoresponse at different temperatures.

Taken together, these results provide different strategies to affect plant growth under different non-stress ambient temperature regimes through either structure-based protein engineering as shown for LUX R146A mutation or via genome editing of specific regulatory elements known to affect growth and thermoresponse such as the LBS and G-box in the *PIF4* promoter. With the increase in global

temperatures due to climate change and the deleterious effects this has on crop productivity, the ability to predictably alter plant growth and thermoresponse is an attractive way to address this global challenge. These results provide a potential foundation for future applications in bioengineering of important crop species.

# ACKNOWLEDGEMENTS

I am thankful to be a part of the LPCV team at CEA Grenoble. It is here that I truly felt the sense of community and collaboration among scientists. I would like to thank my supervisor Dr. Chloe Zubieta for providing me with this wonderful and exciting project and for being the most wonderful PI that every graduate student wished they had. Thanks for correcting all my writing, be it my thesis, project reports, motivation letters for fellowship and presentation. You taught me to be critical about data, repeat experiments till there is conformity in data, nurtured my scientific thinking and ability to come up with new ideas in order to arrive at conclusions. These are the things that I really enjoyed learning while being at your lab. Thanks for being supportive for every experiment that I came to you with. I really enjoyed the independence that you provided me as graduate student.

I would also like to thank Dr. Catarina Silva who was by my side from day 1 at the lab to teach me the basics of protein expression and purification. If it wouldn't be for you, I would still not be confident about working on the Acta or being precise about buffers. It would have been much difficult for me specially coming from a molecular plant biology background, without much experience in biochemistry or structural biology, to take up a project like this; if it was not for your constant guidance I would be lost a bit.

I would like to thank Emmanuel Thévenon for all the help and good vibes at the work bench, thanks for being such a wonderful person to have as a lab mate who motivated me to do a lot of sports, hikes, trails and experiments together and specially for all the assistance for doing the insect cell culture. I would like to thank Dr. Véronique Hugouvieux , Agnés Jourdain for helping me doing a lot of trouble shoot with the plant experiments, seting up the growth chambers and for helping me a lot in the lab. I would like to thank Dr. Xulei Lai for his help in the experiment for doing CD and structural characterizations. I would like to thank Dr. François Parcy and Dr. Renaud Dumas for his critical inputs into my results and helping me design experiments and present data in more meaningful ways. I would like to thank Dr. Dimitris Petroutsos for coming to all my annual thesis advisory committee meetings and for being a constant source of motivation.

I would also like to thank all my friends at CEA, specially Nelson, Elodie, Claudius, Sophie, Tiffanny, Christian, César, Raquel, Eugenia, Pauline, Xenia, Valeria and Serena for being such nice friends. Thanks for all the ski trips, weekend parties and wonderful advices when I was confused.

Last but not the least I am grateful to my family. I am fortunate to have parents who were constantly kept me motivated to pursue science as a career. Thanks mom and dad, if it wasn't for your support and encouragement right from school level to participate in science fairs, science congress and debates, I would not be here where I am today. Also I would like to thank my aunt and teacher Aparna Mohapatra who has motivated me to pursue science right from high-school.

# TABLE OF CONTENTS

# LIST OF TABLES

# LIST OF FIGURES

# CHAPTER 1. **INTRODUCTION**

## 1.1 Circadian Clocks

Over the course of each day, many organisms ranging from cyanobacteria to plants and animals, undergo rhythmic changes in behavior, physiology and biochemistry. Rhythms that occur with a periodicity roughly matching the earth's rotation on its axis and that continue in the absence of external stimuli are termed circadian (Harmer et al., 2001). Circadian rhythms are controlled by an endogenous biochemical oscillator called the circadian clock. The ability to generate daily rhythms is a cellular quality. The cellular clocks form networks that build up the circadian programs in tissues, organs and the entire organism. The circadian clock allows organisms to anticipate rhythmic changes in their environment and accordingly modify their physiology to provide an adaptive advantage (Ouyang et al., 1998). Circadian clocks control a variety of functions across organisms. In cyanobacteria and photosynthetic prokaryotes circadian clock controls photosynthesis, nitrogen fixation and even cell division (Cohen and Golden, 2015; Mitsui et al., 1986). In mammals, the clock influences digestion, regulation of body temperature, hormone secretion and behavior such as time of sleep onset (Harmer et al., 2001).

Circadian rhythms are persistent in the absence of external cues and exhibit temperature compensation, meaning the period of their rhythmicity is roughly consistent even over a wide range of temperatures (Kidd et al., 2015). At the same time, however, circadian clocks are facilely entrained by diurnal oscillations in temperature, meaning that the clock oscillations can be reset by temperature fluctuations. These properties of compensation and entrainment allow the circadian clock to better match endogenous and environmental time. Environmental signals such as changes in light, temperature and nutrient availability act as entrainment cues and are known as Zeitgebers (German for "time givers") (Golombek and Rosenstein, 2010; Moore, 1997).

### The Central Oscillator

A basic circadian system needs to have three principal components: a biological clock, input pathways and output pathways. Each component plays a specific role; the biological clock acts as a central oscillator that generates rhythmicity; input pathways receive and relay environmental cues that entrain the oscillator; and output pathways control diurnal activities. However, this characterization is highly simplified because there are clock input components such as photoreceptors that act as clock outputs as well (Bognár et al., 1999; Emery et al., 1998). Furthermore, clock outputs can feed back on the clock itself, modulating the pace of the central oscillator (Lopez-Molina et al., 1997). Hence, a more comprehensive model is based on multiple overlapping feedback loops that contribute to timekeeping (Roenneberg and Merrow, 1998). However, the central oscillator paradigm has proven conceptually useful and is briefly described in the next section with examples from organisms belonging to different kingdoms of life.

### Types of Circadian Oscillators

In all circadian clocks investigated to date, it has been found that delayed negative feedback loops are at the heart of oscillatory processes. Circadian clocks can be classified as transcription-translation feedback loops (TTFLs) or post-translational oscillators (PTOs) (Hurley et al., 2016). A more detailed discussion on circadian clock based on TTFLs and PTOs loops have been described in previous reviews

but will be briefly summarized in the following sections (Dunlap, 1999; Huang, 2018; Hurley et al., 2016; Robinson and Reddy, 2014; Swan et al., 2018).

TTFLs attain oscillation through delayed negative feedback. Transcription factors that belong to clock input pathways induce the expression of clock genes that then act to negatively regulate each other at the transcriptional level. This creates oscillating patterns of gene expression. In principle, only one component in a loop needs to be cycled in order to attain oscillations. However, in all clock feedback loops described to date, multiple negative elements and some of the positive elements that belong to the oscillator cycle at the mRNA and/or protein level (Brown et al., 2012; Hurley et al., 2016).

By contrast, PTOs utilize timekeeping mechanisms that are independent of transcription. Like TTFLs, PTOs involve a cycle of biochemical processes. However, unlike TTFLs, their behavior is controlled by posttranslational modifications, conformational changes, protein-protein interactions and/or subcellular localization instead of being driven by changes in mRNA expression levels. It is important to note that while there is a distinction between TTFLs and PTOs as general categories of negative feedback systems, many circadian clocks make use of both transcription/translation as well post-translational timing steps to constitute a biochemical oscillator.

Systems based on negative feedback loops need inbuilt delays otherwise the oscillations would soon dampen to constant levels of gene expression. Interestingly, at least one negative component in each feedback loop described below exhibits a significant delay between peak levels of mRNA and protein expression. This delay suggests that post-transcriptional regulation may be a general mechanism for proper clock control even in TTFL based clocks. The negative elements that generate delay in the central oscillator are often the first to be affected by environmental signals that reset the clock, and their phase determines the overall phase of the oscillator. The circadian clocks in cyanobacteria, *Drosophila melanogaster* and mammals will be discussed in the next sections, followed by a detailed description of the plant circadian system.

**The Circadian clock in Cyanobacteria**

Originally, it was assumed that prokaryotes would not have circadian rhythms. However work in the late 1980s and early 1990s showed that in fact cyanobacteria do have circadian rhythms (reviewed in Golden et al., 1997). In photosynthetic bacteria, mutants with short-period, long-period and arrhythmic phenotypes were identified through forward genetic screens (Kondo et al., 1994). All these mutants had changes in one of three adjacent genes: *kaiA, kaiB,* or *kaiC*, (Kai means "cycle" in Japanese), with mutations in *kaiC* accounting for all three possible clock phenotypes (Ishiura et al., 1998).

The *kaiABC* gene cluster was originally categorized as a TTFL based circadian clock. However, subsequently it was shown that the products of this gene cluster can recapitulate circadian biochemical oscillations *in vitro* (Nakajima et al., 2005). This circadian oscillation persists *in vivo* even when transcription and translation are inhibited (Tomita et al., 2005). The ability to be reconstituted *in vitro* demonstrates that KaiABC can act as self-sustained PTO. The mechanism behind this PTO is based on auto-phosphorylation and de-phosphorylation of KaiC. KaiC is both an auto-kinase and auto-phosphatase. KaiA promotes the KaiC auto-kinase activity eventually leading to phosphorylation of

KaiC. KaiB negates the action of KaiA and promotes KaiC de-phosphorylation. This phosphorylation/ de-phosphorylation of KaiC oscillates over 24 hours determining the phase of the clock.

In addition to a PTO based circadian clock, cyanobacteria also exhibit characteristics of a TTFL circadian clock based on the rhythmic expression of the *kaiABC* genes, with mRNA levels peaking near the end of the day at ZT 9–12 (ZT, or circadian time, is used to indicate the subjective time of day in constant growth conditions. By convention, ZT-0 corresponds to subjective dawn (lights on) and ZT-12 corresponds to subjective dusk with 12 hours light and 12 hours dark period (lights off)). Overexpression of *kaiC* suppresses expression of both *kaiB* and *kaiC*, which share a single promoter and are co-transcribed. On the other hand, overexpression of *kaiA* enhances expression of *kaiBC*. The loss of *kaiA* leads to a decrease in expression of *kaiBC* (**Figure 1-1**) (Ishiura et al., 1998). Thus, both PTO and TTFL play roles in maintaining circadian clocks in cyanobacteria. A more detailed discussion of the cyanobacterial central oscillator can be found in Cohen and Golden (2015).

As we move up the evolutionary ladder, the TTFL based circadian clock systems are more prominent. In the next sections the TTFL based circadian systems in drosophila and mouse will be discussed.

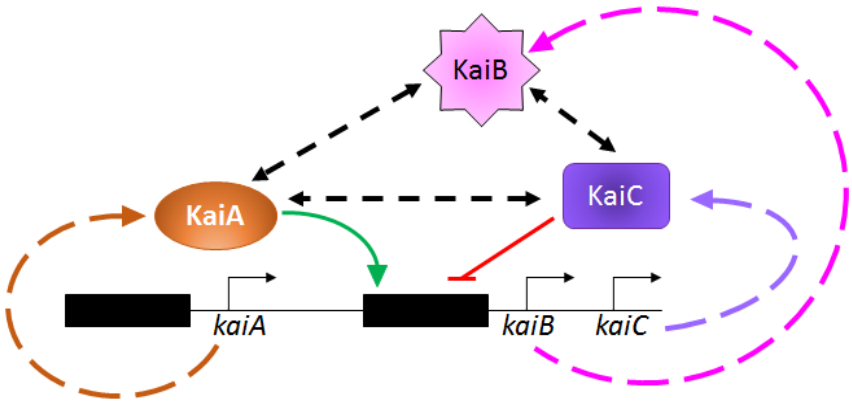



**Figure 1-1 Simplified schematics of feedback loops in cyanobacteria circadian clock.**

Colored dotted lines represent mRNAs and link genes with their respective proteins Protein-protein interactions are indicated with dotted black arrows. Green arrows indicate positive effect of a component on some process or promoter, and red lines indicate an inhibitory action. KaiA acts a positive component while KaiB and KaiC act as negative component in the feedback loops of cyanobacteria.

**The Circadian clock in Drosophila melanogaster.**

The molecular-genetic study of circadian rhythms was initiated in Drosophila. The Drosophila circadian clock is based on the TTFL model. The identification and eventual cloning of the *period* (*per*) gene was

based on its perturbation of fly eclosion (emergence of adults from their pupal cases) and activity rhythms (Konopka and Benzer, 1971; Reddy et al., 1984). The next fly clock gene to be cloned was *timeless* (*tim*). It was identified by its ability to bind to the PER protein and by positional cloning of a mutant gene responsible for altered activity rhythms (Gekakis et al., 1995; Myers et al., 1996). Both *per* and *tim* mRNA levels cycle, the peak expression occurring at the early night around ZT-14 (Hardin et al., 1990; Sehgal et al., 1995). Protein levels also cycle, but the protein expression peak is delayed by a phase relative to the mRNA peak. The protein expression peak for PER and TIM is around the middle of the night at ZT18 (Hunter-Ensor et al., 1996; Myers et al., 1996; Zeng et al., 1996). PER interacts with TIM via a PAS (Per-Arnt-Sim) protein-protein interaction domain, but there are no other obvious functional domains in PER or TIM. The PER-TIM heterodimer is nuclear localized and inhibits the expression of *per* and *tim*, resulting in a negative feedback loop. However, the mechanism of this repression was not discovered until later. In addition, both PER and TIM undergo phosphorylation/de-phosphorylation cycles via the kinases Double-Time (Dbt) kinase and Casein Kinase 2 (CK2) (Peschel and Helfrich-Förster, 2011) and phosphatase, protein phosphatase 2A (PP2A) (Sathyanarayanan et al., 2004). This suggests that post-translational modifications also play a role in the oscillation of PER and TIM protein levels.

The mechanisms of PER and TIM protein activity were unknown until two other Drosophila clock components *clock* (*clk*) and *cycle* (*cyc*) were discovered. These two genes, *clk* and *cyc*, encode proteins containing basic helix-loop-helix (bHLH) DNA-binding domains and PAS domains. The steady-state mRNA levels of *clk* cycle and peak near subjective dawn (ZT-12) (Bae et al., 2000; Darlington et al., 1998) however CLK protein levels do not (Houl et al., 2006). Contrary to its name, CYC does not cycle, neither at the mRNA nor protein level (Bae et al., 2000; Rutila et al., 1998). However, in *clk* or *cyc* mutants, *per* and *tim* mRNA levels are drastically reduced and do not cycle, indicating that these bHLH factors regulate transcription of PER and TIM (Allada et al., 1998; Rutila et al., 1998). This is thought to be a direct effect of CLK/CYC heterodimers binding directly to E-box elements found in the *per* and *tim* promoters and inducing gene expression (Darlington et al., 1998; Hogenesch et al., 1998). This induction of gene expression is antagonized by the actions of PER and TIM (Darlington et al., 1998). It has been shown that binding of PER and TIM to the CLK/CYC complex prevents DNA binding but doesn't disrupt their association (Bae et al., 2000; Lee et al., 1998, 1999). Hence, CLK and CYC are the positive elements of the transcriptional loop, whereas PER and TIM mediate negative feedback (**Figure 1-2**).

In addition to the major feedback loop described above, there exists a second loop interlocked to the first. In this loop, the CLK–CYC complex drives sequential expression of a transcriptional activator, PDP1 (PAR-domain protein 1), and a repressor, VRI (Vrille), of *clk* expression (Cyran et al., 2003; Glossop et al., 2003). PDP1 and VRI maintain rhythmic expression of *Clk* mRNA by feeding back on *clk* expression in a rhythmic manner. However, given that CLK protein levels do not cycle, the purpose of the mRNA cycling is unclear (Houl et al. 2006). It is thought that the second loop stabilizes the system and provides greater precision (Cyran et al. 2003; Glossop et al. 2003). A more detailed description of *Drosophila* circadian systems and their function can be found in the review by Dubowy & Sehgal (Dubowy and Sehgal, 2017).

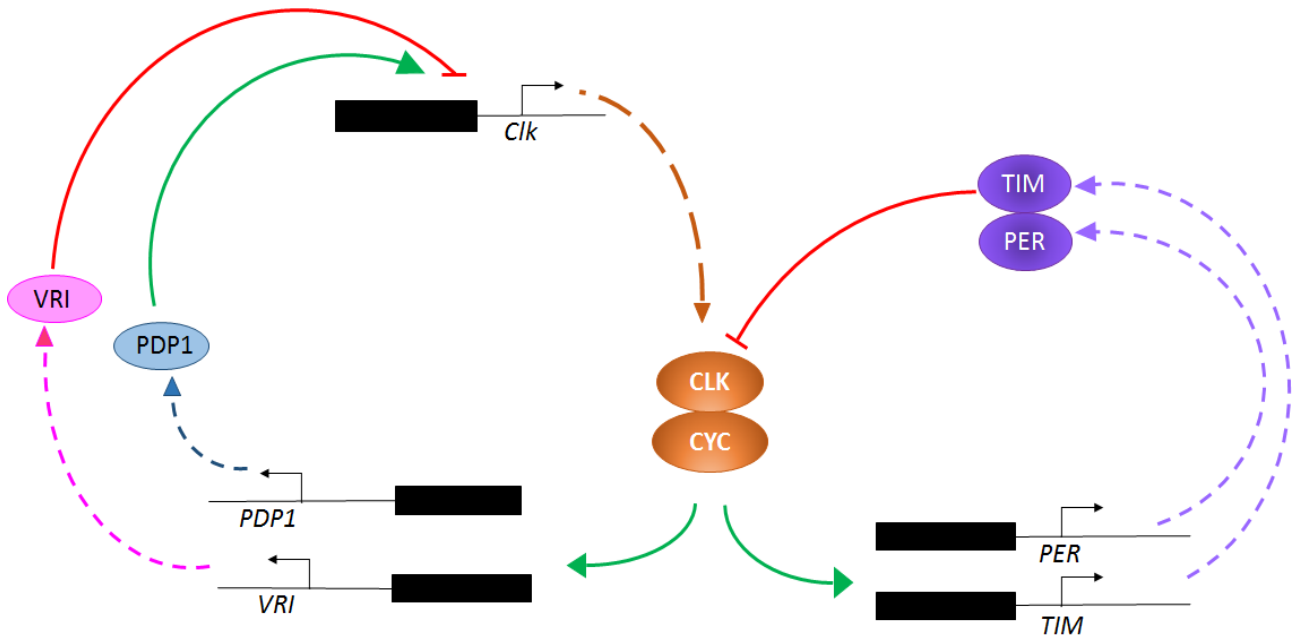



**Figure 1-2 Simplified schematics of feedback loops in *Drosophila* circadian clock**

CYC-CLK acts a positive component while PER-TIM act as negative component in the feedback loop while a secondary feedback loop is composed of PDP1(Positive component) and VRI(Negative component) controlling *Clk* promoter. (Colored dotted lines represent mRNAs. Green arrows indicate positive effect of a component on some process or promoter, and red lines indicate an inhibitory action. Black boxes indicate promoter. While proteins are indicated by ovals. )

**The circadian clock in mammals.**

The mammalian circadian clock resembles the fly clock. The main mechanism of the mammalian clock is transcriptionally and translationally regulated feedback loop systems. A multitude of mammalian genes were found to be homologous to the *Drosophila* clock genes. These include three *Period* homologues (*mPer1*, *2*, and *3*), a *Tim* homologue (*mTim*), a *Clk* homologue (*mClk*), and a gene homologous to *cyc* called *BMAL1 (Brain and Muscle Arnt-like protien1)*. (For simplicity, the mammalian circadian genes are prefixed with an "m" while their Drosophila counterparts are prefixed with a "d".) Apart from *mClk*, which was discovered in a forward genetic screen the rest were primarily identified by sequence homology to their fly counterparts.

Compared to flies, the mammalian circadian clock has an additional layer of complexity arising from the presence of a master circadian clock in the SCN (the suprachiasmatic nucleus) and peripheral clock systems in different tissues (Damiola et al., 2000; Stokkan et al., 2001). The SCN is thought to coordinate and regulate peripheral clocks. Light based inputs from the outside world are directly received by the SCN from the retina via the retino-hypothalamic tract. The neurons within the SCN are coupled and together they produce electrical and molecular circadian rhythms of a robustness not seen

in isolated neurons and other types of cells (Liu et al., 2007). For simplicity, the discussion would be limited to gene function in the SCN.

Like the fly clock, the positive elements of the mouse clock are the bHLH/PAS transcription factors called mCLK and BMAL1. These TFs form a heterodimer and activate the transcription of *Bmal1, Cry1, Cry2, Per1* and *Per2*. *mPer* transcripts levels are reduced in both *mClk* and *BMAL1* mice mutants (Gekakis et al., 1998), supporting the role of mCLK and BMAL1 as positive elements of the mouse clock just like their fly homologues. The mCry and mPer proteins form heterodimers and disrupt the mClk-BMAL1 complex, resulting in downregulation of mClk-BMAL1 targets, including *cry* and *per*. In rodents, *BMAL1* mRNA and protein levels cycle, both peaking around the middle of the subjective night at ZT-18. In contrast, the *mClk* mRNA does not cycle (Shearman et al., 2000). All three mPER proteins have some antagonistic effects on mCLK/BMAL1-mediated gene activation and this action is independent of mTIM. (Griffin et al. 1999, Kume et al. 1999, Sangoram et al. 1998). Unlike the Drosophila system, the role of mTIM in the mammalian circadian system is not clear because no association was found between mPER and mTIM proteins in the SCN (Field et al., 2000). mPer proteins do interact with Cry proteins and this heterodimer impedes mClk-BMAL1 gene activation, possibly fulfilling the role of the Drosophila dPER-dTIM heterodimer.

The negative feedback on the mammalian circadian system is instead supplied by two cryptochromes (cry), *mCry1* and *mCry2 and mPER*. Cryptochromes may bind pterin and flavin chromophores and are related to the DNA repair enzyme photolyase, but have no repair activity. In plants and flies, cryptochromes act as blue-light photoreceptors and transmit light information to the circadian clock (reviewed in Sancar 2000 and Devlin & Kay 2001). However mCRYs have not been reported to be involved light-dependent activity in mammals to date. Further it has been shown that the interactions between mCRYs and mCLK/BMAL1 are light independent (Griffin et al., 1999). In the mammalian circadian clock, mCRYs inhibit mCLK/BMAL1-mediated gene activation just as effectively as dPER/dTIM complexes inhibit dCLK/CYC (Figure 1-3) (Griffin et al., 1999; van der Horst et al., 1999; Kume et al., 1999; Okamura et al., 1999; Vitaterna et al., 1999).

A second feedback loop involves the retinoic acid-related orphan receptor (ROR) (α, β, and γ) and REV-ERB (an orphan receptor encoded on the noncoding strand of the thyroid alpha gene) (α and β) proteins. The RORs act as transcriptional activators and REV-ERBs are repressors. The BMAL1/CLOCK binds to E-box elements present in *Ror* and *Rev-erb* genes and activates their transcription. RORs and REV-ERBs in turn control rhythmic transcription of the BMAL1 gene (Preitner et al., 2002)

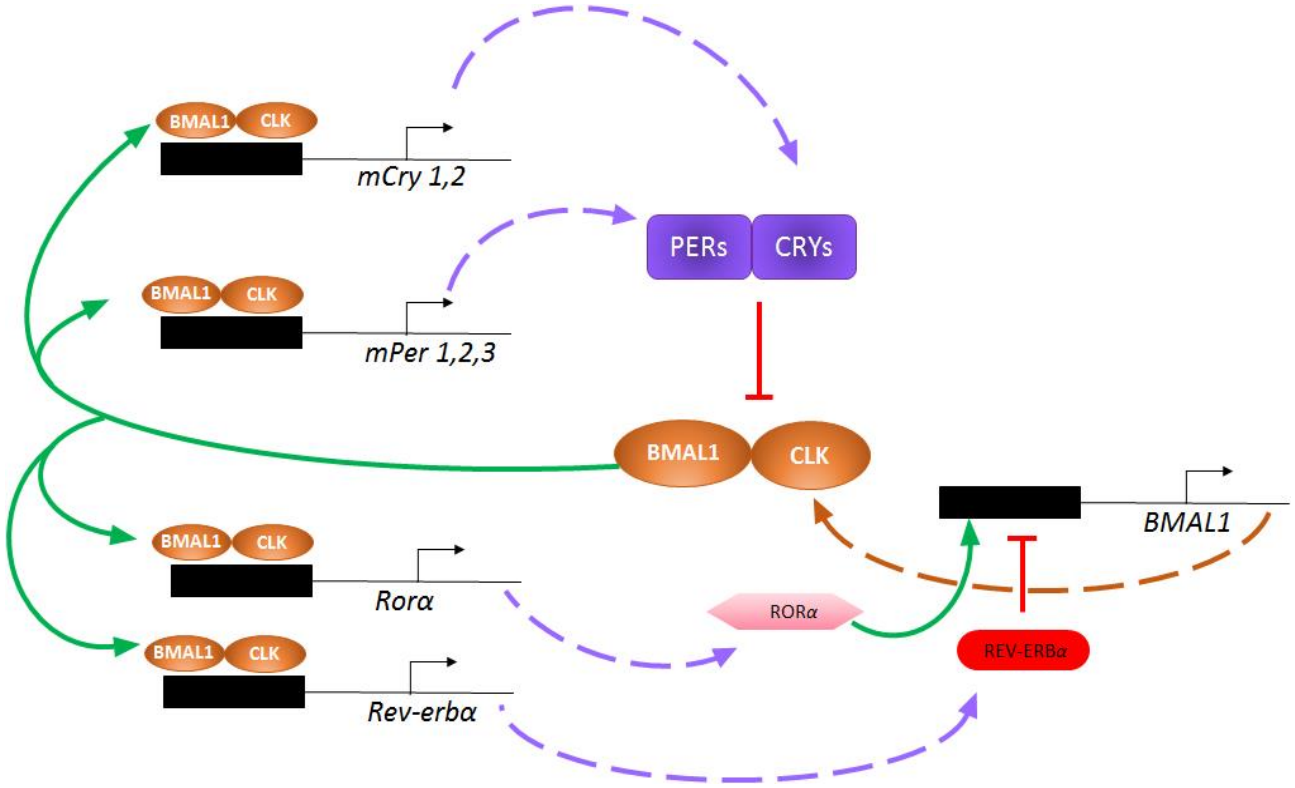



**Figure 1-3 Simplified schematics of feedback loops in mammalian circadian clock.**

The core mechanism of the mammalian circadian clock and its link to energy metabolism. The cellular oscillator is composed of a positive loop (CLOCK and BMAL1) and a negative loop (CRYs and PERs). The CLOCK:BMAL1 heterodimer binds to enhancer sequences located in the promoter region of *Per* and *Cry* genes, activating their transcription. After translation, PERs and CRYs inhibit CLOCK:BMAL1, resulting in decreased transcription of their own genes. CLOCK:BMAL1 heterodimer also induces the transcription of *Rev-erbα* and *Rorα*. RORα and REV-ERBα which also regulate *Bmal1* expression. While RORα stimulates, REV-ERBα inhibits *Bmal1* transcription.

Colored dotted lines represent mRNAs. Green arrows indicate positive effect of a component on some process or promoter, and red lines indicate an inhibitory action.

The SCN oscillator is also responsible for controlling peripheral clocks and behavior rhythms through neural signals (e.g. sympathetic nerves) and humoral signals (e.g. hormone and cytokines). A more detailed description of the mammalian circadian systems and its function can be found in reviews by (Buhr and Takahashi, 2013; Honma, 2018; Partch et al., 2014; Reppert and Weaver, 2001)

Having briefly discussed the circadian systems in different organisms ranging from cyanobacteria to mammals, it is clear that circadian systems are composed of both transcriptional feedback loops and post-translational modification oscillators. As complexity in organisms increase, so does the complexity in their circadian system. This complexity is evident in plant circadian systems as well

In the next section, a detailed description of plant (*Arabidopsis thaliana*) circadian systems will be presented with its main components detailed and compensation/entrainment processes described with a particular focus on the role of the Evening Complex, a three protein complex consisting of LUX ARRYTHMO (LUX), EARLY FLOWERING 3 (ELF3) and EARLY FLOWERING 4 (ELF4). The evening complex is the main subject of this thesis. The thesis deals with molecular mechanism of evening complex function in mediating thermal responses. It focuses on mechanisms of the EC formation, structural aspects of DNA binding by LUX and binding affinity studies. Also, effects of cis elements mutation on a EC target - PHYTOCHROME INTERACTING FACTOR 4 (PIF4)- promoter.

## 1.2 Circadian systems in plants.

The first reported study of circadian systems in plants dates back to 1729 by the French astronomer de Mairan who demonstrated that the daily leaf movements of heliotropes persisted in consistent darkness. This presented an idea of an endogenous rhythm which de Marian related to the sleep rhythms of bedridden humans (de Mairan, 1729). It took another 30 years before de Mairan's observations were independently repeated (Hill, 1757; Duhamel duMonceau, 1759; Zinn, 1759). These studies excluded temperature variation as possible factor for driving leaf movement rhythms. It took another century before the period lengths of these leaf movements were accurately measured and were determined to be approximately 24 hours, making them circadian (Candolle, 1832). It was further demonstrated by Candolle that these rhythms could be inverted by reversing the alternation of light and dark.

It was realized that apart from the leaf movement rhythm, there existed rhythms in germination, stomatal movement and gas exchange, growth, enzyme activity, photosynthetic activity, flower opening, and fragrance emission (Cumming and Wagner, 1968). In 1985 genetic studies of plant clock revived when Kloppstech described circadian rhythm in pea. It was demonstrated that the abundance of three nuclear-encoded transcripts encoding the light-harvesting chlorophyll a/b binding protein (LHCB; also called CAB), the small subunit of ribulose-1,5-bisphosphate carboxylase/oxygenase, and an early light-induced protein were circadian in nature (Kloppstech, 1985). A similar observation was made in wheat, where it was shown that the transcription rate for the *Cab-1* gene was under circadian control (Nagy et al., 1988). During the same time Arabidopsis was emerging as a powerful system to combine forward genetic analysis with molecular gene cloning techniques. Hence using Arabidopsis as a model system, it was soon demonstrated that the transcription rate and transcript accumulation of Arabidopsis *Cab* (Millar and Kay, 1991) and a number of other genes (McClung and Kay, 1994) were under circadian control.

## 1.3 The Arabidopsis Circadian Clock

In the model plant *Arabidopsis thaliana* a number of biological processes are regulated by the circadian clock. To understand the molecular mechanism of clock function in Arabidopsis, mutant screening and genetic mapping-cloning approaches have been used. The tissues for RNA extraction, RNA gel blotting, and nuclear run-on analyses had to be harvested at intervals of 3-4 hours over fairly lengthy time courses. Forward genetic analysis also required a sensitive, reliable, and nondestructive assay that could score the circadian activity of individual seedlings without killing them. These requirements were met by luciferase assay system that offered a versatile noninvasive approach. Firefly luciferase (LUC) catalyzes oxidative decarboxylation of luciferin with the simultaneous release of a photon at 560 nm. This photon emission can be quantified with luminometers or with sensitive charge-coupled device cameras (Welsh et al., 2005). It was demonstrated that a short promoter fragment of the *Arabidopsis LHCB1*3 (CAB2)*

could drive rhythmic transcription and mRNA accumulation of *LUC* mRNA. The luminescence was detected as rhythmic light emission from individual Arabidopsis seedlings bearing the *LHCB:LUC* transgene (Millar et al., 1992). Fortunately, *LUC* mRNA is sufficiently unstable and its accumulation tracks the transcription rate and when driven by the LHCB promoter, it is rhythmic. After this initial demonstration in Arabidopsis, LUC use in circadian studies spread to other organisms, including Drosophila and mammals (Welsh et al., 2005).

With the development of the LUC reporter assay system, it was possible to screen for Arabidopsis clock mutants. Arabidopsis seeds bearing the *LHCB:LUC* transgene were subjected to chemical mutagenesis and the seedlings arising from these plants were screened for discovering clock mutant. The first clock gene to be discovered in *Arabidopsis* using this approach was the TIMING OF CAB EXPRESSION1(*TOC1*) gene. (Millar et al., 1995). Soon a number of other *Arabidopsis* circadian genes were discovered using the same approach. *ZEITLUPE (ZTL )*(Somers et al., 2000), *TEJ*(from Sanskrit for "fast") (Panda et al., 2002), *TIME FOR COFFEE* (TIC )(Hall et al., 2003), *LUX ARRHYTHMO* [*LUX* or *PHYTOCLOCK 1 (PCL1 )*] (Hazen et al., 2005a; Onai and Ishiura, 2005), *FIONA1 (FIO1)*(Kim et al., 2008) and *PROTEIN ARGININE METHYL TRANSFERASE 5 (PRMT5 )* (Hong et al., 2010), were identified as key clock genes. Other genes pertaining to circadian functions, such as of *EARLY FLOWERING 3 (ELF3)* (Zagotta et al., 1996) , *ELF4*(Doyle et al., 2002) , *GIGANTEA (GI )* (Fowler et al., 1999) and LATE ELONGATED HYPOCOTYL (LHY) (Schaffer et al., 1998) were discovered screening for mutants impaired in biological processes that are regulated by circadian clock, such as photoperiodic flowering and hypocotyl elongation. Circadian genes *CIRCADIAN CLOCK ASSOCIATED 1 (CCA1)* and *CCA1 HIKING EXPEDITION (CHE)* were identified by isolating proteins that bind to rhythmic gene promoter (Wang et al. 1997, Wang and Tobin 1998, Pruneda-Paz et al. 2009). While four homologs of TOC1, *PSEUDO-RESPONSE REGULATOR 9 (PRR9), PRR7 , PRR5* and *PRR3* were discovered through reverse genetics approach. (Matsushika et al. 2000, Eriksson et al. 2003, Michael et al. 2003, Yamamoto et al. 2003, Para et al. 2007)

Through these classical genetic approaches, at least 25 genes have been identified to be associated with clock functions. Of these 25 genes, 9 genes are classified as core clock genes in Arabidopsis **(Table 1-1)**. In the past decade, studies have revealed molecular functions of clock-associated genes that were previously undetermined. In the following sections, recent studies on the function of circadian gene circuits in *Arabidopsis* will be summarized.

| Gene | AGI code | Phenotype of knockout plants | Molecular Function | References |
|---|---|---|---|---|
| *CCA1* | At2g46830 | Short period (redundant with *LHY*) | Transcription factor | Arabadi et al., 2002, Mizoguchi et al., 2002, Farre et al., 2005 |
| *LHY* | At1g01060 | Short period (redundant with *CCA1*) | Transcription factor | Arabadi et al., 2002, Mizoguchi et al., 2002, Farre et al., 2005 |
| *CKB3* | At3g60250 | Not reported | β subunit of CK2 | Sugano et al., 1998, 1999 |
| *CKB4* | At2g44680 | Not reported | β subunit of CK2 | Perales et al., 2006 |
| *PRR9* | At2g46790 | Long period (redundant with *PRR7*) | Transcription factor | Farre et al., 2005, Salome and McClung, 2005, Nakamichi et al., 2010 |
| *PRR7* | At5g02810 | Long period (redundant with *PRR9* and *PRR5*) | Transcription factor | Farre et al., 2005, Salome and McClung, 2005, Nakamichi et al., 2005, Nakamichi et al., 2010 |
| *PRR5* | At5g24470 | Short period (redundant with *PRR7*) | Transcription factor, Interact with TOC1 | Nakamichi et al., 2005, Nakamichi et al., 2010, Wang et al., 2010 |
| *PRR3* | At5g60100 | Short period | Interact with TOC1 | Para et al., 2007 |
| *TOC1* | At5g61380 | Short period | Transcription factor | Strayer et al., 2000, Pruneda-Paz et al., 2010 |
| *CHE* | At5g08330 | Not apparent | Transcription factor | Pruneda-Paz et al., 2010 |
| *PRMT5* | At4g31120 | Long period | Methyl transferase | Sanchez et al., 2010, Hong et al., 2010 |
| *ZTL* | At5g57360 | Long period (redundant with *FKF1* and *LKP2*) | Blue-light receptor, F-box | Somers et al., 2000, Kiba et al., 2007, Kim et al., 2007 |
| *FKF1* | At1g68050 | Not apparent (redundant with *ZTL* and *LKP2*) | Blue-light receptor, F-box | Imaizumi et al., 2003, Sawa et al., 2007, Baudry et al., 2010 |
| *LKP2* | At2g18915 | Not apparent (redundant with *ZTL* and *FKF1*) | Blue-light receptor, F-box | Imaizumi et al., 2003 Baudry et al., 2010 |
| *GI* | At1g22770 | Reduced robustness | Interact with ZTL family | Fowler et al., 1999, Park et al., 1999, Kim et al., 2007, Sawa et al., 2007 |
| *LUX (PCL1)* | At3g46640 | Arrhythmic in constant light (LL) | Transcription factor | Hazen et al., 2005, Onai and Ishiura, 2005, Helfer et al., 2011 |
| *ELF4* | At2g40080 | Arrhythmic in LL | Transcription factor | Doyle et al., 2002, Dixon et al., 2011 |
| *ELF3* | At2g25930 | Arrhythmic in LL | Transcription factor | Hicks et al., 1996, Dixon et al., 2011 |
| *TEJ* | At2g31870 | Long period | Poly (ADP-ribose) Glycohydrolase | Panda et al., 2002 |
| *LWD1* | At1g12910 | Not apparent (redundant with *LWD2*) | Associae with *PRR9* promoter | Wu et al., 2008 Wang et al., 2011 |
| *LWD2* | At3g26640 | Not apparent (redundant with *LWD1*) | Unknown | Wu et al., 2008 |
| *FIO1* | At2g21070 | Long period | Unknown | Kim et al., 2008 |
| *TIC* | At3g22380 | Short period | Unknown | Hall et al., 2003 |
| *JMJD5 (JMJ30)* | At3g20810 | Short period | Probable histone demethylase | Jones et al., 2010, Lu et al., 2011 |
| *MYB3R2* | At4g00540 | Not reported | Transcription factor | Hanano et al., 2008 |
| *bHLH69* | At4g30980 | Not reported | Transcription factor | Hanano et al., 2008 |

**Table 1-1 List of genes involved in the Arabidopsis circadian clock.**
Genes colored in red act in the morning, green-colored genes are active from early daytime until midnight, and those in light blue and orange act during the night (adapted from Nakamichi, 2011 ).

**Genetic circuit in the *Arabidopsis* circadian clock.**

In Arabidopsis, the first basic clock model was the positive-negative feedback loop, which is based on CCA1/LHY and TOC1, where CCA1/LHY act as the repressor while TOC1 acts as an activator of gene expression (Alabadí et al., 2001). However, this early model didn't account for other components that were discovered eventually. Hence to include newer genes that were discovered in the Arabidopsis circadian clock, a new model referred to as the "bar code clock" model based on the *PRR* gene family was developed (Matsushika et al., 2000). The *PRR* genes are expressed through the day and the *PRR* genes (*PRR1*, *PRR3*, *PRR5*, *PRR7*, and *PRR9*) were subjected to a circadian rhythm at the level of transcription. Furthermore, in a given 24 h period, the *PRR* mRNAs started accumulating sequentially after dawn with 2-3 h intervals in the order of *PRR9* > *PRR7* > *PRR5* > *PRR3*> *PRR1*. However, this model didn't explain why the sequential transcription events were not significantly affected by the photoperiod conditions (e.g. long or short days), and the expression of *PRR9* was first boosted always after dawn. Suggesting there were missing elements in the Arabidopsis circadian clock to be explored. This model also lacked the integration of the feedback loops that exist between *LHY/CCA1* and *TOC1* and could not predict the effect of different mutations in the circadian genes.

A clock model which predicts the effect of mutations and incorporates our current knowledge of regulation and molecular function of clock associated genes is known as "The Repressilator" model (Herrero et al., 2012). This model proposes that three classes of repressors constitute the genetic circuit; (i) the morning-phase protiens CCA1 and LHY that repress ELF4 and LUX (ii) the evening phase proteins ELF3, ELF4 and LUX that repress PRR7 and PRR9 (iii) the mid-day phased proteins PRR9 and PRR7 that repress CCA1 and LHY (Helfer et al., 2011). This proposed synthetic genetic circuit which is a cyclic negative feedback loop that can produce an oscillating pattern of gene expression (**Figure 1-4**). The repressilator model is a sustainable oscillator that is dependent on similar decay rates of protein and mRNA and a large amount of proteins at its peak level. (Elowitz and Leibler, 2000; Shalit-Kaneh et al., 2018). This model is able to explain the behavior of wt clock and predicts the effect of *ko* mutants of gene oscillations.

Through these evolving models, interactions between different clock components and their expression patterns started to become clearer. It was possible to build a mathematical model based on temporal evolution of the concentration of various clock components using ordinary differential equations. The latest mathematical models are based on temporal gene expression and protein abundance profile of clock genes. These models captures key features of clock on a qualitative level, namely the entrained and free-running behaviors of the wild type clock, as well as the defects found in knockout mutants. (De Caluwé et al., 2016; Locke et al., 2006; Pokhilko et al., 2010).

In the next section we will discuss in more detail, the "Dawn-Phased", the "Mid-day Phased" and the "Evening Phased" genes that constitute the Arabidopsis circadian clock. Since the thesis mainly deals with the Evening Complex and its role in temperature sensing, in this section "The evening phase" will be discussed in greater details.

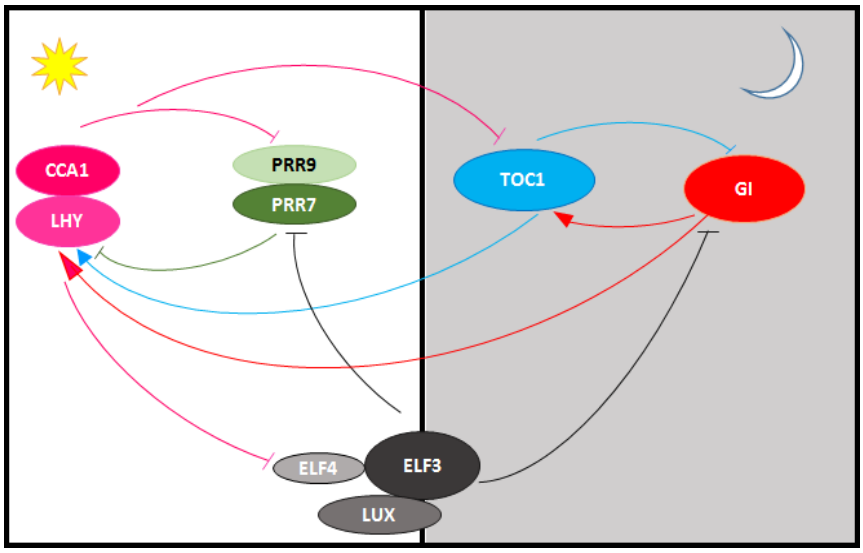



**Figure 1-4 Schematic diagram of feedback loops in Arabidopsis central oscillator.**

CCA1 and LHY repress *PRR9* and *PRR7*, and repress *TOC1*, *ELF4* and *LUX*. ELF4, LUX and ELF3 repress *PRR9* and *PRR7* and *GI*. PRR9, PRR7 repress *CCA1* and *LHY*. While GI and TOC1 activates CCA1 These interactions illustrate the 'genetic circuit', which underlies the 24 h endogenous cycle.

**The Dawn-Phase.**

*CCA1* and *LHY* are the dawn phase genes encoding single MYB domain containing transcription factors (Schaffer et al., 1998; Wang and Tobin, 1998). Constitutive overexpression of either gene, leads to arrhythmia, whereas *CCA1* or *LHY* loss-of-function mutations retain rhythmicity but with a shortened period (Green and Tobin, 2002; Mizoguchi et al., 2002). *CCA1* and *LHY1* are expressed in early morning and work by repressing their target genes by binding to the CCA1-binding element (AACAATCT or

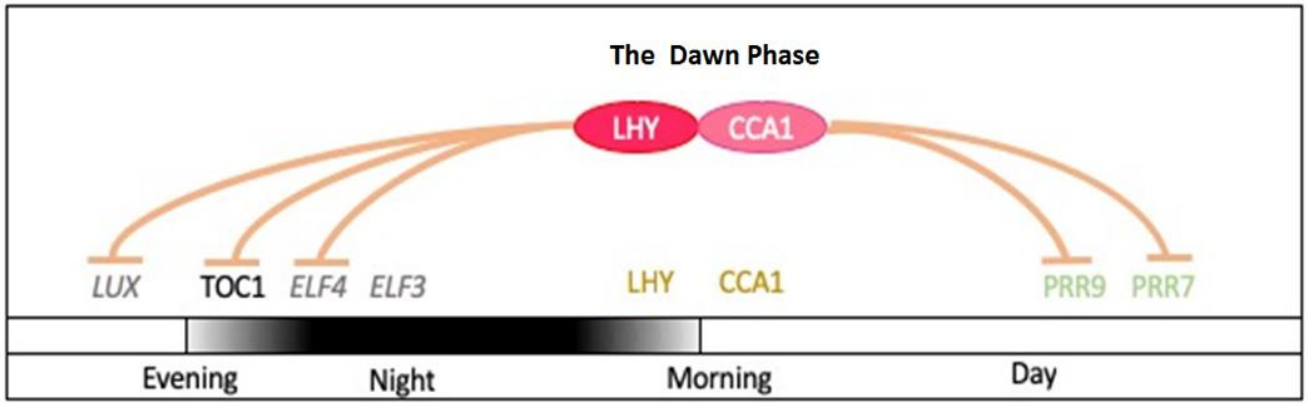



**Figure 1-5 The Dawn-Phased Genes**

In the morning, LHY and CCA1 proteins repress the evening-phase genes LUX, TOC1 and ELF4, and repress the Mid-day phased genes PRR9 and PRR7.

AAAAATCT) and evening elements (AAAATATCT)(Alabadí et al., 2001; Wang and Tobin, 1998). CCA1/LHY repress *TOC1*, *ELF4*, *ELF3* and *LUX* which are components of the evening phase and also *PRR9* and *PRR7* which are components of the mid-day phase loop.

**The Mid-day Phase.**

After the peak expression of the dawn-phased genes, sequential expression of PRR9, PRR7 and PRR5 (PRRs) occurs. The *PRR9* mRNA peaks at ZT-4 while the *PRR7* mRNA peaks at ZT-8 and PRR5 at ZT-12 (Farré et al., 2005; Nakamichi et al., 2010). The PRRs possess a N-terminal pseudo-receiver (PR) domain and a CONSTANS, CONSTANS-LIKE1 AND TOC1(CCT) motif at their C-terminal region (Makino et al., 2000; Matsushika et al., 2000; Strayer et al., 2000). The PRR's function as negative regulators of *CCA1* and *LHY* by associating to the promoters of CCA1/LHY and repress these genes during the day until midnight. Each PRR protein works at a specific time; during the early daytime (ZT-4 to ZT-8) PRR9 is active followed by PRR7 which is active until Zt-20 and PRR5 functioning from ZT 8 until midnight ZT-20 (**Figure 1-6**)(Nakamichi et al., 2010). The PRRs act through transcriptional repression. It has been demonstrated that PRR5, PRR7 and PRR9 associate with TOPLESS/TOPLESS-RELATED co-repressors, to repress the transcription of CCA1 and LHY (Wang et al., 2013). The CCT domain of PRRs share sequence similarity with the DNA binding domain of yeast HEME ACTIVATOR PROTEIN 2 (HAP2) and CO, which binds to CCAAT boxes in eukaryotic promoters (Wenkel et al., 2006). It has also been demonstrated that the CCT domain of several circadian regulated PRRs, including PRR7, binds DNA directly *in vitro* (Gendron et al., 2012). However the DNA binding specificity of the PRRs remains unknown.

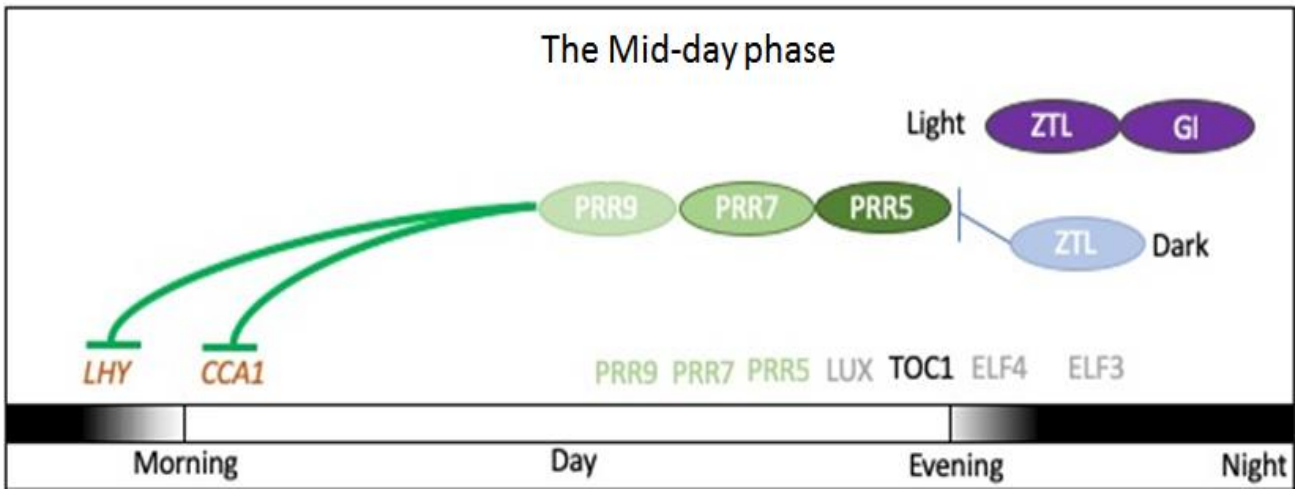



**Figure 1-6 The Mid-Day Phased Genes**

From early daytime until midnight, PRR9, PRR7 and PRR5 repress the morning-phase genes CCA1 and LHY. Blue light enhances interaction of ZTL and GI. In the dark, the ZTL–GI complex is decoupled, allowing ZTL to promote the degradation of PRR5.

**The Evening Phase**

The evening phase includes TOC1 and a protein complex consisting of ELF3, ELF4 and LUX called the Evening Complex (EC). The EC proteins, ELF3, ELF4 and LUX are expressed from the evening until midnight. While *ELF4* and *LUX* mRNA expression peak at ZT-12, *ELF3* expression peaks at ZT-16. All EC components are essential for sustaining circadian rhythms under constant light conditions. Mutation in any of these genes leads to arrhythmia. EC genes are required for full CCA1 and LHY expression as evidenced by decreased expression of CCA1/LHY in the *elf4*, *elf3* and *lux, loss-of-function* mutants (Doyle et al., 2002; Hazen et al., 2005a; Onai and Ishiura, 2005).

The EC is highly regulated by other clock components. The evening element (EE) which is denoted by 5'-AAAATATCT-3', is present in both *ELF4* and *LUX* promoters and is over-represented in promoters of evening phased genes, supporting the idea that EC transcripts are regulated by the clock (Harmer et al., 2000). The dawn-phased clock transcription factors CCA1 and LHY suppress expression of evening-phased genes by binding to the EE. Therefore the nightly peaks of ELF4 and LUX expression is likely regulated by CCA1/LHY. ELF3 differs from ELF4 and LUX due to the fact that it doesn't have a canonical EE. However it does have a EE like element (AATATCT) and two CCA1 binding sites (CBS, AA(A/C)AATCT) which is used by CCA1 for suppressing ELF3 activity. Recently, genome-wide target of CCA1 have been identified through ChIP-Seq showing that CCA1 occupies the promoter region of all EC components (Kamioka et al., 2016; Nagel et al., 2015) suggesting that CCA1 is a major regulator of EC expression. Apart from CCA1/LHY, other clock factors such as LNK1 and LNK2 (LNK stands for NIGHT LIGHT-INDUCIBLE AND CLOCK-REGULATED GENE) have been shown to activate the expression of *ELF4* (Rugnone et al., 2013) while an afternoon-peaking protein called RVE8 (REVEILLE 8) antagonizes CCA1 and can activate the expression of *ELF4* and *LUX* through binding to the EE element (Hsu and Harmer, 2014; Hsu et al., 2013).

As the evening progresses into the night, TOC1 has been found to suppress the expression of both *LUX* and *ELF4* (Huang et al., 2012). TOC1 also activates CAA1 expression by antagonizing the repression activity of CHE (Pruneda-Paz et al., 2009). TOC1 itself is targeted for degradation by ZTL(Más et al., 2003b) while its stabilized by PRR3 via protein-protein interactions(Para et al., 2007)All these data show that the EC expression is collectively regulated through various clock components.

The EC also regulates *Arabidopsis* circadian clock components. The EC is nuclear localized (Liu et al., 2001), mediating night-time repression of clock genes *TOC1*, *LUX*, *GI*, *PRR7* and *PRR9* (Dixon et al., 2011; Kolmos et al., 2009; Mizuno et al., 2014) and indirectly promoting expression of morning oscillators *CCA1* and *LHY* (Dixon et al., 2011; Kolmos et al., 2009). EC functions mainly through transcriptional repression (**Figure 1-7**). This is demonstrated by the fact that complementation experiments of *lux* mutants in which LUX fused to a strong repression domain leads to functional compensation , while LUX with a strong activation domain fails to complement the loss of function. (Helfer et al., 2011; Nusinow et al., 2011). In addition, the EC is able to regulate its own levels by binding to the LBS (LUX Binding Site) on the LUX promoter suggesting autoregulation (Helfer et al., 2011).

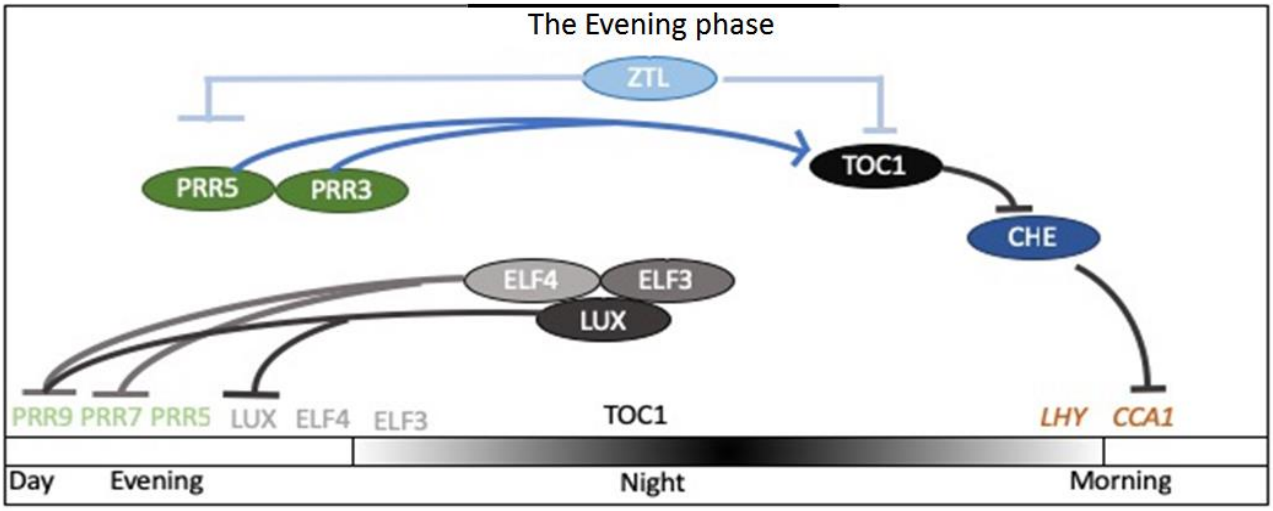



**Figure 1-7 The Evening-Phased Genes**

In the evening, LUX represses PRR9 and LUX expression, and ELF4 and ELF3 repress PRR9 and PRR7. PRR5 enhances nuclear localization of TOC1, and PRR3 stabilizes TOC1 in the evening. On the other hand, ZTL promotes the degradation of PRR5 and TOC1 at night (in darkness). TOC1 acts to activate CCA1 expression by antagonizing CHE on the CCA1 promoter.

## 1.4 Inputs to the Arabidopsis circadian clock

Circadian clocks can be entrained by environmental cues to synchronize the internal oscillator with the external environment. These environmental cues are generally light and temperature. In Arabidopsis, light is considered to be the most powerful and best-characterized entrainment signal. Although Arabidopsis circadian clock is temperature compensated, meaning overall clock period remains constant over a range of temperature, temperature changes or fluctuations can act to reset the clock output patterns. However very little is known about the mechanisms of temperature entrainment. In this section we will discuss both light and temperature entrainment of the Arabidopsis circadian clock.

### Light Inputs

Light signaling pathways, involving red/far-red light photoreceptors called phytochromes (PHY) and blue light photoreceptors called cryptochromes (CRY), regulate clock components to achieve entrainment in plants (reviewed by Fankhauser and Staiger, 2002). Arabidopsis has five phytochromes PHYA-E and two cryptochrome CRY1 and CRY2 (Oakenfull and Davis, 2017). PHYB is one of the most active phytochromes and relays light quality information by binding to clock components. Six clock components have shown to be bind to PHYB-CCA1 and TOC1 under far-red light, LUX under red light and LHY, GI and ELF3 under both red/far-red light (Yeom et al., 2014). Mutation in *phyb* leads to long period (Somers et al., 1998), thus indicating the major role of PHYB in maintaining free running rhythms. Blue light information is also conveyed to the clock by F-box proteins ZEITLUPE (ZTL), FLAVIN KELCH FLAVIN (FKF1) and LOV KELCH PROTEIN 2 (LKP2) by binding to GI forming a stable complex under blue light condition (Kim et al., 2007). The UV-B light is sensed by

photoreceptor UV RESISTANCE LOCUS 8 (UVR8) which interacts with COP1 under UV-B light and is shown to be necessary for UV-B light entrainment of the circadian clock (Yin et al., 2015).

**Temperature Inputs**

Circadian clocks have a stable period over a range of temperature making them temperature compensated. However temperature changes can act on clock output patterns. Very little is known about the mechanism of temperature entrainment. It has been shown that C-REPEAT/DRE BINDING FACTOR 1 (CBF1) mediates cold temperature inputs to the clock by binding to the *LUX* promoter and enhancing its expression (Chow et al., 2014). The EC has been implicated in ambient temperature responsiveness of the clock because EC mutants have mis regulation of *PRR7, GI* and *LUX* show differential repression at temperatures from 16-28°C due to changes in EC activity (Mizuno et al., 2014). Apart from this, the EC plays a major role of directly coordinating the expression of hundreds of key regulators of phytohormone signaling, growth and response to the environment. Recent findings using RNA-seq and ChIP-seq also show that the ability of EC to bind targets genome-wide depends on temperature. These experiments showed that EC binds to greater number of targets at lower temperature compared to higher temperature (Ezer et al., 2017a).

Recently it has also been shown that the red/far-red reversibility of PHYB is affected at higher ambient temperature. At higher ambient temperatures it was seen that there was an increase in the rate of reversion from active $P_{fr}$ state to inactive $P_r$ state, the abundance of the biologically active $P_{fr}$-$P_{fr}$ dimers pools and the size of associated nuclear bodies were also reduced. Only the active form of PHYB is able to interact with DNA and repress targets such as PIF4. This suggests that PHYB is also responsible for conveying temperature information to the clock (Legris *et al.*, 2016). It has also been reported that phytochrome null mutants display constitutive warm temperature responses and the warm temperature transcriptome becomes de-repressed at low temperatures (Jung et al., 2016). Also there is an enrichment of G-Box elements which are bound by PHYB in close proximity to multiple EC binding sites providing a mechanism for integrating environmental information (Ezer et al., 2017a).

**Regulation of outputs from the plant circadian clocks**.

In Arabidopsis more than 90% of expressed genes oscillate under cycling conditions, while approximately 30 % of the expressed genes continue to cycle even under constant conditions. This suggests that almost a third of the Arabidopsis transcriptome is circadian clock regulated (Covington et al., 2008; Michael et al., 2008). The circadian clock controls several developmental processes throughout the life cycle of the plant. Hypocotyl growth, photoperiodic flowering and cold acclamation are the best described pathways that are regulated with circadian clock output (Inoue et al., 2018). In this section, control of hypocotyl elongation and photoperiodic flowering by the circadian clock will be discussed in detail.

**Hypocotyl elongation**

The hypocotyl is the stem of a germinating seedling. Hypocotyl elongation in Arabidopsis is very well characterized growth that is regulated by the circadian clock. In Arabidopsis, hypocotyl elongation is

rhythmic, with maximal elongation occurring at dawn (Nozue et al., 2007). This rhythmic growth is regulated by bHLH transcription factors called PHYTOCHROME INTERACTING FACTOR (PIF) with PIF4 and PIF5 playing major roles. PIF4 and PIF5 accumulate to high levels during late night and are degraded to low levels in the presence of light. The EC represses *PIF4* and *PIF5* expression in early evening (Nusinow et al., 2011). Furthermore, ELF3 physically interacts with PIF4 to inhibit its transcriptional activity at early night. During the day PIF4 and PIF5 are targeted for degradation by light activated PHYB (Lorrain et al., 2009). Thus, PIF4 and PIF5 activity are tightly regulated by circadian clock components, resulting in restriction of their activity to late night. Thus hypocotyl growth under diurnal conditions is largely controlled by circadian clock (Inoue et al., 2018).

**Regulation of flowering time by circadian clock**

Photoperiodic flowering is the most characterized event in plant development that is regulated by the circadian clock. Flowering in Arabidopsis is controlled by a zinc-finger containing transcription factor called CONSTANS(CO) by positively regulating *FLOWERING LOCUS T* (*FT*). CO is strictly regulated by the circadian clock and light signaling pathways. Circadian clock components PRR9, PRR7 and PRR5 negatively regulate *CYCLING DOF FACTOR 1* (*CDF1*) which is a negative regulator of *CO* (Nakamichi et al., 2007) thus promoting flowering.

Although the two EC components ELF3 and ELF4, were originally discovered in screening for mutants that had disrupted flowering time in *A.thaliana*, their exact roles in mediating transition to flowering were unclear (Doyle et al., 2002; Hicks et al., 2001). ELF3 was found to be a substrate adaptor for COP1-dependent degradation of GI, resulting in reduced expression of *CO* and *FT* resulting in delayed flowering. (Yu et al., 2008). Additionally, SHORT VEGETATIVE PERIOD (SVP), a MADS-box transcription factor, is also known to repress *FT*. It was shown that SVP directly interacts with ELF3 and accumulates in ELF3 overexpression line which is consistent with their late flowering phenotype due to reduced *FT* expression (Yoshida et al., 2009). Also the EC target PIF4 binds to *FT* promoter in a temperature dependent manner and interacts with CO to induce flowering at higher temperature under non inductive short day conditions (Kumar et al., 2012). As PIF4 and PIF5 are also required for warm-night induced early flowering (Thines et al., 2014), its speculated that the EC indirectly regulates flowering by modulating expression of the *PIFs*.

## 1.5 Thesis objectives

Temperature fluctuations occur naturally during plant growth and reproduction. However, it has been shown that even small changes in mean ambient temperature leads to profound changes in plant growth and development (Quint et al., 2016). In crops like rice, a 1°C change in mean minimum temperature during the wet growing season can result in 10% reduction in grain yield (Peng et al., 2004). Likewise, over the past two decades in Europe a 10% stagnation in yield of wheat and barley can be attributed to climate change (Moore and Lobell, 2015). Similar effect on increase in mean temperature were observed for tomatoes, where increase in the mean temperature leads to contraction of the growing season of tomatoes leading to lower yields (Ventrella et al., 2012).

The EC plays a central role in the circadian clock and is likely a key player in thermosensitive growth and development. However the molecular mechanisms underlying this activity were unknown. The broad objective of my thesis was to understand the mechanisms of EC formation, DNA-binding and activity. In order to achieve these ambitious goals, I used a combination of *in vitro*, structural and *in planta* experiments to probe EC activity.

In **chapter 2**, the individual components of EC will be discussed. Their known biochemical properties and interactions with other proteins will be reviewed, followed by their predicted structures. Details about how each component of EC was purified for performing various in vitro experiments will be discussed. Also results pertaining to invitro reconstitution of the evening complex will be discussed in this chapter.

In **chapter 3**, binding affinity and crystal structure of LUX will be discussed. The crystal structure of LUX MYB domain provided us with insights of important amino acids responsible its DNA Binding activity. Based on this knowledge experiments were done to understand the effect of site directed mutagenesis (SDM) on binding affinity of LUX, which is presented in this chapter.

In **chapter 4**, results from *in planta* validation of mutagenized version of LUX will be presented. In order to see that binding affinity of the single amino acid mutated version of LUX is conserved *in planta*, the mutated version of LUX along with the wildtype version of the protein were introduced in the *lux* mutant background. The goal was to understand how important was the mutation when all other components of EC were present. In this chapter, phenotypical data will be presented and compared between various mutant lines and wildtype plants. Along with this, a comparison will be done on how *PIF4* expression changes in different mutant lines with respect to the wildtype plants.

In **chapter 5**, effects of *PIF4* promoter manipulation on thermo-responsiveness will be discussed. EC binds to certain regions in the *PIF4* promoter. Experiments were performed on wildtype Arabidopsis plants using CRISPR/Cas9 to mutate regions where EC or putative EC interacting partners bind to investigate the effects of mutation on thermotolerance. In this chapter results will be presented pertaining to growth and developmental changes occur due to these mutations on the *PIF4* promoter.

In **chapter 6,** general discussion and conclusion of work done during the thesis are presented with the scope of works that need to be done in future. A general discussion of how probably the EC might be

involved in temperature sensing is presented in this chapter. Also, how the work done in Arabidopsis could be translated into crop plants to make them thermotolerant at higher ambient temperatures are presented in this chapter.

# CHAPTER 2. **BIOCHEMISTRY AND PURIFICATION OF ELF3, ELF4 AND LUX**

## 2.1 Introduction

All of the components of the EC- LUX, ELF3 and ELF4 - are plant specific proteins, with little or no homology to proteins in other kingdoms of life apart from the ~52 amino acid MYB DNA-binding domain (DBD) of LUX (Hazen et al., 2005a; Onai and Ishiura, 2005). ELF3 (Hicks *et al.*, 2001, Liu *et al.*, 2001) and ELF4 (Kolmos *et al.*, 2009) have no identifiable domains of known homology and functional characterization thus making it difficult for building accurate structural models by comparisons with other proteins. In this chapter, there are two section, the first section deals with EC components while the second section deals with in vitro reconstitution of the EC.

**The evening complex components.**

In this section, the Evening Complex components will be discussed in details. The details of what is known about the individual proteins in this complex (LUX, ELF3 and ELF4) and their interaction with other proteins will be discussed. This section would also include results pertaining to predicted structure and disorder of these protein, followed by results from expression and purification of individual EC components.

## 2.2 LUX ARRHYTHMO (LUX)

LUX contains a MYB DNA binding domain. MYB-like TFs are widely distributed in eukaryotic organisms (Riechmann et al., 2000). Plants contain further large numbers of MYB genes in contrast to animals and fungi. Arabidopsis alone contains 197 member forming the largest TF family (Riechmann and Ratcliffe, 2000). MYB TFs have two distinct regions, an N-terminal conserved MYB DNA binding domain and a diverse C-terminal modulator region responsible for the regulatory activity of the protein. A MYB domain is a 51-53 amino acid domain, forming three α- helices. Based on the number of adjacent MYB domain repeats that are present, the MYB family is divided into four classes, 1R, R2R3, 3R and 4R MYBs. (Dubos et al., 2010). LUX belongs to a sub family of 1R with heterogeneous members (~ 50 members) that have a single MYB repeat (Imamura et al., 1999).

### 2.2.1 Disorder analysis and structure prediction for LUX

Regions of protein disorder in LUX were predicted using SPOT-disorder server (Hanson et al., 2017). The software predicted most of the C-terminal and N-terminal to have intrinsic disorder with disorder probability greater than 0.5. The MYB-type DNA Binding domain stretches from Gly139 to Lys200. This region has a disorder probability of less than 0.5 (Figure 2-1a).

Secondary structure prediction of LUX was done using the PredictProtein server (Rost et al., 2004). Prediction algorithms suggested that the DNA binding domain contained 3 helical regions along with a short β-strand in the beginning of the DNA binding domain. Presence of a short helix and three short β-strands were predicted in the N-terminal while on the C-terminal there was one short helix predicted (Figure 2-1b).

Based on these prediction it was decided that recombinant expression of both the full length and the MYB domain of LUX (From here on referred to as LUX DBD )would be done. LUX DBD was of particular interest because this region of the full length protein is predicted to be stable and folded.

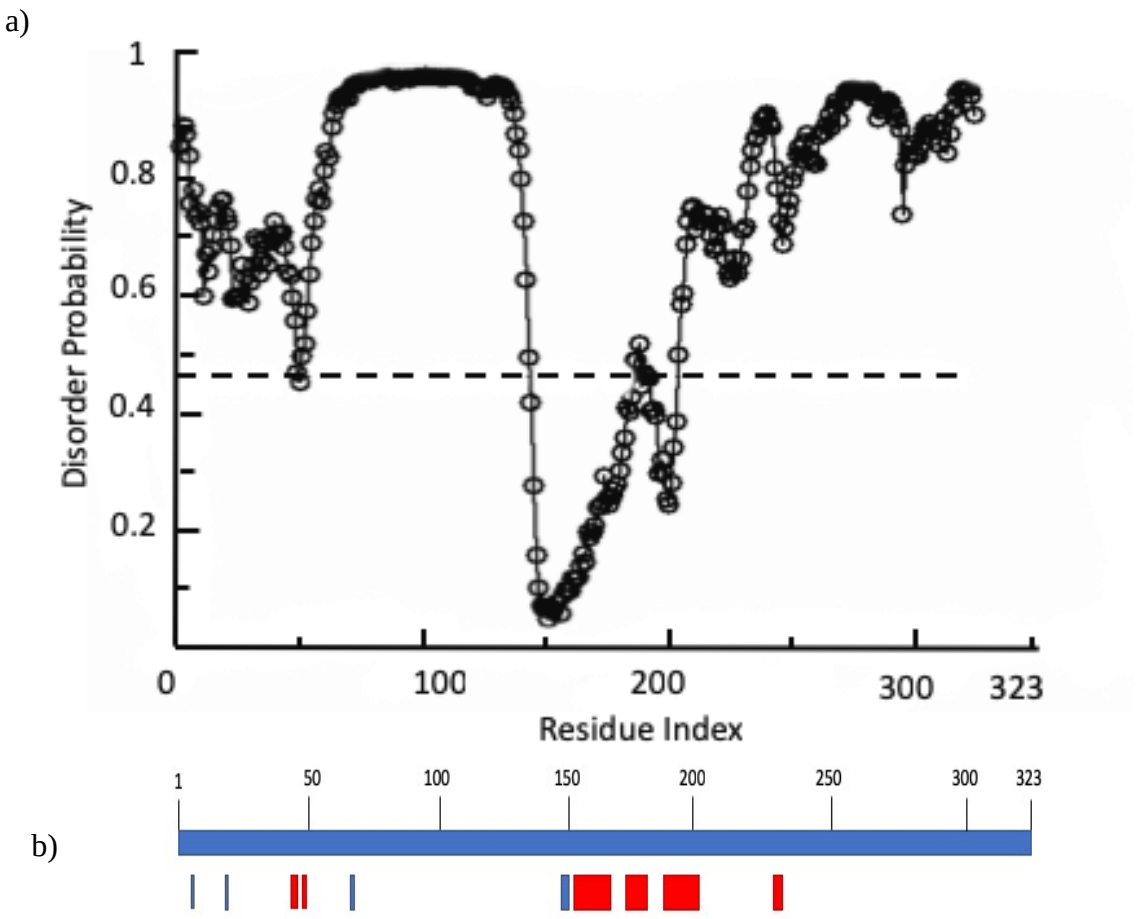



**Figure 2-1: Secondary Structure and Binding Site prediction of *LUX***

**a)** Disorder prediction using SPOT-D shows that with standard .48 threshold, most of the N-terminal and C-terminal of LUX is probably unstructured. The only structured part contains the DNA binding domain. **b)** PredictProtein predicts the presence of helixes in the structured DNA binding domain (indicated in Red) and a few β-strands (indicated in blue) numbers indicate protein length. (Amino acid sequence with structure prediction can be found in the appendix)

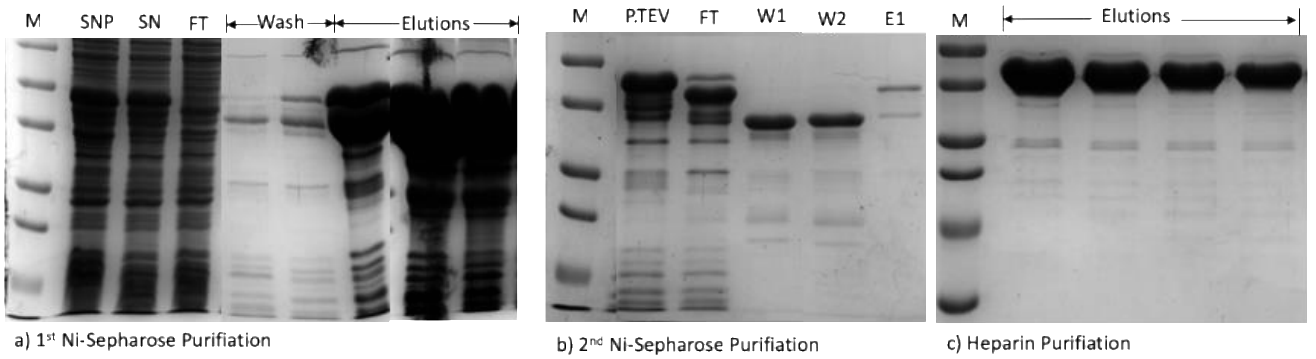
### 2.2.2 Purification of LUX full length protein.

LUX full length (FL) is a 323 amino acid long protein with a molecular weight of 34.6KDa and a theoretical pI of LUX FL is 5.47. LUX FL was expressed with a 6xHistidine tag followed by TEV cleavage site to obtain histidine tag free protein for downstream experiments. Obtaining good yields of stable and soluble LUX FL was a challenge. In order to address this a number of different expression trials were performed and purification protocols investigated. We started with Tris buffer at pH 7.5 and

changed pH and salt concentration of the buffer to obtain soluble LUX. The attempts to purify LUX in Tris buffer were of limited success as the yield was very low (100µg/L) and protein precipitation was issue which had to be dealt with. HEPES and CAPS buffer were tested after Tris buffer. It was found that CAPS buffer at pH 10.5 with 500mM NaCl was the best for extraction of LUX-FL and LUX DBD. A possible explanation for the effect of CAPS pH 10.5 on stability of LUX and LUX DBD during purification could be a consequence of structural modification of the protein accompanying change in pH. Hence the CAPS buffer was used for all subsequent purifications of LUX-FL and LUX DBD. LUX-FL was successfully purified and concentrated to a final yields of 1-2 mg/L.

The presence of 6xHis-tag on LUX FL allowed its purification on a Nickel – Sepharose column. It was found that there were considerable amount of high and low molecular weight impurities in the elution (**Figure 2-2 a-b)**. LUX is a DNA binding protein, hence it should be possible to separate LUX from impurities on Heparin Column which is negatively charged thus aiding separation of DNA binding protein. Hence it was decided to add an extra step of Sepharose™ HPLC purification. Although it was possible to separate high molecular weight impurities to a large extent, the smaller molecular weight impurities were not clearly separated (**Figure 2-2 c**).

The Heparin purified LUX was concentrated and a final S200 purification step was done to separate remaining impurities. In the size exclusion column purification using a S200 10/300 GL column, it was seen that LUX eluted at a 13.6ml elution volume (**Figure 2-3a**). With the S200 purification we were able to further separate higher molecular weight impurities while there were still some lower molecular weight impurities remaining (**Figure 2-3b**).




a) 1st Ni-Sepharose Purifiation

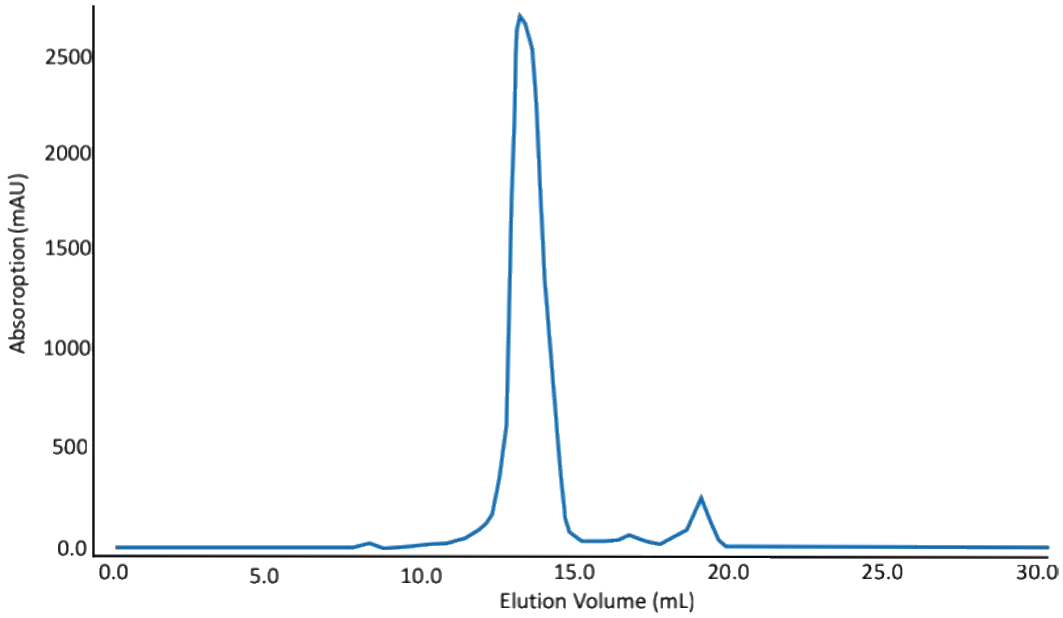
b) 2nd Ni-Sepharose Purifiation

c) Heparin Purifiation

**Figure 2-2 LUX Full Length Purification**

**(a)** Samples from the Ni-Sep purification of LUX-FL resolved on 8% SDS-PAGE. The lanes indicate Dual Colour Marker (M) pellet (SNP), supernatant (SN), Flow through (FT), Wash, Elutions **(b)** 2nd Ni-Seph purification after TEV protease cleavage of LUX FL protein. The lanes indicate Dual Colour Marker (M), P.TEV is the LUX FL sample before TEV, FT is LUX FL after TEV cleavage passed over a Ni-Speharose Column while W1 and W2 are samples obtained from clumn washing that contain the TEV protease. E1 is elution sample containing un cleaved LUX FL. **(c)** Purification of TEV cleaved LUX FL on Heparin-Sepharose ™ column, M is the dual color precission plus marker lane and the rest of the lanes contain a representative elution sample fractions from Heparin purification

a)




b)

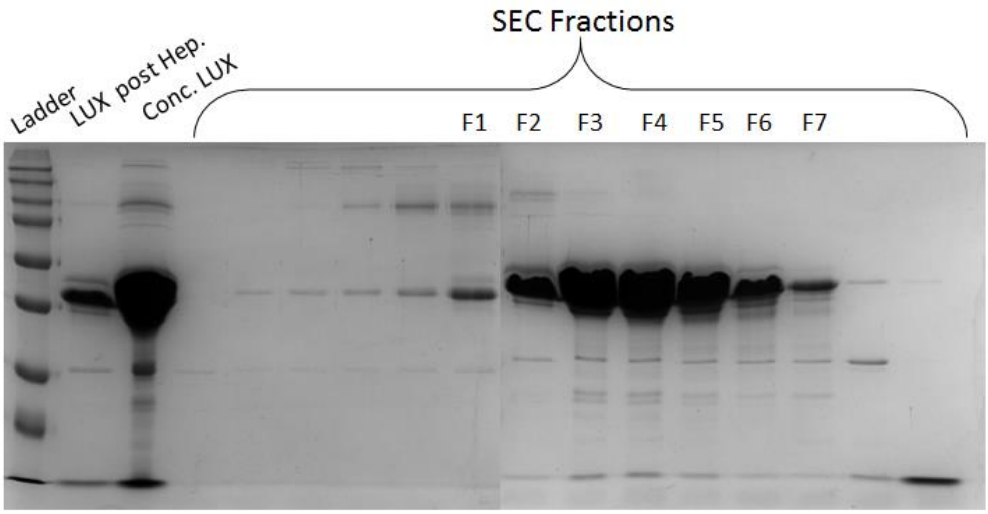



**Figure 2-3 Purification of LUX FL over S200™ Column.**

**a)** Size exclusion chromatograph of LUX FL on a S200 10/300 GL Column reveals that LUX-FL elutes at 13.6 mL elution volume. Elution fraction were collected every 0.5ml elution volume and resolved on a SDS PAGE to observe the proteins present in the elution. **b)** SDS PAGE for S200 peak fraction obtained at from 11.5ml to 15ml of elution volume. Lane 1 is Dual Color Protein Ladder, lane 2 is LUX sample loaded from post Heparin separation, lane 3 is concentrated sample of Heparin purified LUX used for S200 purification and rest of the lanes are samples from the S200™ purification. Lanes F3, F4 and F5, F6 were pooled together.

### 2.2.3 Purification of LUX DBD

The DNA binding domain of LUX was cloned with the MYB domain along with 10 extra amino acids on the N-terminal to the MYB domain. Hence LUX DBD is 62 amino acids in length with a molecular weight of 7.2KDa. With the 6xHistidine tag and TEV cleavage site that was cloned to the N-terminal of LUX DBD, the molecular weight is 10.18 KDa. The theoretical pI of the protein is 10.07. Using same buffer conditions (CAPS buffer at pH 10.5 with 500mM NaCl) previously optimized for full length LUX purification, it was possible to obtain good amount of soluble LUX-DBD for purification. However there were considerable amount of impurities in the elution fraction after the 1st Nickel Sepharose Purification (**Figure 2-4a**). The TEV cleavage wasn't very successful as it was found that in the 2nd nickel column purification (**Figure 2-4b)** there were considerable amount of uncleaved LUX DBD that was present in the elution fraction. TEV protease works best at a pH range of 6-9. The elution buffer for LUX DBD was at pH 10.5, this could be a possible reason for incomplete TEV cleavage. Also another factor that could affect TEV cleavage is temperature, TEV protease works best in the range of 29-34°C. The mixture of purified LUX DBD and TEV protease was left at 4°C overnight, which could serve as another reason for incomplete cleavage.
An additional TEV cleavage was done for two hours at room temperature followed by an additional Ni-Sep Purification. Although the TEV cleavage was successful considerable amount of cleaved LUX DBD was also present in the elution fraction (**Figure 2-4c**).

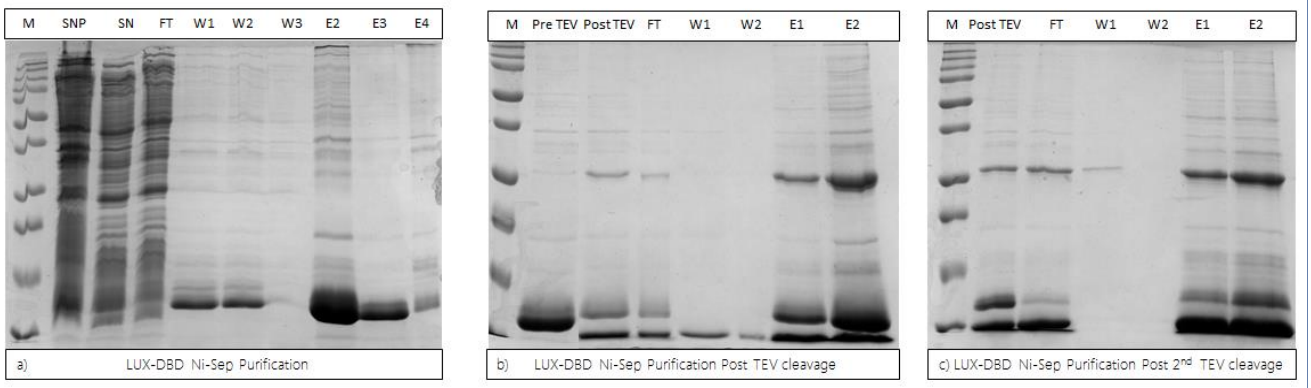



**Figure 2-4 Nickel –Sepharose Purification of LUX DBD.**

**a)** 1st Nickel Sepharose affinity column based purification of LUX-DBD. Lane M is marker, lane SNP is pellet obtained after centrifugation, lane SN is supernatant fraction, lane FT is flow-through fraction from 1st Ni-Seph purification, lanes W1 and W2 are wash fraction, and lanes E1 to E3 contain elution fractions
**b)** 2nd Nickel Sepharose Affinity column based purification to obtain 6x histidine tag removed LUX DBD.Lane PreTEV is LUX DBD without TEV treatment, lane Post TEV is LUX DBD after overnight TEV treatment, FT is flow-through fraction from 2nd Ni-Seph Purification, W1 and W2 are wash fraction while lane E1 and E2 are elution fraction. **c)** 2nd TEV cleavage reaction was carried out and a 3rd Ni-Sep purification was done to obtain greater amounts of His tag cleaved LUX-DBD. Lane post TEV is LUX DBD after 2 hours of TEV treatment at room temperature, lane FT is Flow through fraction, and lanes E1 and E2 are wash fractions while lanes 6 and 7 are elution fractions.

The samples obtained from the final Nickel-Sepharose purification were further separated on Heparin-Sepharose ™ column to obtain pure LUX-DBD Protein. It was possible to get rid of most contaminants and obtain pure LUX-DBD protein by Heparin Sepharose™ purification. Pure LUX-DBD obtained was concentrated upto 12.7mg/ml concentration and 50µL aliquots of LUX DBD at this concentration was flash frozen with liquid nitrogen and stored at -80ºC for use in future experiments.

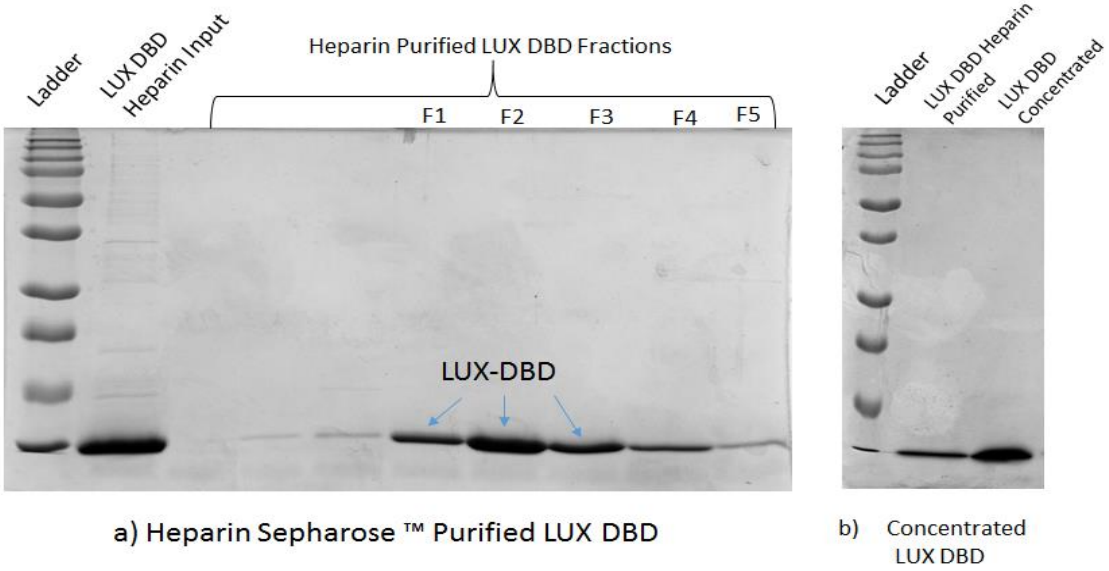



**Figure 2-5 Purification of LUX DBD over Heparin-Sepharose™ Column.**

**a)** Lane 1 is Dual Colour Protein Ladder, lane 2 is LUX DBD sample loaded for Heparin Sepharose purification. Lanes F1, F2, F3, F4 and F5 are samples fractions from the Heparin Purification. These samples were pooled together. **b)** Lane 2 is Heparin Purified LUX-DBD before concentration and Lane 3 is post heparin purification concentrated LUX DBD.

## 2.3 EARLY FLOWERING 4 (ELF4)

EARLY FLOWERING 4 (ELF4) is a small protein of 111 amino acids. ELF4 was also identified through genetic screen for photoperiod mutant. ELF4 promotes clock accuracy and is required for sustaining rhythms in the absence of daily light dark cycles. Mutation in ELF4 gene leads to attenuated expression of CCA1, suggesting that ELF4 is essential for the upregulation of CCA1. The *elf4* mutant has elevated levels of *TOC1* expression suggesting that ELF4 as a negative regulator of TOC1. Upon ELF4 overexpression it was seen that TOC1 levels were low and CCA1 and LHY levels were high. On the other hand the *cca1 lhy* double mutant show an increased expression of ELF4(Kikis et al., 2005). Collectively these data suggest a model where ELF4 has a central position in interconnected feedback loop. Hence apart from acting through the evening complex ELF4 also works as a core clock gene in driving morning expression of CCA1 and LHY (Kikis et al., 2005; McWatters et al., 2007). However, ELF4 lacks a DNA binding motif thus the mechanism through which it activates expression of *CCA1* and *LHY* is not known. ELF4 is also known sequester GI in the nucleus affecting GI binding to *CO* promoter (Kim et al., 2013). Hence ELF4 might be involved in a similar mechanism to remove repression of *CCA1* and *LHY* expression.

### 2.3.1 Disorder analysis and structure prediction for ELF4

Regions of protein disorder in *ELF4* were predicted using SPOT-disorder server (Hanson et al., 2017). The software predicted small regions of the N-terminal and C-terminal to have intrinsic disorder with disorder probability greater than 0.5. Region stretching from Val14 to Gly93 are predicted to be structured with a disorder probability less than 0.5 **(Figure 2-6a)**. Secondary structure prediction of *ELF4* was done using the PredictProtein server (Rost et al., 2004). Prediction algorithms suggested that the structured region consists of mainly helices **(Figure 2-6b)**.

Apart from structure prediction, previous structural studies done using far UV Circular Dichroism (CD) showed that ELF4 has a strong α-helical and disordered secondary structure. In the same study no β-sheet contributions were evident from CD spectra (Kolmos et al., 2009).

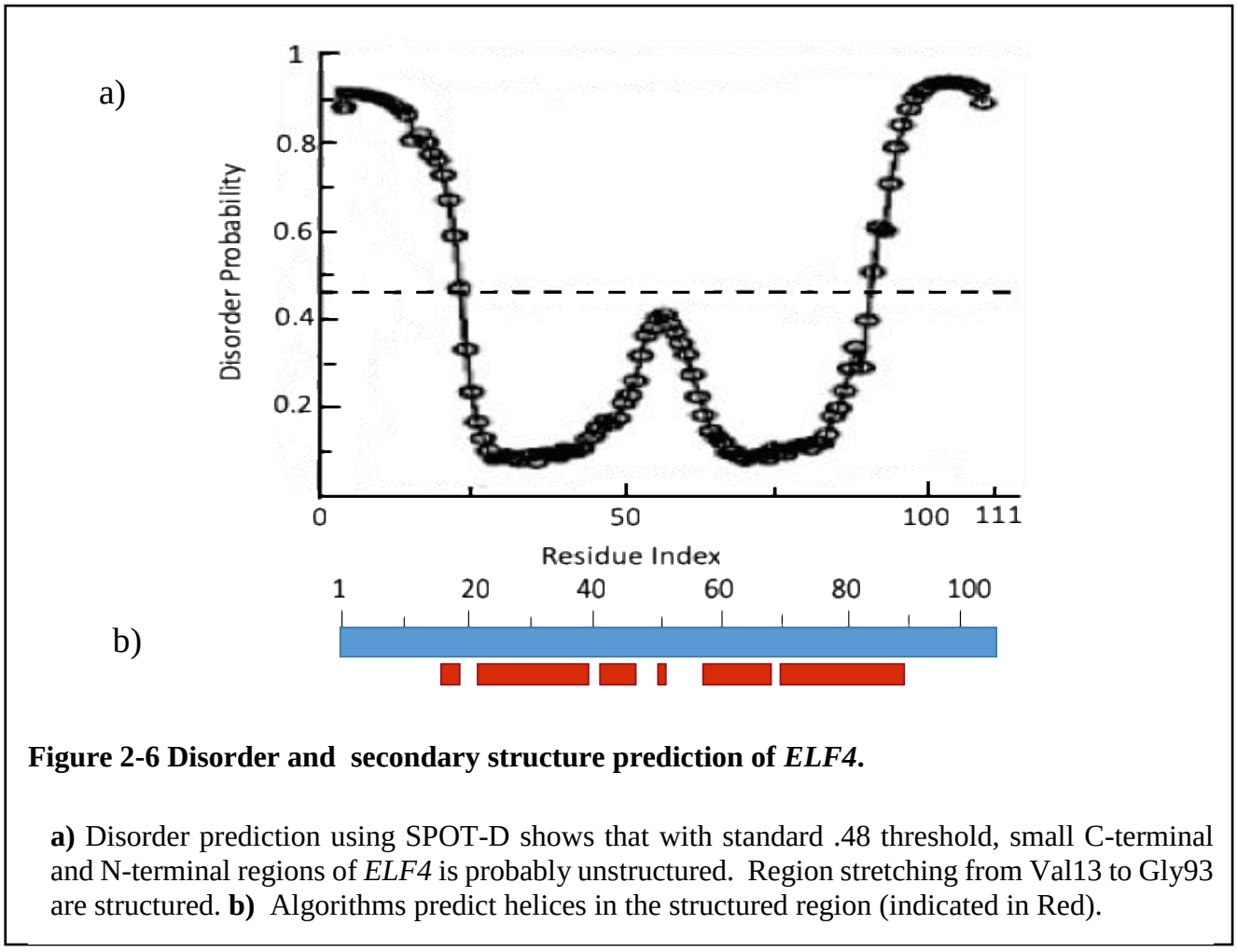



**Figure 2-6 Disorder and secondary structure prediction of *ELF4*.**

**a)** Disorder prediction using SPOT-D shows that with standard .48 threshold, small C-terminal and N-terminal regions of *ELF4* is probably unstructured. Region stretching from Val13 to Gly93 are structured. **b)** Algorithms predict helices in the structured region (indicated in Red).

### 2.3.2 Purification of ELF4 full length protein.

ELF4 is a small protein of 111 amino acid with a molecular weight of 12.37 kDa with a pI of 8.16. A 6x Histidine tag followed by a TEV peptide sequence is also added to the protein for aiding protein

purification. Hence with the 6xHis-tag and TEV peptide, the molecular weight of the expressed protein in *E.coli* turns out to be 15.32KDa. With initial trials in the Tris buffer at pH 8.0 and 300mM NaCl, it was possible to obtain soluble ELF4. The 20mM imidazole wash stringent enough to get rid of impurities that bound to the nickel column. However in the elution fraction it was observed that there was degradation products from the Full length ELF4 (**Figure 2-7a**).Precipitation was not observed in the elution fraction. The eluted protein was stable after overnight TEV cleavage to remove the 6xhistidine tag which was successful (**Figure 2-7b)**. However multiple degradation products were also observed apart from 6x Histidine cleaved ELF4. ELF4 Obtained was passed through Superdex®200 column for purifying ELF4 based on size exclusion chromatography.

The S200 purification using a HiLoad16/60 S200 column was used to separate degradation products from ELF4. An absorption peak was observed between the elution volume of 90ml to 110ml (**Figure 2-8a**). Elution fractions were separated every 2ml. A sample of 12µl from each fraction corresponding to the peak were resolved on a 15% SDS-PAGE to observe proteins purity of different fractions (**Figure 2-8b**). It was observed from the SDS- PAGE that there were degradation products that were also present in the S200 purification peak. The elution fractions from 92ml-100 ml had better concentration of Pure ELF4, hence they were pooled together and concentrated for future use in the *in-vitro* EC reconstitution experiments.

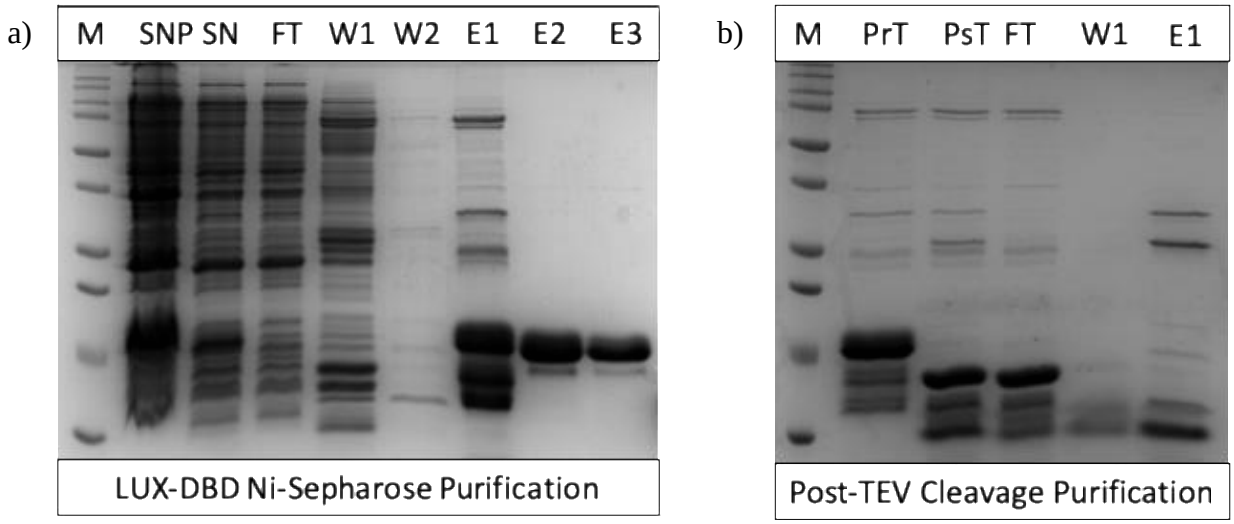



**Figure 2-7 Purification of ELF4**

**a)** 1st Nickel Sepharose affinity column based purification of ELF4. Lane M is marker, lane SNP is pellet obtained after centrifugation, lane SN is supernatant fraction, lane FT is flow-through fraction from 1st Ni-Seph purification, lanes W1 and W2 are wash fraction, and lanes E1 to E3 contain elution fractions. **b)** 2nd Nickel Seph purification to obtain 6x histidine tag removed ELF4. Lane PrT is ELF4 without TEV treatment, lane PsT is ELF4 after overnight TEV treatment, Lane FT is flow-through fraction from 2nd Ni-Seph Purification, Lane W1 is wash fraction while lane E1 is elution fraction.

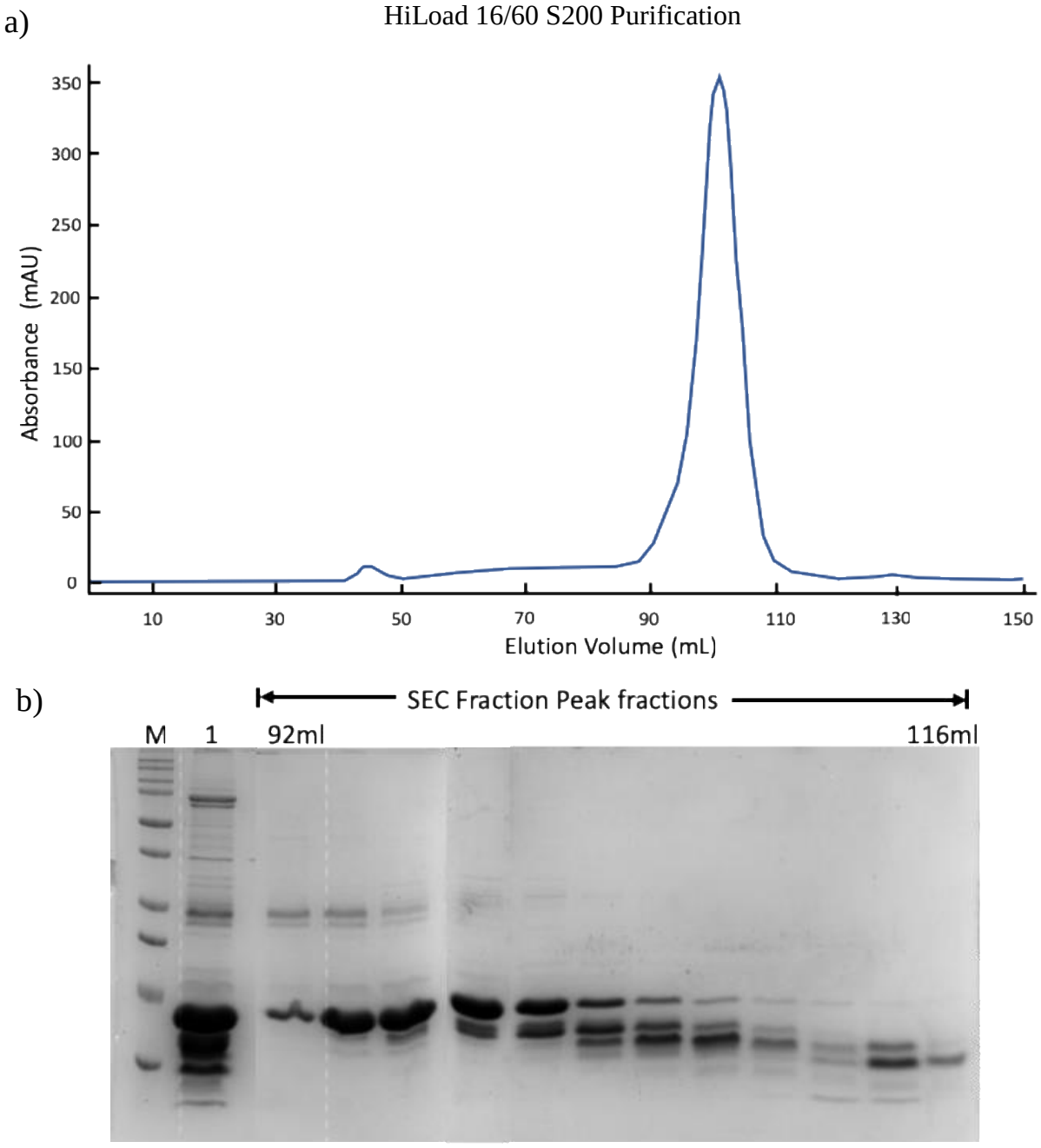



**Figure 2-8 S200™ Purification of ELF4**
ELF4 protein was purified over a Hiload 16/60 Superdex200® Column. **a)** Chromatogram from the S200 purification showing that ELF4 elutes between 90~ 110 ml elution volume. **b)** SDS-PAGE showing samples collected from the peak with an interval of 2 ml starting from 92 ml. Lane M is marker, lane 1 is input sample used for the S200 purification, rest of the lanes are samples from the S200 purification peak observed between 92-116 ml.

## 2.4 EARLY FLOWERING 3 (ELF3)

ELF3 was identified through genetic screens for photoperiod mutants and was mapped initially to the middle of chromosome 2. It was found to regulate circadian rhythms (Zagotta et al., 1996). ELF3 is a 695 amino acid long plant specific nuclear protein with no known functional domains (Hicks et al., 2001).Within the circadian network, ELF3 is highly interconnected and directly binds to multiple proteins such as PhyB, COP1, B BOX DOMAIN PROTIEN 19 (BBX19), PIF4, ELF4, LUX, BROTHER OF LUX (NOX), SHORT VEGETATIVE PHASE (SVP), TOC1 and GI **(Table 2-1).** This high degree of interaction of ELF3 suggests that it might be functioning as hub between different interactors (Huang et al., 2016). Classically ELF3 associated proteins are mainly involved in circadian clock pathways (ELF4, LUX and NOX) (Nusinow et al., 2011) and the light signaling pathway such as PhyB and COP1 (Liu et al., 2001; Yu et al., 2008).

**Table 2-1 List of Proteins Directly Interacting With ELF3. Abbreviations:** AA amino acid; AGI Arabidopsis Genome Initiative; Y2H Yeast Two Hybrid;

| Protein Name | AGI locus | Minimum ELF3 Fragment for ıction in Y2H Assays | Refs |
|---|---|---|---|
| **ELF4** | AT2G40080 | AA 261–484 | (Nusinow et al., 2011) |
| **LUX** | AT3G46640 | FL | (Nusinow et al., 2011) |
| **NOX** | AT5G59570 | FL | (Nusinow et al., 2011) |
| **GI** | AT1G22770 | AA 1–261 | (Yu et al., 2008) |
| **COP1** | AT2G32950 | AA 1–261 | (Yu et al., 2008) |
| **PhyB** | AT2G18790 | AA 1–440 | (Liu et al., 2001) |
| **PIF4** | AT2G43010 | AA 442–695 | (Nieto et al., 2015) |
| **TOC1** | AT5G61380 | AA 515–695 | (Huang et al., 2016) |
| **SVP** | AT2G22540 | FL | (Yoshida et al., 2009) |
| **BBX19** | AT4G38960 | FL | (Wang et al., 2015) |

### 2.4.1 Structure prediction and binding site analysis for ELF3

Regions of protein disorder in ELF3 were predicted using SPOT-disorder server (Hanson et al., 2017). The software predicted ELF3 to be largely unstructured, apart from that a small middle region (Pro318-Lys351) and a short fragment in the C-terminal region. Both these regions have a disorder probability of less than 0.5. Most of ELF3 looks disordered with no predicted domains (**Figure 2-9a)**.

**Purification of ELF3 full length protein.**

ELF3 is a protein of 695 amino acids and has a molecular weight of 77KDa. The theoretical pI of ELF3 full length protein is 8.63. In our preliminary test for expression it was found to be insoluble and

remained in the pellet fragment of the lysed cells after centrifugation. Hence it was decided to do an ESPRIT library to find soluble domains of ELF3 that could be expressed and purified.

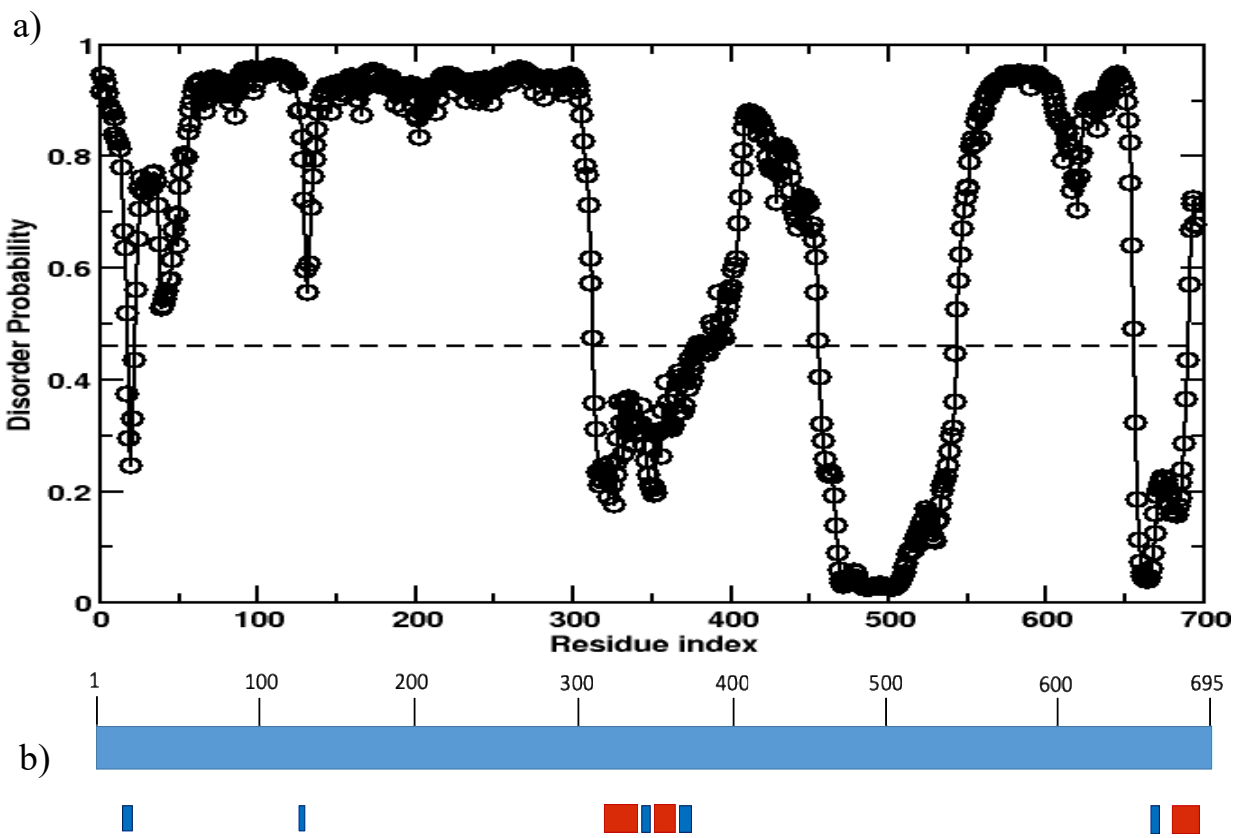



**Figure 2-9: Secondary Structure and Binding Site prediction of *ELF3*.**

**a) Disorder prediction using SPOT-D shows that with standard .48 threshold, small stretch in the middle and small N-terminal regions of ELF3 is probably structured. Regions stretching from Pro318-Lys369 and Ala672-arg689 are predicted to be structured. b) Secondary structure prediction algorithms predicts the presence of short α-helices (in red ) and strands (in dark blue).**

**ELF3 Co-ESPRIT library to obtain Soluble ELF3 Fragments.**

High throughput soluble expression library construction (ESPRIT) method was applied for obtaining soluble fragments of ELF3(Yumerefendi et al., 2010). ESPRIT uses limited exonuclease treatment to generate a library of truncation constructs. From the library 43 constructs were obtained for soluble ELF3 fragments. Upon scale-up, we were able to purify 11 soluble constructs (**Figure 2-10b**). Of these 11 soluble constructs, there were 8 unique soluble fragments (**Table 2-2**). When aligned against full length ELF3, it was found that the middle domain of ELF3 was not obtained as a soluble fragment. Obtained soluble fragments spanned the N-terminal and C-terminal. The soluble fragments were used for in-vitro EC reconstitution experiments.

| No. | Construct name | Amino acid length | Fragment on ELF3 full length |
|---|---|---|---|
| **1.** | ELF3 #01 | 238 | L388-E625 |
| **2.** | ELF3 #02 #40 #46 | 204 | R048-D251 |
| **3.** | ELF3 #03 | 225 | K035-H259 |
| **4.** | ELF3 #20 #9 | 127 | M503-N630 |
| **5.** | ELF3 #28 | 235 | G053-A287 |
| **6.** | ELF3 #34 | 157 | E040-A196 |
| **7.** | ELF3 #37 | 210 | N423-P633 |
| **8.** | ELF3 #47 | 209 | R117-L325 |

**Table 2-2 ELF3 soluble fragments**
ELF3 constructs generated from the ESPRIT library with their amino acid length and their corresponding amino acid fragments aligned on ELF3 full length.

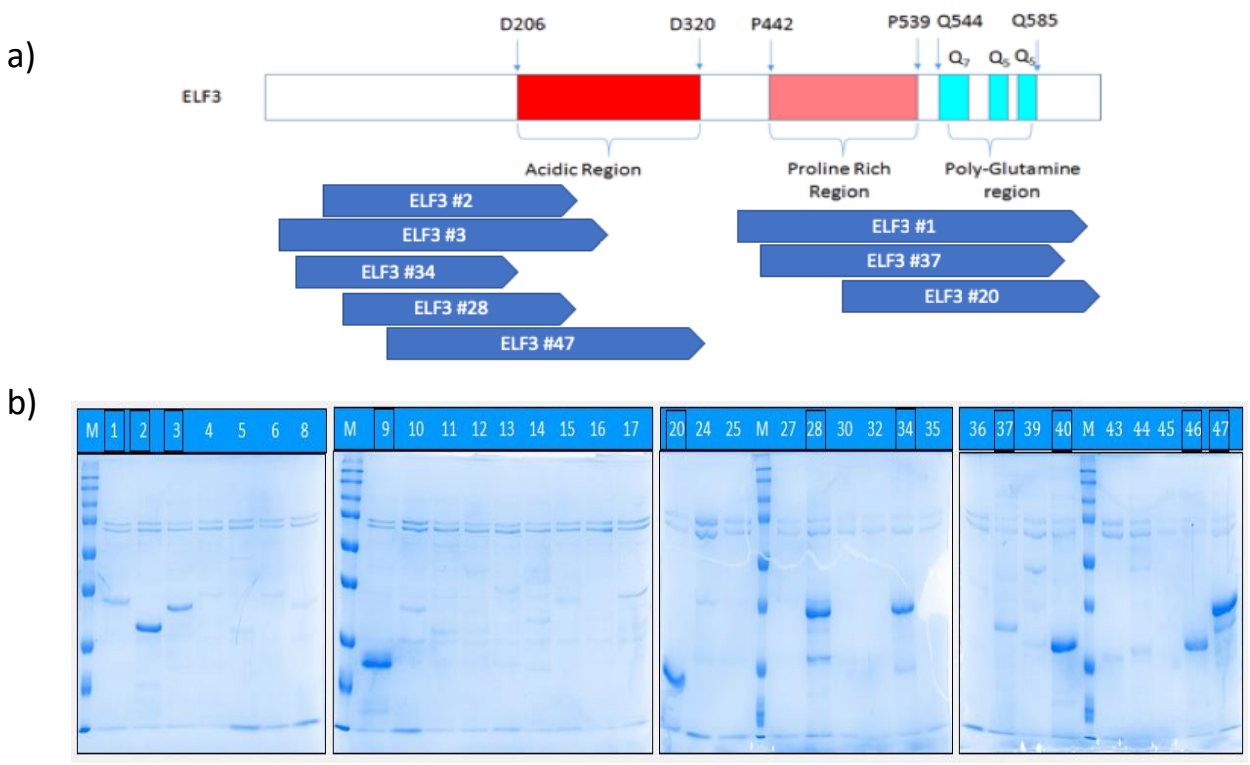



**Figure 2-10 ELF3 Soluble fragments obtained from Co-ESPRIT library.**

a) Schematics alignment of soluble fragments of ELF3 generated from the ESPRIT library aligned with the ELF3 Full length protein which is 695 amino acids long. The eight soluble fragments obtained are aligned with respect to their amino acid sequence corresponding to the full length. (b) Small scale expression test of soluble fragments of ELF3 generated from ESPRIT library resolved on 5% SDS-PAGE. M indicates dual color protein marker and the numbers indicate different constructs from the ELF3 ESPRIT library generated. The soluble fragments are marked with a black box on the number.

The soluble fragments generated from the ESPRIT library was missing the middle domain of ELF3 which has been reported to be important for binding of ELF4 to ELF3(Herrero et al., 2012) and might be important for forming the EC. Hence it was important to obtain full length ELF3 in order to perform *in-vitro* binding assay with all EC components. Since ELF3 full length expression was not successful in *E.coli* cells, it was decided to express ELF3 in insect cells alone and with its EC partners, ELF4 and LUX in an effort to stabilize the complex . From the initial small scale purification experiments it was clear that there was high expression of ELF3 in insect cells and the expression increased every 24 hours (**Figure 2-11b**). However, there was a problem with solubility as seen from the gels, it was clear that expressed ELF3 was present in the pellet fraction suggesting that ELF3 was not soluble.

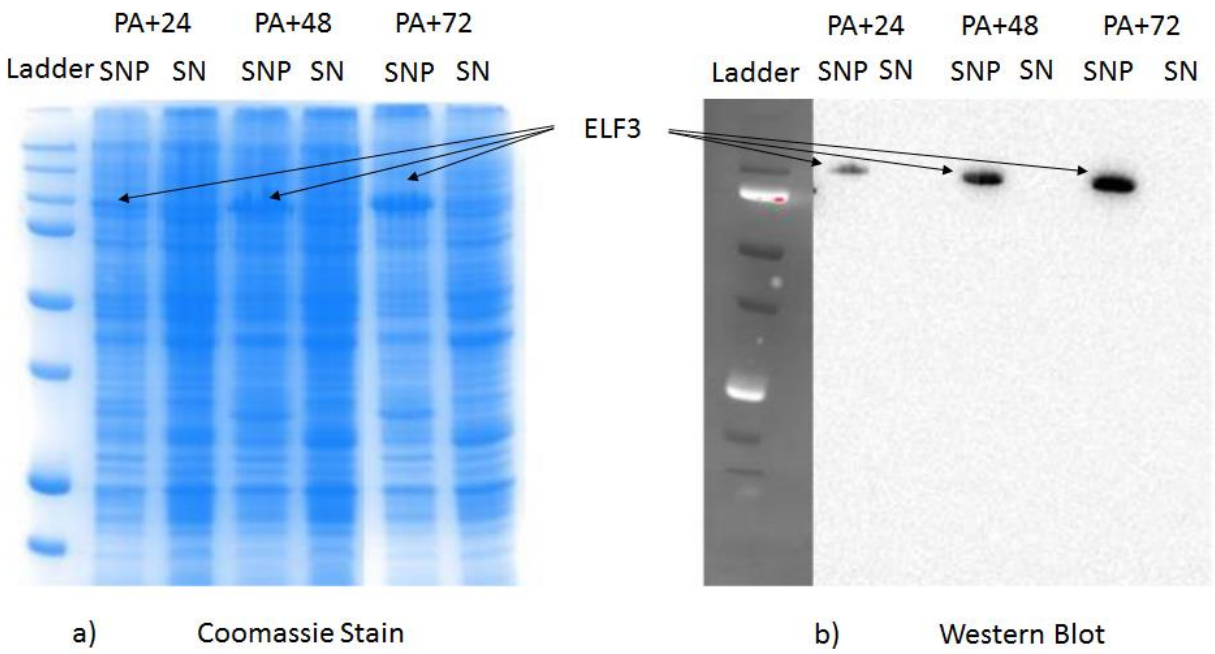



**Figure 2-11 Small Scale ELF3 Insect cell expression test for solubility and expression quantity.**

a) Samples from Pellet (SNP) and Supernatant (SN) collected every 24 hours after Proliferation Arrest (PA) were resolved on a 15% SDS gel and were stained by Coomassie Brilliant Blue (CBB). Lane 1 is ladder and rest are samples collected after proliferation arrest as indicated. b) Western Blot of samples used in CBB staining using α-His antibodies to detect 6x-His tagged ELF3. Lane 1 is ladder and the rest are samples collected after proliferation arrest.

**Buffer optimization for purification of ELF3 from insect cells.**

Since ELF3 is insoluble in Phosphate buffer saline (PBS buffer), it was necessary to find a buffer that would be suitable for ELF3 extraction and purification . Different buffer conditions spanning a pH range from 4-10.5, varying salt concentrations and additives were tried. However, these trials were unsuccessful. An alternative strategy was adopted to purify ELF3 using 8M urea with a rapid refolding protocol. In this protocol, 8M urea was used to solubilize the protein and the protein was then purified via a Ni-Sepharose column. Contaminants were removed with a wash step with 8M urea and a second

wash was done using 0M Urea to rapidly refold the protein on the column. The protein was then eluted with 200 mM imidazole under non-denaturing conditions.

ELF3 obtained in the first elution fraction of the first Ni-Seph purification had protein contaminants while the second elution fraction was relatively pure. (**Figure 2-12a**). ELF3 obtained was buffer exchanged to a buffer containing 40mM CAPS pH9.7 and 100mM NaCl. Western Blot was performed before against the 6xHis tag with α-His antibodies to confirm ELF3 (**Figure 2-12b**). The 6xHis tag was removed through cleavage by TEV protease overnight at 4°C. The cleaved protein was applied to a 2nd Ni-Sepharose Column to relmove TEV protease and uncleaved protein (**Figure 2-12c**). Hence ELF3 obtained from flow through. Size exclusion chromatography was performed to obtain pure ELF3. In the Size exclusion chromatography profile, a major peak was observed at 8.62mL of elution volume. 3ml of sample was collected corresponding to the major peak observed at elution volume between 6.7ml to 9.5mL (**Figure 2-13a**). Samples obtained were resolved on a 15% agarose gel to look at the purity of the sample and size of the elutant to confirm that the elutant was ELF3 (**Figure 2-13b**). Based on gel filtration results, ELF3 seems to appear as higher order oligomer or soluble aggregates.

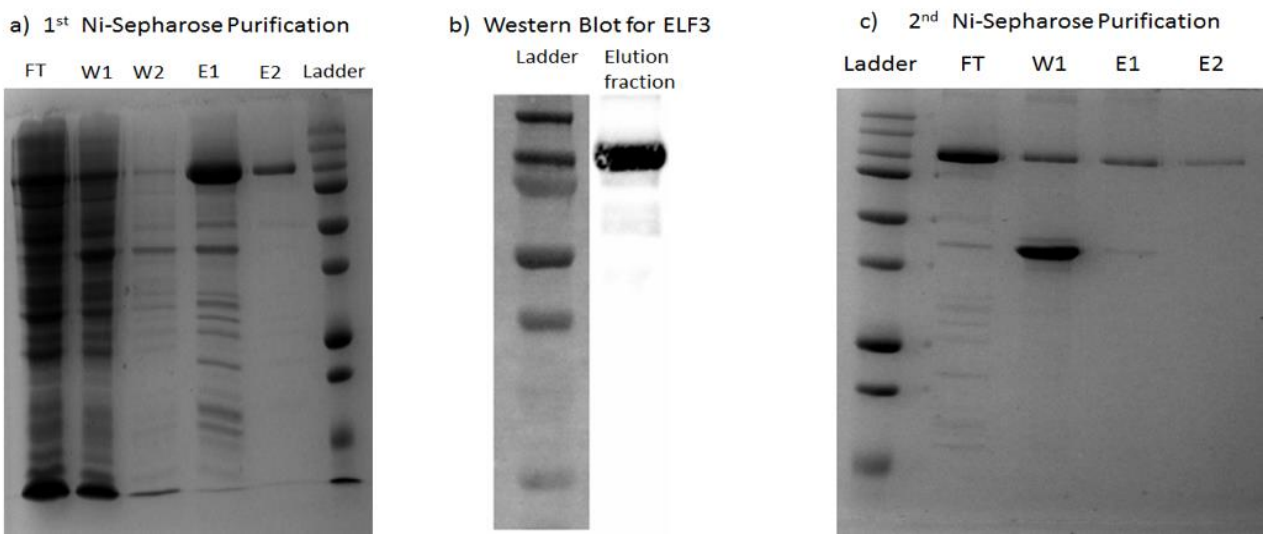



**Figure 2-12 Purification of ELF3 from insect cells**

a)Samples from different purification fractions of 1st nickel column purification resolved on 16% SDS PAGE Fragments from the 1st nickel. b)Western blot of elution fraction using anti his antibodies to stain against 6xHis ELF3 protein produced in the insect cells. c) Samples from different purification stages from the 2nd Nickel column purification after the TEV protease treatment to remove 6xHistidine tag.

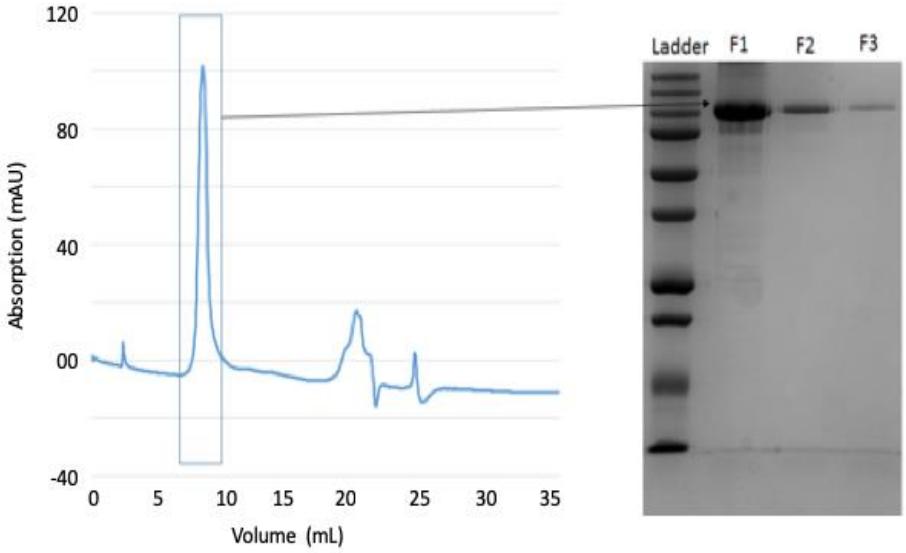



**Figure 2-13 S200 purification of ELF3**

Elution profile of ELF3 FL purified on S200 Column. A peak corresponding to 100 mAU was observed at ~8mL elution volume.

## 2.5 *In-vitro* EC reconstitution

Once all components of the EC were expressed and purified, experiments were performed to reconstitute EC using its constituent proteins *in vitro*. Since purified ELF4, LUX and different fragments of ELF3 (corresponding to the N-terminal and C-terminal) were available, it was decided to incubate them together and perform size exclusion chromatography to verify if a complex could be formed *in vitro*. In this section initially the results pertaining to the EC reconstitution using soluble ELF3 fragments, ELF4 and LUX will be presented.

To test if a core EC could be formed reactions were carried out with 20µM each of ELF3#28(N-terminal) and ELF3#1(C-terminal ) along with 50 µM each of LUX and ELF4 in a final volume of 500 µL. The reaction was incubated at room temperature for 45 minutes. After the incubation, the reaction mixture was passed over an S200 column for purifying the complex. 2 peaks were observed from the S200 run (**Figure 2-14a**). The first peak was observed at an elution volume of 13.66 ml while the second peak was observed at an elution volume of 17.8 ml (**Figure 2-14a**). The peaks correspond to molecular weight of approximately 158 KDa and 29 KDa.

The elution samples collected from these fractions were concentrated and resolved on an 18 % SDS gel (**Figure 2-14b**). Western blot was done using antibodies against the 6x histidine tag which is present on all the proteins that were used for the experiment (**Figure 2-14b**). From the western blot it became clear that in the first peak there are two ELF3 fragments along with excess of LUX eluting while in the second peak there is excess ELF4 and some amount of ELF3 fragments and LUX. Also degradation of at least one of the proteins was observed. Based on size exclusion no peak on higher molecular weight was observed, suggesting that the fragments tested were not suffiecient to form a complex.

An alternative method of refolding all three proteins together from 8 M urea was used to obtain a complex. In this protocol equimolar amounts of all three proteins were combined and step-wise dialysis from 8 M to 0M urea was performed. This protocol was successful to form small amounts of active EC, which was used for EMSA experiments. EMSA experiments using a probe from the *PRR9* promoter with an EC binding site were performed. As shown below, ELF4 and ELF3 alone didn't interact with DNA (**Figure 2-15a left**), this was expected as both of these proteins lack any predicted DNA binding domains. LUX on the other hand binds DNA and this binding is not affected by addition of ELF4. However, addition of ELF3 leads to disappearance of the LUX-DNA band and appearance of a higher molecular weight band that corresponds to the EC bound DNA **(Figure 2-15a left)**. Interestingly, when ELF3 is added to a solution containing LUX and its target DNA without ELF4, it was seen that addition of ELF3 attenuates binding ability of LUX to its target DNA **(Figure 2-15a right)**. No bands of higher molecular weight were observed when ELF3 was added to LUX-DNA complex. This suggests that ELF3 impedes LUX binding to its target and that ELF4 is essential for restoring DNA binding capability of the EC **(Figure 2-15b)**. A similar observations has been made for ELF3-PIF4 interaction, where ELF3 prevents DNA binding activity of PIF4 by sequestering PIF4 (Nieto et al., 2015).

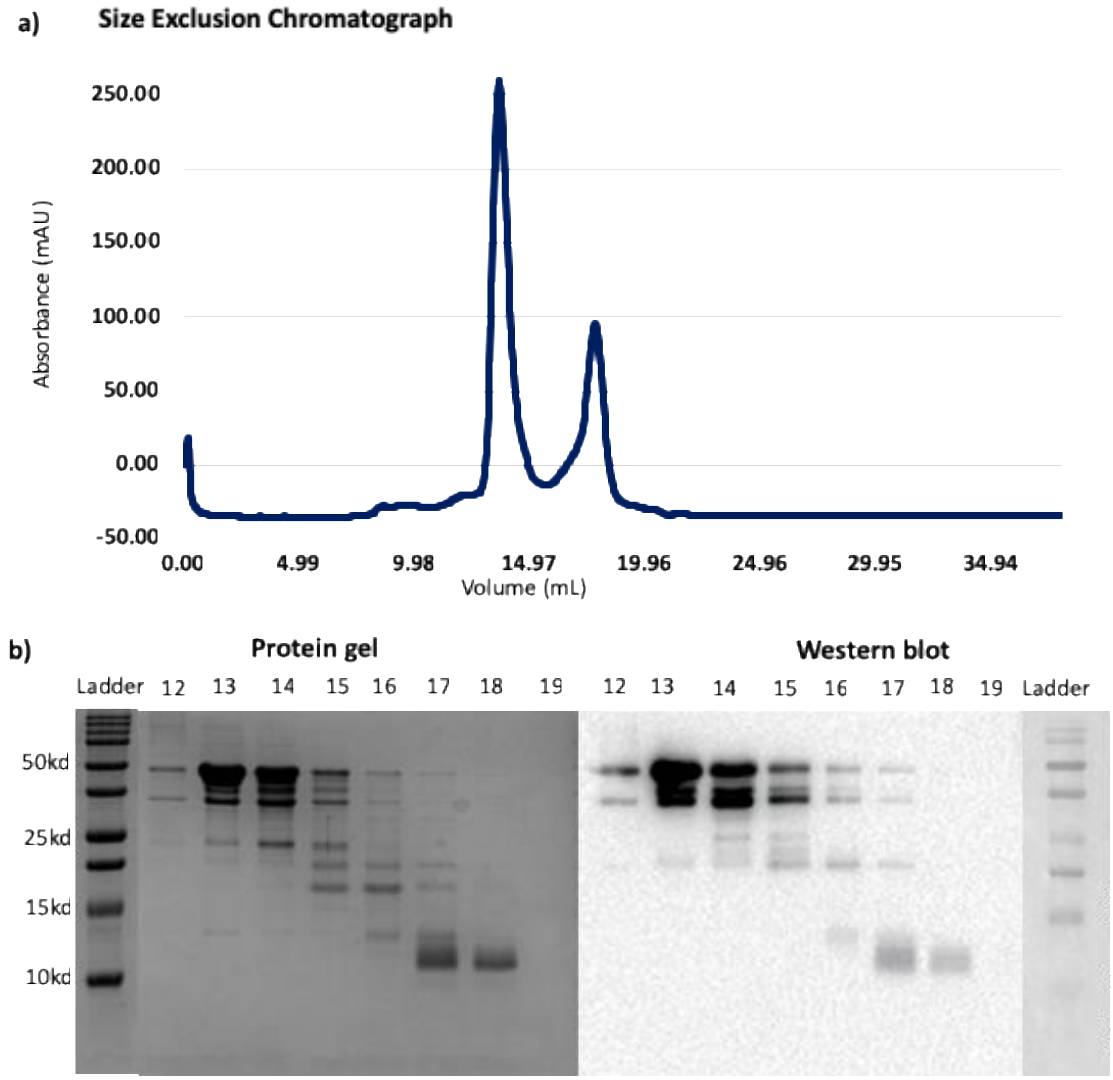
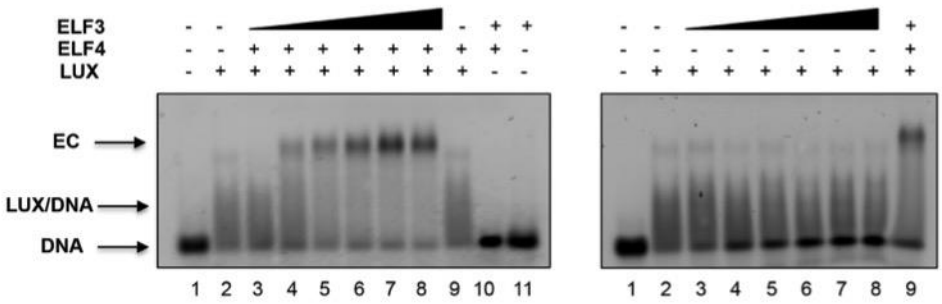



**Figure 2-14 *In Vitro* EC reconstitution trial.**

**a)**Size exclusion chromatograph reveals there are two peaks that emerge when the *in vitro* reconstituted EC is passed through an S200 column. The first peak is observed at elution volume of 13.66 ml while the second peak is observed at 17.81ml. **b)** Protein gel and western blot of elution fraction obtained from S200 purification. Numbers on the lane indicate the fraction volume from which samples were taken. An antibody against 6x hisitidine was used in the western blot to blot the EC components.

**a)**




**b)**

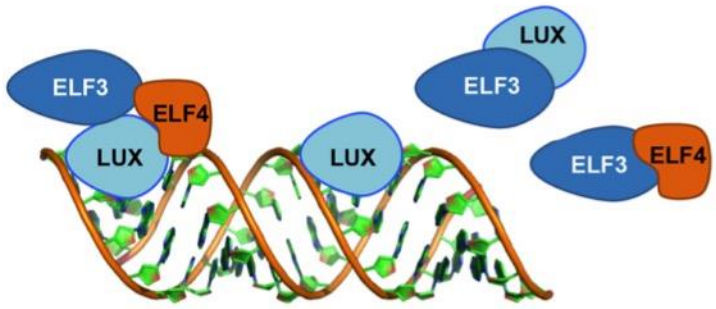



**Figure 2-15 EC formation *in-vitro* and interaction with DNA.**

**a)** EMSAs of LUX-ELF3 and the EC, in 2% agarose gels. DNA concentration was 30 nM. (Left) Reconstitution of the EC with LUX and ELF4 concentrations held constant at 200 nM and 1000 nM, respectively and increasing ELF3 concentrations. Lanes are (1) DNA alone, (2) DNA+LUX, (3-8) DNA+LUX+ELF4 with increasing concentrations of ELF3 (220 nM, 450 nM, 890 nM, 1.3 µM, 1.8 µM and 2.2 µM) (9) DNA+LUX+ELF4, (10) DNA+ELF3+ELF4 and (11) DNA+ELF3 at 1.8 µM. (Right), LUX-ELF3 interactions, with LUX concentration kept at 200 nM. Lanes are (1) DNA alone, (2) DNA+LUX, (3-8) DNA+LUX with increasing concentrations of ELF3 (as per left panel) (9) DNA+LUX+ELF4+ELF3 (EC; with 1 µM ELF4 and 1.8 µM ELF3). With increasing ELF3 concentration, the free DNA band increases in intensity. **b)** Schematic depiction of LUX, ELF3 and ELF4 interactions and DNA-binding. LUX is depicted in light blue, ELF3 in dark blue and ELF4 in orange.

## 2.6 Conclusion

This chapter dealt with the production of the EC components. LUX, LUX-DBD and ELF4 were all purified to homogeneity. ELF3 presented challenges for both expression and solubility. Bacterial and insect cell expression was unsuccessful. Construct library generation yielded multiple soluble constructs, however these constructs did not interact with LUX and/or ELF4. A refolding protocol was developed that allowed us to obtain small amounts of active EC. This complex was used in DNA binding assays. These assays allowed us to understand the role of each protein in complex formation and DNA-binding. These results demonstrate that LUX is able to target the entire EC site specifically to DNA and that ELF4 is required to maintain LUX's DNA binding capacity. ELF3 is able to bind LUX, however the LUX-ELF3 complex exhibits reduced ability to bind DNA, possibly due to occlusion of the DBD by ELF3. This sequestering activity of ELF3 has been observed for another ELF3 partner, PIF4. In order to more fully understand the specificity determinants of DNA-binding, the LUX DBD was studied using protein crystallography as described in the following chapter.

## 2.7 Methodology

### 2.7.1 ELF3, ELF4, LUX and LUX DBD vector design and expression tests in *E.coli*.

**Cloning of ELF4, ELF3, LUX and LUXDBD**

CDS sequence for ELF4, ELF3, LUX and LUXDBD were cloned into the expression vector pESPRIT002 using the AatII and NotI restriction sites (Tarendeau et al., 2007). All constructs contained a TEV protease cleavable N-terminal 6x-His tag. The vector maps of the respective plasmids can be found below**(Figure 2-16)**.These constructs were tested for small-scale expression tests in *E.coli*.

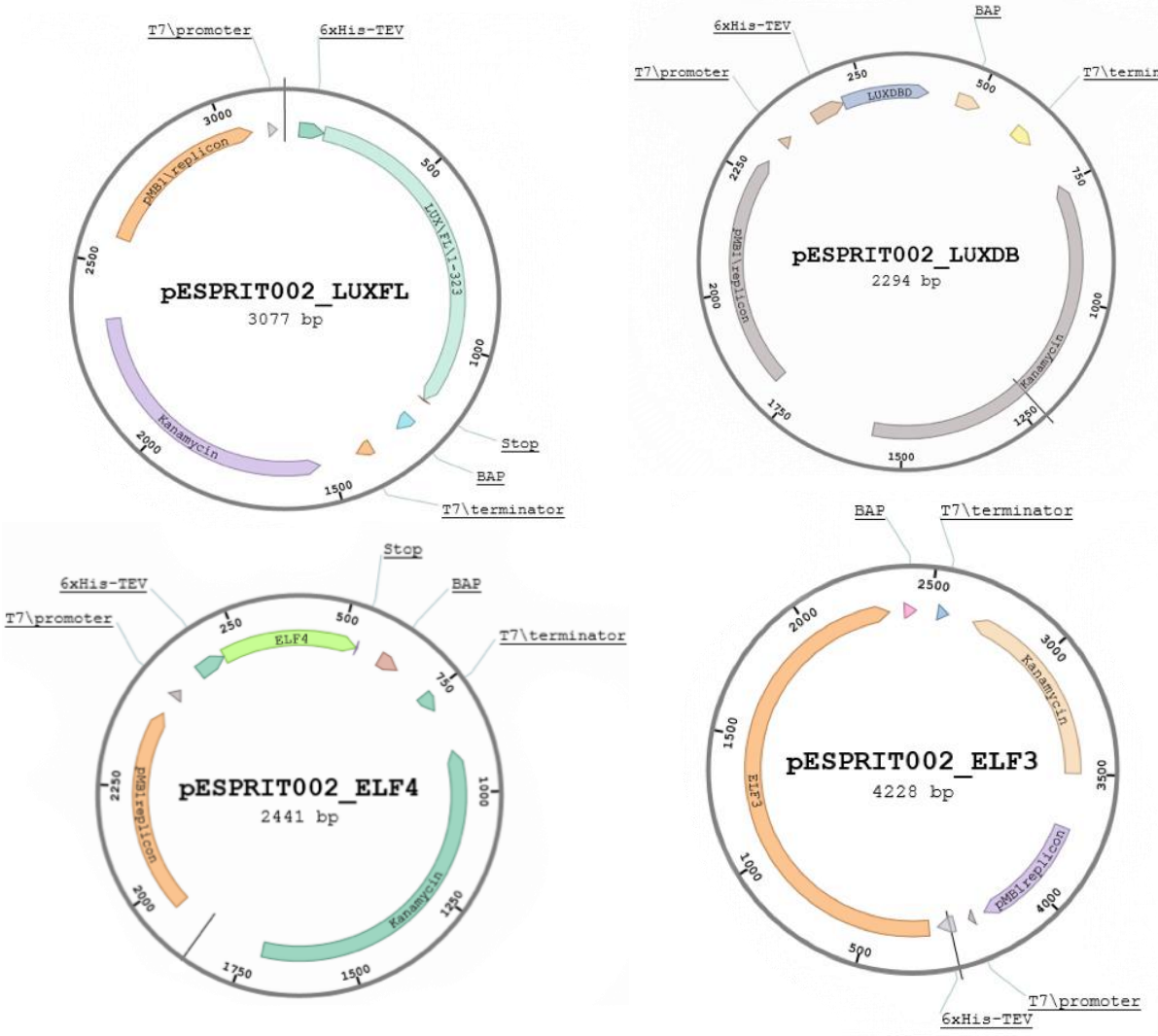



**Figure 2-16 ELF3, ELF4, LUX and LUX DBD expression constructs.**

**Protein Expression in *E.coli***

All constructs were overproduced in *Escherichia coli* (*E.coli*) Rosetta2 (DE3) pLysS cells. For small-scale expression tests, cells were grown at 37ºC in 1L Luria-Bertani (LB) culture medium supplemented with chloramphenicol (37 mg/mL) and kanamycin (50 mg/mL). Cells were grown until an $OD_{600}$ of 0.7-0.8 was reached. Once an $OD_{600}$ of 0.7-0.8 was reached, the temperature was reduced to 20ºC and protein expression induced by addition of 1 mM isopropyl-β-D-thiogalactopyranoside (IPTG); expression was continued overnight (~16 h) and the cells were harvested by centrifugation for 30 min. at 6000 rpm and 4ºC.

**Protein purification and buffer condition optimization.**

For small-scale purification tests we started with the buffers **(Figure 2-17)** mentioned below for different stages of protein purification.

**Resuspension buffer**

| Component | concentration |
|---|---|
| Tris pH8.0 | 20 mM |
| NaCl | 500 mM |
| TCEP | 1 mM |
| Bezonase | 5µL/50ml of buffer |
| Protease Inhibitor Cocktail | One tablet in 50ml of buffer |

**Wash Buffer**

| Component | concentration |
|---|---|
| Tris pH8.0 | 20 mM |
| NaCl | 500 mM |
| Imidazole | 20mM |
| TCEP | 1 mM |

**Elution buffer**

| Component | concentration |
|---|---|
| Tris pH8.0 | 20 mM |
| NaCl | 500 mM |
| Imidazole | 200mM |
| TCEP | 1 mM |

**Dialysis buffer**

| Component | concentration |
|---|---|
| Tris pH8.0 | 20 mM |
| NaCl | 300 mM |
| TCEP | 1 mM |

**S200 Buffer**

| Component | concentration |
|---|---|
| Tris pH8.0 | 20 mM |
| NaCl | 100 mM |
| TCEP | 1 mM |

**Figure 2-17 Buffers used for purification of proteins**

Harvested cells were resuspended in resuspension buffer. Cells were vortexed and agitated in 30mL of Resuspension Buffer and lysed by sonication. Cell debris were removed via centrifugation for 40 min. at 18000 rpm and 4ºC and the supernatant applied onto a 1 mL Ni-Sepharose High-Performance resin column, pre-equilibrated resuspension buffer. The column was washed with 25 CV of wash buffer and the protein eluted with elution buffer. Fractions of interest were pooled and dialyzed overnight at 4ºC

against dialysis buffer and in the presence of 2% (w/w) TEV protease, in order to cleave the N-terminal 6xHis tag. The protein sample was then passed onto a second Ni-Sepharose column in order to deplete 6xHis tagged TEV protease and uncleaved protein and subsequently applied to a size exclusion Superdex 200 Hi-Load 16/60 column (GE-Healthcare), pre-equilibrated with S200buffer (20 mM Tris pH 8.0, 100 mM NaCl, 1 mM TCEP). Protein fractions were pooled and concentrated. This purification protocol was successful for ELF4. Micrograms quantities of LUX and LUX DBD could be obtained using these buffers. However, for larger scale purifications Tris was replaced with 200 mM CAPS, pH 10.5 for Ni-NTA column purification and dialysis buffer replaced with 50 mM CAPS, pH 9.70 .

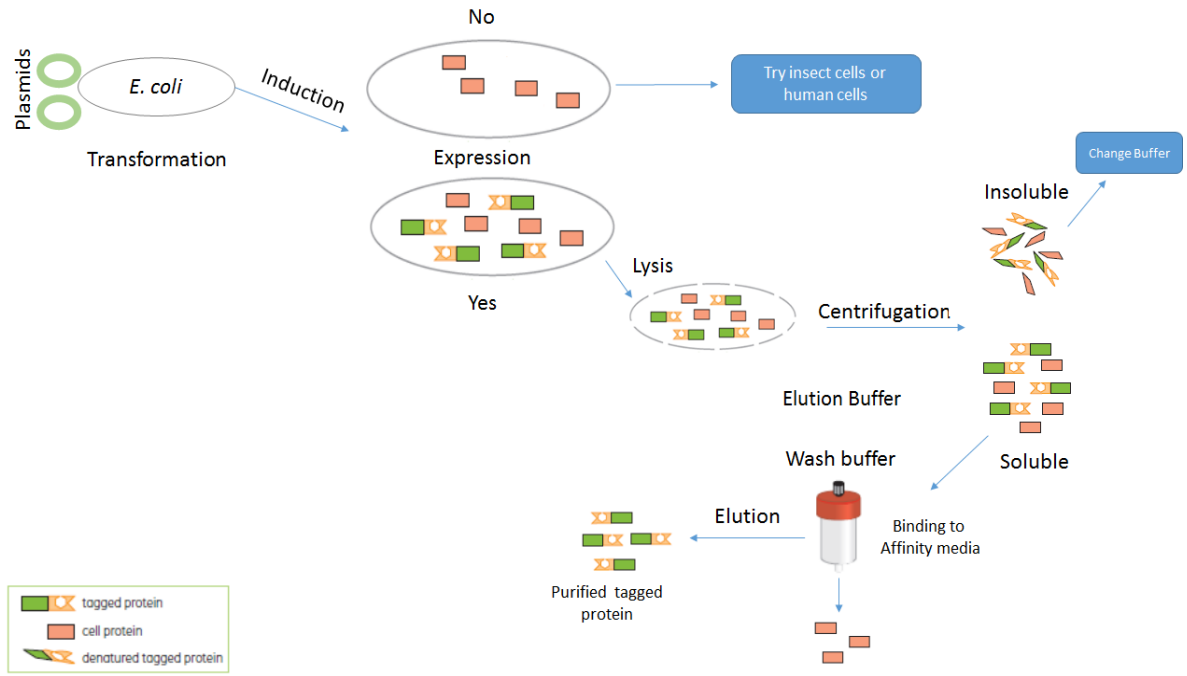



**Figure 2-18 Process diagram for purification of EC components.**

### 2.7.2 Soluble ELF3 Construct Generation BY ESPRIT technology.

**Generation of random construct library**

As ELF3 did not express in E. coli, a construct library using ESPRIT was generated in order to obtain soluble protein. ESPRIT uses exonuclease III based degradation for the generation of nested deletions of different regions in gene of interest (Henikoff, 1984; Ostermeier and Lutz, 2003). The ELF gene was cloned into a vector encoding a C-terminal biotin acceptor peptide and, at the end of the insert to be truncated, a pair of restriction sites that leave exonuclease III sensitive 5′ and resistant 3′ overhangs. The vector also harbors an N-terminal hexa-histidine tag to facilitate direct purification testing of constructs. After the initiation of the exonuclease III reaction, small aliquots are removed at interval of 1min for an hour and pooled in a quenching solution of 2M NaCl (Tarendeau et al., 2007). Single stranded overhangs are removed by using Mung bean nuclease and then T4 DNA ligase was used to re-circularize the vector.

**Colony blot analysis of protein expression**

ELF3 expression was induced in a high-density colony array format of 56,000 colonies, 28,000 in duplicate. Colony blots were prepared by in situ lysis on nitrocellulose membranes (Bussow et al., 1998) and blots are probed with fluorescent streptavidin to detect the biotinylation status of the C-ter biotin acceptor peptide as an indicator for solubility. An additional quality filter was introduced by simultaneously probing the array with a monoclonal antibody against the N-ter hexahistidine tag; only clones exhibiting both N- and C-ter tags by colony blot analysis were analysed further since these comprised un-degraded, intact protein constructs.

**Protein expression and purification from liquid cultures**

Isolation of the most efficiently biotinylated constructs leads to an enrichment of putatively soluble clones that, in the second level of screening, were expressed in small-scale liquid format, lysed and purified on Ni2+ NTA affinity resin. Soluble, purifiable constructs were visualized by SDS–PAGE and their sequence boundaries determined by DNA sequencing.

### 2.7.3 ELF3 Full length and EC expression using insect cells.

#### Cloning of ELF3 full length and EC into insect cells

ELF3 was cloned into pACEBac1 using SalI and NotI restriction sites with a TEV protease cleavable 6x His-tag at both N-terminal and C-terminal ends of the gene. ELF4 was cloned into vector pIDS using NcoI and XhoI restriction sites while LUX was cloned into Vector pIDK using KpnI and XhoI restriction sites (Vijayachandran et al., 2011). For co- expression of all three protiens, Cre-LoxP recombination was performed to fuse pIDK-LUX and pIDS-ELF4 to pACEBac1-ELF3 **(Figure 2-19**)

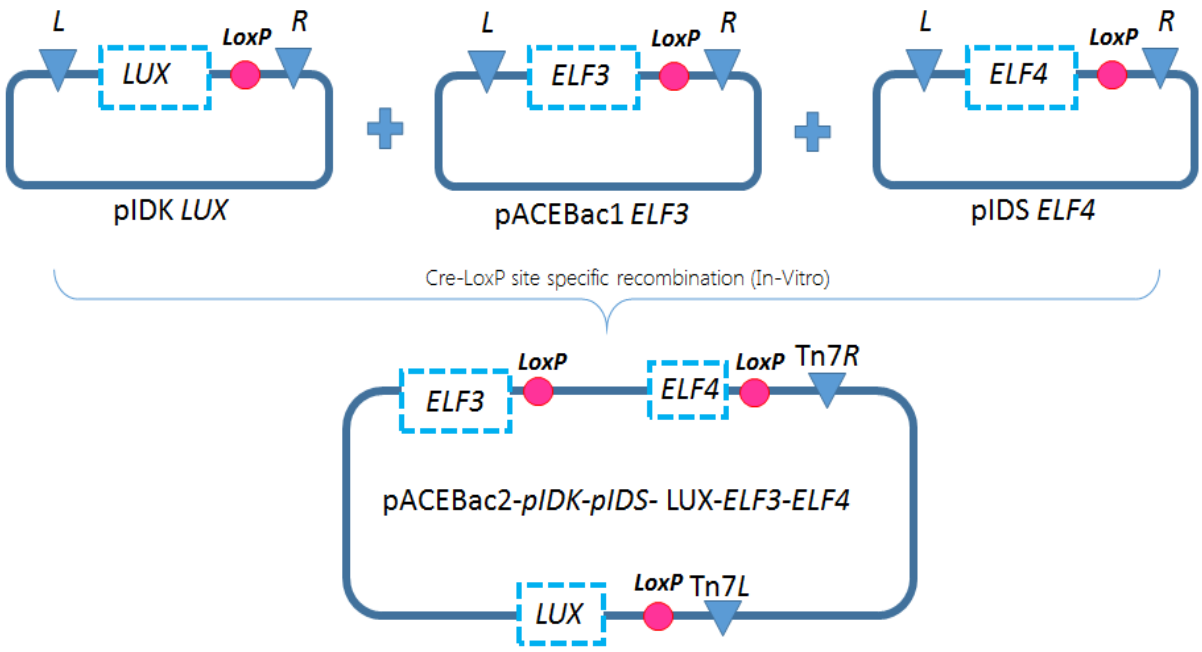



**Figure 2-19 Cloning ELF3, ELF4 and LUX for expression in insect cells.**

#### Expression of ELF3 full length and EC in insect cells

ELF3 and EC co expression constructs were produced in *Sf*21 insect cells (Invitrogen) using the baculovirus expression system. Briefly, the generated construct was transformed into chemically competent DH10 EmBacY cells (harbouring the baculoviral EmBacY genome) (Bieniossek et al., 2008; Trowitzsch et al., 2010). Positive clones were identified by blue/white screening in the presence of IPTG and BluoGal and used for downstream bacmid isolation. *Sf*21 insect cells were transfected at a density of $0.3x10^6$ cells/mL in a 6-well plate format (**Figure 2-20**). Primary baculovirus stock ($V_0$) was harvested 60 h after transfection and used for infecting 25 mL new *Sf*21 cell cultures yielding $V_1$ stock. $V_1$ baculovirus stock was collected every 48h after cell proliferation arrest, stored at 4ºC and used to launch ELF3 expressions ($V_2$) (500 mL cell cultures at $1.0x10^6$ cells/mL with 0.1% (v/v) $V_1$ virus). Amplification of the virus and protein expression were followed by monitoring YFP (Yellow

Fluorescent Protein) expression from the viral backbone. 20 mL Samples were collected every 24h hours after proliferation arrest (PA) and cells were harvested (2000g, 15 min, 4ºC) in PBS buffer and sonicated. After sonication the supernatant and pellets were recovered and resolved on 15% SDS gel.

**Purification of ELF3 and EC from Insect cells.**

Pellets from insect cells expressing ELF3 and EC were resuspended in 5x volume of resuspension buffer and were lysed by freeze/thaw cycles in liquid nitrogen and ice/water bath. The lysate was centrifuged for 1h at 18000 rpm in 4ºC. The supernatant obtained after centrifugation was applied to a 0.5mL Ni-Sepharose High-Performance resin column, pre equilibrated in resuspension buffer. The column was washed with 15 CV volumes of wash buffer. Afterwards, on column rapid refolding protocol was followed to obtain ELF3 in a urea free buffer. For this, the column was washed with 5 volumes of wash buffer 2 and then protein was subsequently eluted using elution buffer. Fractions of interest were pooled together and were used for downstream characterization experiments. All buffers needed for the purification are listed below **(Figure 2-21).**

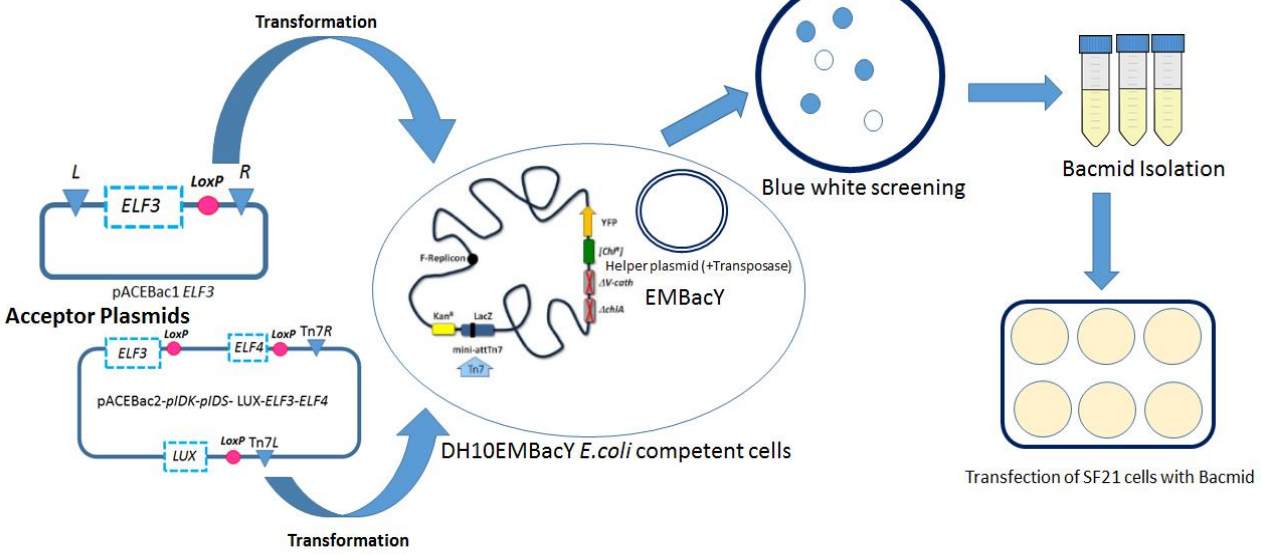



**Figure 2-20 Process Diagram for Expression of ELF3 and EC in insect cells**

**Resuspension Buffer**

| | |
|---|---|
| Urea | 8M |
| NaCl | 300mM |
| TCEP | 1mM |

**Wash Buffer**

| | |
|---|---|
| Urea | 8 M |
| NaCl | 300 mM |
| Imidazole | 30 mM |

**Wash Buffer 2**

| | |
|---|---|
| CAPS pH 10.5 | 200 mM |
| NaCl | 300 mM |
| Imidazole | 30 mM |

**Elution buffer**

| | |
|---|---|
| CAPS pH 10.5 | 200 mM |
| NaCl | 300 mM |
| Imidazole | 300 mM |

**Figure 2-21 Buffers used for Purification of ELF3.**

***In-vitro* EC reconstitution and EMSA assay.**

For LUX-ELF3 and LUX-ELF3-ELF4 experiments, all tested complexes were reconstituted by mixing the proteins of interest in 6 M urea in dialysis buttons, followed by a step-wise dialysis in order to incrementally reduce the urea concentration to 0 M, allowing protein complex refolding (6 M, 5 M, 4 M, 2 M, 1 M urea + 1 mM TCEP in 30 min. steps, and finally 50 mM sodium phosphate pH 7.6 + 100 mM NaCl + 1 mM TCEP for one hour). LUX and ELF4 concentrations were 200 nM and 0 or 1 μM, respectively, while the ELF3 concentration was varied from 220 nM to 2.2 μM. The DNA oligomer used was from the *PRR9* promoter (5´-ATGATGTCTTCTCAA**GATTCGA**TAAAAATGGTGTTG-3') and its concentration maintained constant at 30 nM in all EMSA experiments. Proteins and DNA were incubated at room temperature for 40 min. in binding buffer (10 mM Tris pH 7.0, 1 mM MgCl2, 1 mM TCEP, 6% glycerol, 28 ng/μL herring sperm DNA (Roche), 20 μg/mL BSA, 2.5% CHAPS and 1.25 mM spermidine) and protein-DNA complexes (LUX-ELF3 and LUX-ELF3-ELF4) run on a 2% agarose gel using TBE buffer 0.5x in non-denaturing conditions at 4ºC. Gels were scanned using a Chemidoc scanner (Biorad).

# CHAPTER 3. **BINDING AFFINITY AND CRYSTAL STRUCTURE OF LUX MYB DOMAIN.**

## 3.1 Introduction

The LUX DNA binding domain contains a plant-specific GARP family signature motif SH(A/L)QK(F/Y). The GARP acronym is built of gene names that belong to this family: Golden2 in maize (Hall et al.), ARR-B proteins from Arabidopsis (Imamura et al., 1999), and Chlamydomonas Psr1 (Wykoff et al., 1999). The GARP family of transcription factors (TFs) contains closely related genes involved in very diverse function such as plant development, hormone signaling, organ biogenesis, circadian clock oscillation, pathogen resistance and nutrient sensing systems in plants.

Earlier studies have revealed a great deal of insight into LUX DNA binding specificity through Protein Binding Microarrays. PBMs provide a rapid means of comprehensive characterization of in-vitro DNA binding specificities of TFs. Similar PBMs for LUX FL have revealed LUX binding preference over all possible 8-mers. From the PBMs it was found that LUX is a sequence specific DNA-binding protein that selectively binds to the sequence GAT(A/T)CG (Helfer et al., 2011).

Structure based sequence alignment suggests that the LUX DBD is structurally similar to the GARP transcription factor ARR10 from Arabidopsis with 60% sequence identity **(Figure 3-1**). NMR based structural studies done on ARR10 revealed the presence of highly conserved three helix bundle fold characteristic of Myb domains (Hosoda, 2002). In the same study it was observed that ARR10 had a relatively low DNA-binding affinity in the high-nanomolar to low-micromolar range (Hosoda, 2002). In order to understand the atomic level determinants of LUX DNA binding activity, structural and mutagenesis studies were performed on LUX FL and LUX DBD.

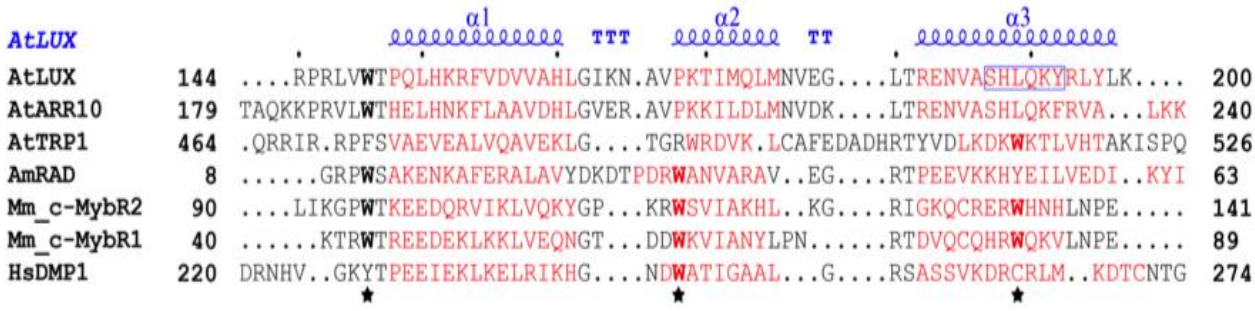



**Figure 3-1 Structure based Sequence alignments of MYB domains in comparison with LUX DBD.**

Structure-based sequence alignment of representative MYB domains; the three regularly spaced hydrophobic residues characteristic of MYB domains are indicated with a star. Depicted in red are the α-helices derived from each corresponding structure. The secondary structure annotation of LUX DBD is depicted in blue on top of the aligned sequences (α, alpha helices; TT, strict β-turn; TTT, strict α-turn).
The * represents the conserved tryptophan residues responsible for the hydrophobic cluster formation.

### 3.2 Determining the binding affinity of LUX FL and LUX-DBD for target DNA motifs .

For understanding the binding affinity of LUX FL and LUX-DBD, EMSAs were performed for different target motifs that were identified from the protein binding microarray experiments for LUX. The binding affinities were determined for the DNA motifs mentioned below

| | |
|---|---|
| Oligo PRR9 | 5´-ATGATGTCTTCTCAAGATTCGATAAAAATGGTGTTG-3’ |
| Oligo PBM1 | 5’- ATGATGTCTTCTCAAGATACGCTAAAAATGGTGTTG-3’ |
| Oligo PBM2 | 5’-ATGATGTCTTCTCAAGATCTTATAAAAATGGTGTTG-3’ |
| Oligo PBM3 | 5’-ATGATGTCTTCTCGGATCCGATAAAAATGGTGTTG-3’ |
| Oligo PBM4 | 5’-ATGATGTCTTCTCGAATATTCGATAAAAATGGTGTTG-3’ |

**Table 3-1 List of Oligonucleotides used for EMSA Assay**

Since it was known that LUX binds to the *PRR9* promoter, a fragment of the *PRR9* promoter containing a well-characterized LBS was assessed(Helfer et al., 2011). Both LUX FL and LUX DBD were tested with a constant DNA concentration of 10nM and a protein concentration varying from 0-1000nM. The EMSAs were further processed with ImageJ® and Graphpad® Prism software to find the binding affinity by determining the protein concentration at 50% free DNA. From these experiments it was found that LUX-DBD binds target *PRR9* LBS with better affinity compared to LUX FL. The $K_d$ of LUX-DBD for the *PRR9* LBS was found to be ~ 36nM while for LUX-FL the $k_d$ was found to be ~ 98nM **(Figure 3-2)**

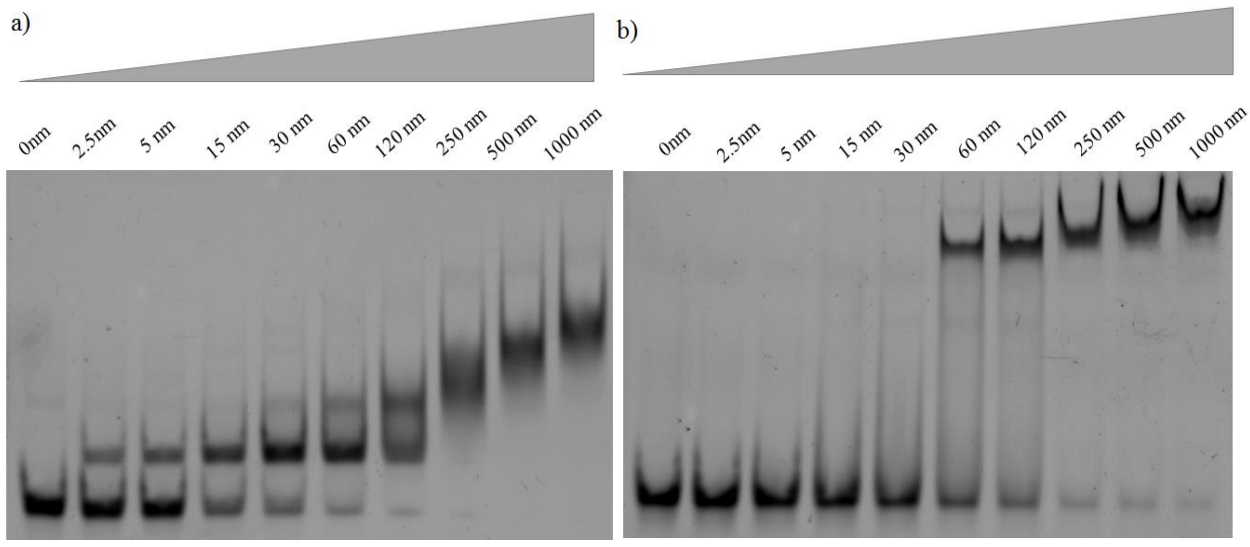



**Figure 3-2 Mobility shift assay for AGATTCGA (PRR9) DNA motif.**
Representative 8% polyacrylamide gel electrophoretic mobility shift assays (EMSAs) for $LUX^{MYB}$ (**a**) and LUX full-length (**b**). The DNA concentration was 10 nM with protein concentration increasing from 0 to 1000 nM as indicated.

In addition to PRR9 motif, binding affinities were determined for all high scoring sequences that were obtained from PBM experiments as mentioned in **Table 3-1**. The $K_d$'s of the LUX DBD and LUX FL proteins for PBM2, PBM3 and PBM4 is summarized in **Table 3-2.** From EMSAs with the LUX binding motifs from the PBM, it was observed that PBM1 sequence (GATACG) was the most preferred motif LUX DBD. The $K_d$ of LUX-DBD for PBM1 was found to be ~ 6 nM. On the contrary, LUX FL bound best to the AGATCTTA (PBM2) motif The $K_d$ of LUX-FL for PBM2 was found to be ~ 92 nM. However the difference for binding affinity of LUX FL for PBM1, PBM2 and PRR9 were very similar. LUX DBD alone exhibited higher affinity as compared to the full-length protein for all DNA oligomers tested with $K_d$'s ranging from 6.5 nM to 43 nM for stronger and weaker affinity oligomers, respectively. The full-length protein exhibited slightly lower affinity over the sequences tested, with $K_d$'s in the 90-180 nM range.

| LUX binding site | Binding affinity ($K_d$) | |
|---|---|---|
| Sequence of 8-mer motifs | LUX DBD (nM) | LUX FL (nM) |
| AGATTCGA (PRR9) | 36.7 ± 2.9 | 93.3 ± 5.8 |
| AGATACGC (PBM1) | 6.5 ± 1.4 | 98.3 ± 2.9 |
| AGATCTTA (PBM2) | 13.8 ± 1.8 | 92.5 ± 3.5 |
| GGATCCGA (PBM3) | 16.5 ± 2.1 | 117.5 ± 10.6 |
| ATATTCGA (PBM4) | 43.0 ± 4.2 | 177.5 ± 3.5 |

**Table 3-2 Binding affinity of LUX-DBD and LUX-FL for different 8-mer binding Motifs.**

The observed differences between the binding affinities of LUX FL and LUX-DBD could be because of the presence of the unstructured N- and C-termini which may partially occlude the MYB domain. This may explain the higher $K_d$'s of LUX FL when compared with LUX DBD for the same binding motifs.

### 3.3 Determining Crystal structure of LUX DBD

Crystallization trails with the full-length LUX protein were not successful. This is possibly because LUX full-length lacks predicted secondary structure in the N- and C-terminal regions of the protein. However, co-crystallization of seleno-methionine substituted LUX DBD with a 10-mer DNA dsDNA, 5'-TAGATACGCA (forward), 5'-ATGCGTATCT (reverse containing the core binding motif (underlined) had been previously done at our lab previously and yielded small crystals that diffracted to 2.8 A ° resolution (Silva et al., 2016). Crystal optimization was done in order to increase the resolution of the data and full data sets from a single crystal were obtained for LUX DBD with the above mentioned 10-mer dsDNA motif with a resolution of 2.14 Å and has been submitted to PDB with the PDBid 5LXU (Zubieta et al.). Apart from this we also crystallized LUX DBD with a second 10-mer dsDNA motif having the sequence 5'- TATATTCGAA which lacks the highly conserved guanine at position 3 and replaces the adenine with a thymine at position 6, a conservative change in the LBS consensus sequence AGAT(A/T/C)C. The crystal for this LUX DBD dsDNA motif was obtained at a resolution of 1.67 Å

and has also been submitted under PDBid 6QEC. The data collection and refinement statistics for both the structure are in **Table 3-3**.

| LUX DBD - 5'-TAGATACGCA-3' | |
|---|---|
| **Data collection on** | ESRF BEAMLINE ID29 |
| **Space group** | $P2_1$ |
| **Cell dimensions** | |
| ***a, b, c* (Å)** | 42.16, 32.83, 53.76 |
| **α, β, γ (°)** | 90, 98.6, 90 |
| **Resolution (Å)** | 42.-2.14 (2.22-2.14)* |
| $R_{sym}$ **or** $R_{merge}$ | 6.1 (60) |
| $I / \sigma I$ | 12.5 (2.0) |
| **Completeness (%)** | 91 (57) |
| **Redundancy** | 2.8 (1.7) |
| | |
| **Refinement** | |
| **Resolution (Å)** | 41.7-2.14 |
| **No. reflections** | 7572 |
| $R_{work}$ / $R_{free}$ | 19.5/23.2 |
| **No. atoms** | 975 |
| **Protein** | 491 |
| **DNA** | 409 |
| **Water** | 75 |
| ***B*-factors** | |
| **Protein** | 56 |
| **DNA** | 58 |
| **Water** | 56 |
| **R.m.s. deviations** | |
| **Bond lengths (Å)** | 0.01 |
| **Bond angles (°)** | 1.06 |

| LUX DBD- 5'- TATATTCGAA | |
|---|---|
| **Data collection on** | ESRF BEAMLINE ID29 |
| **Space group** | $P2_1$ |
| **Cell dimensions** | |
| ***a, b, c* (Å)** | 32.76, 51.80, 36.00 |
| **α, β, γ (°)** | 90, 110.6, 90 |
| **Resolution (Å)** | 31.-1.67 (1.9-1.67)* |
| $R_{sym}$ **or** $R_{merge}$ | 6.1 (60) |
| $I / \sigma I$ | 12.5 (2.0) |
| **Completeness (%)** | 91 (57) |
| **Redundancy** | 2.8 (1.7) |
| | |
| Refinement | |
| **Resolution (Å)** | 31.-1.67 |
| **No. reflections** | 9695 |
| $R_{work}$ / $R_{free}$ | 19.8/23.9 |
| **No. atoms** | 1105 |
| **Protein** | 491 |
| **DNA** | 409 |
| **Water** | 130 |
| ***B*-factors** | |
| **Protein** | 56 |
| **DNA** | 58 |
| **Water** | 56 |
| **R.m.s. deviations** | |
| **Bond lengths (Å)** | 0.004 |
| **Bond angles (°)** | 0.671 |

**Table 3-3 Crystallization data and statistics**

Data collection and refinement statistics of $LUX^{MYB}$-5'-TAGATACGCA (Left) and $LUX^{MYB}$-5'-TATATTCGAA (Right).

LUX DBD adopts a classic three-helix bundle conformation characteristic of PFAM00249 MYB and SANT domains. The hydrophobic core of MYB domains usually consists of three regularly spaced bulky hydrophobic residues, most often tryptophans forming a cluster in the hydrophobic core , with an 18 or 19 amino acid spacing(Kanei-Ishii et al., 1990). In LUX, however, the second and third tryptophan residues are replaced by a proline (Pro171) and leucine (Leu192). The hydrophobic core is further stabilized by edge to face interactions of Phe157 (helix 1) and Tyr195 (helix 3), pi stacking of Trp149 (helix 1) with His191 (helix 3) and edge to face interactions of His191 and Phe157**( Figure 3-3)**.

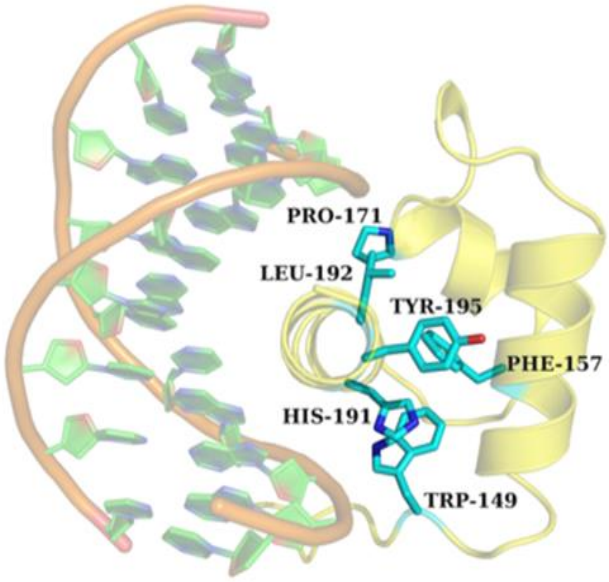



**Figure 3-3 Close up view of the hydrophobic core residues of LUX-DBD**

Close-up view of the hydrophobic core residues of $LUX^{MYB}$. The protein is shown as a yellow semi-transparent cartoon with side chains as sticks and colored by atom with carbons in cyan. Residues are labelled.

LUX DBD sequesters DNA primarily through helix 3 that lies in the major groove and contains a plant-specific GARP family signature motif, SH(A/L)QK(F/Y). When examined for electrostatic charge distribution, it was found that the electrostatic surface of LUX DBD demonstrates a highly electropositive face that acts as the main DNA binding surface. Its Helix 3 that harbors this electropositive face (Figure 3-4). This electropositive face is contributed by the presence of 5 basic amino acid (2 Lysine, 2 Arginine and 1 Histidine residues) in the third helix **(Figure 3-4)** making it the main DNA binding surface.

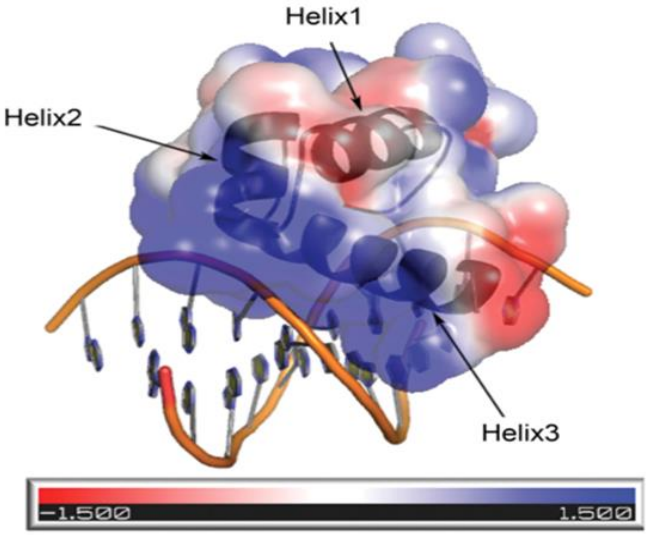



**Figure 3-4 Electrostatic surface representation of LUX DBD in complex with DNA.**

In this figure the electropositive surfaces are colored in blue while the electronegative surfaces are colored in red blue with helices lying in the interior indicated by arrows

For DNA base readout in the crystal structure with PBM1 motif (5'-TAGATACGCA), hydrogen-bonding interactions between Lys194 and guanine at position 3 (G3) and T*8, Gln193 and A4 and A*7 and Ser190 and A4 (where an asterix denotes the antisense strand) are important. Arginine196 contacts the sugar-phosphate backbone, further stabilizing the protein-DNA complex. In addition, Arg185 interacts with the DNA backbone through its secondary amine and participates in a water-mediated hydrogen-bonding network with G*5 **(Figure 3-5).** Helix 2 and 3 form a helix-turn-helix motif, constituting an electropositive groove for the negatively charged DNA and acting as the primary interface with the LUX binding site (LBS). Lys172 of helix 2 interacts with the sugar-phosphate backbone via van der Waal's interaction, helping to orient the DNA and to allow helix 3 to lie fully in the major groove. Residues from helix 3 account for the majority of the base readout and sugar-phosphate backbone interactions between the protein and DNA **(Figure 3-5)**. While no residues in helix 1 directly interact with the DNA, Arg146, part of the unstructured N-terminal extension, intercalates into the minor groove and interacts largely via van der Waal's forces and a water-mediated hydrogen bonding network with adenine and thymine/guanine of the bound DNA (T**AG**ATACGA and T**AT**ATTCGAA) **(Figure 3-6).** Interestingly, the Arg146 residue adopts different conformations in the two structures and different hydrogen bonding networks, suggesting plasticity in how this region of the protein binds DNA. Due to the flexible nature of the Arg146 tail, the protein can accommodate different DNA sequences in its binding motif. As Arg146 seems to act as a general "clamp" targeting the DNA minor groove, this residue was targeted for mutagenesis.

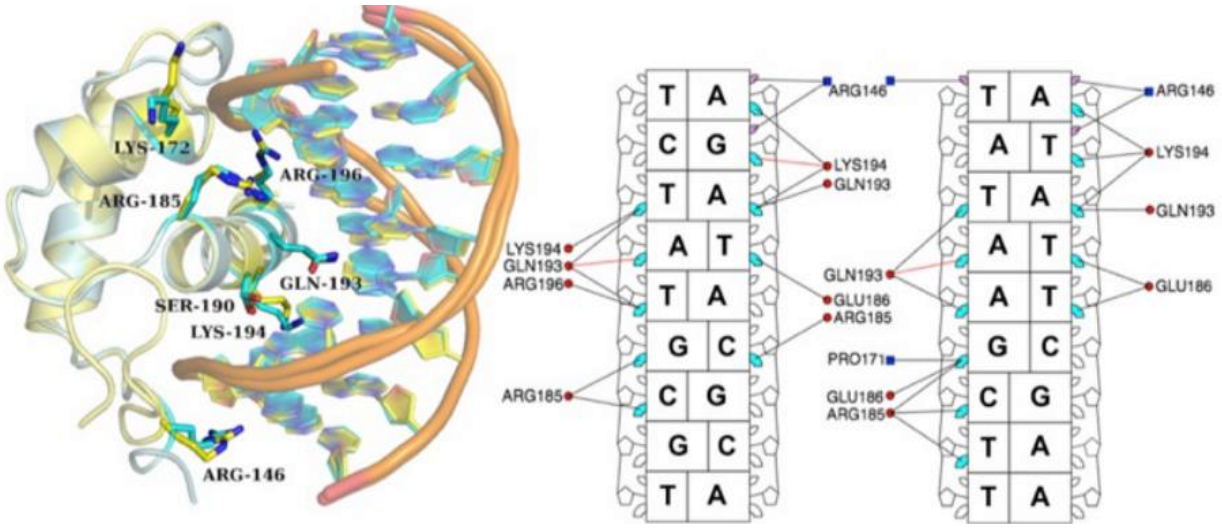



**Figure 3-5 Schematic representation of LUX DBD domain and DNA interactions.**

*Left*, overlay of $LUX^{MYB}$-DNA structures. Amino acid residues interacting with the DNA are shown as sticks and colored by atom with carbons in cyan (5'-TAGATACGCA) or yellow (5'-TATATTCGAA), DNA is shown as a cartoon. *Right*, simplified schematic from DNAproDB (Sagendorf *et al*).

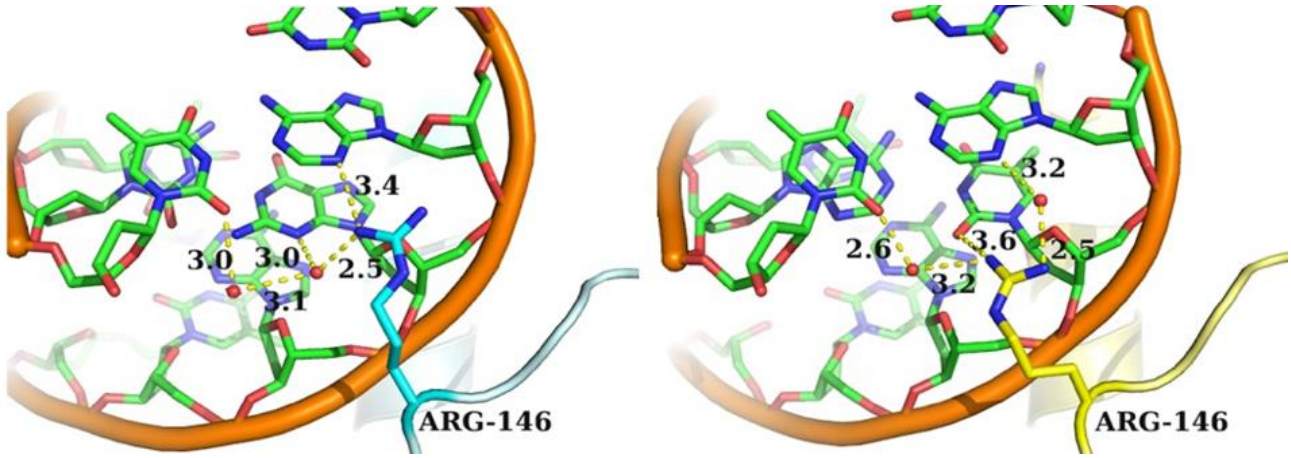



**Figure 3-6 Close up View of Arg146 interaction with different DNA sequences.**

Close up view of Arg146 interaction with different DNA sequences. Hydrogen bonding interactions are shown as dotted lines and distances labelled water molecules are shown in red spheres Arg146 adopts different conformations in the two structures.

### 3.4 Effect of Arg146Ala mutation on binding affinity of LUX FL and LUX DBD.

**EMSA assays to probe binding affinity of LUX$^{R146A}$ full length and LUX$^{R146A}$ DBD proteins.**

To understand the role of the arginine at 146 position of LUX, site directed mutagenesis was done to change the arginine at position 146 to an alanine in both the LUX FL and LUX DBD. These proteins were expressed and purified using the same protocols as described before for the wildtype version of the protein.

Binding affinities were determined for the mutated (R146A) version of LUX DBD and LUX FL. Like the previous EMSA experiments, PRR9 (AGATTCGA) motif was the first motif to be tested. For the PRR9 motif it was found that the dissociation constant of LUX$^{R146A}$ DBD was approximately 50nM and for the LUX$^{R146A}$ FL it was found to be ~ 105nM (Figure 3-7). Hence when compared to the wildtype proteins, we see that the dissociation constant of the LUX$^{R146A}$ DBD for PRR9 motif increased by less than one fold, while for LUX$^{R146A}$ FL there wasn't a very significant increase in the dissociation constant ($K_d$). $K_d$ of wildtype LUX FL for the PRR9 motif was found to be 93nM while for the LUX$^{R146A}$ FL was found to be ~105nM.

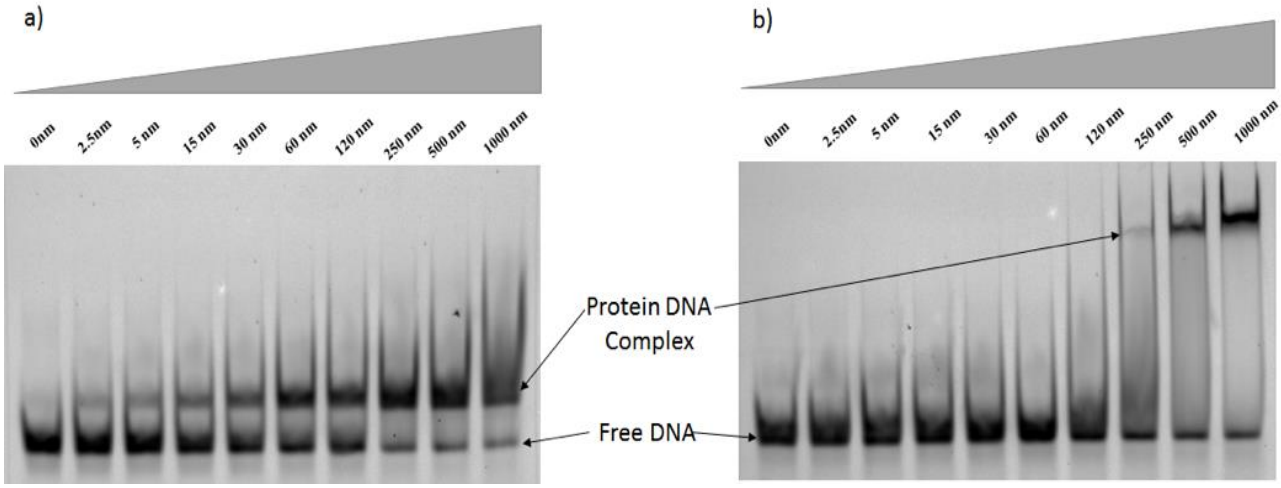



**Figure 3-7 Mobility shift assay for AGATTCGA (PRR9) DNA motif with R146A mutated proteins.**

Representative 8% polyacrylamide gel electrophoretic mobility shift assays (EMSAs) for LUX$^{R146A}$DBD (a) and LUX$^{R146A}$ full-length (b). The DNA concentration was 10 nM with protein concentration increasing from 0 to 1000 nM as indicated

After the PRR9 motif, the $K_d$'s of the mutant proteins were determined for PBM1 motif which was used for crystal structure determination. For the PBM1 motif, $K_d$ of LUX$^{R146A}$ DBD was approximately 50nM and for the LUX$^{R146A}$ FL it was found to be ~ 336nM (Error! Reference source not found.). When c ompared to the wildtype proteins, we see that the $K_d$ of the LUX$^{R146A}$ DBD for PBM1 motif increased around 6.5 folds compared to the wildtype LUX DBD, while for LUX$^{R146A}$ FL there was a 2.5 folds increase in the $K_d$. $K_d$ of wildtype LUX DBD for the PBM1 motif was ~ 6.5nM while same for the LUX

FL was found to be ~ 98.3nM. Consequently the $K_d$'s of the mutant proteins were determined for PBM2, PBM3 and PBM4 motif which is summarized in **Table 3-4**. For the PBM2 motif, $K_d$ of $LUX^{R146A}$ DBD was approximately 63 nM and for the $LUX^{R146A}$ FL it was found to be ~ 204nM. When compared to the wildtype proteins, we see that the $K_d$ of the $LUX^{R146A}$ DBD for PBM2 motif increased around 3.5 folds compared to the wildtype LUX DBD, while for $LUX^{R146A}$ FL there was a 1.2 folds increase in the $K_d$. $K_d$ of wildtype LUX DBD for the PBM2 motif was ~ 13.8nM while same for the LUX FL was found to be ~ 92.5 nM from previous experiments.For the PBM3 motif, $K_d$ of $LUX^{R146A}$ DBD was ~ 120 nM and for the $LUX^{R146A}$ FL it was found to be ~ 164nM (**Error! Reference source not found.**). When c ompared to the wildtype proteins, we see that the $K_d$ of the $LUX^{R146A}$ DBD for PBM3 motif increased around 6.2 folds compared to the wildtype LUX DBD, while for $LUX^{R146A}$ FL there was a 0.3 folds increase in the $K_d$. $K_d$ of wildtype LUX DBD for the PBM3 motif was ~ 16.5nM while same for the LUX FL was found to be ~ 117.5nM from previous experiments. For the PBM4 motif, $K_d$ of $LUX^{R146A}$ DBD was found to be ~ 137 nM. However for $LUX^{R146A}$ FL from the EMSAs it was found that the dissociation constant would be in micromolar range. When compared to the wildtype proteins, we see that the $K_d$ of the $LUX^{R146A}$ DBD for PBM4 motif increased around 2 folds compared to the wildtype LUX DBD, $K_d$ of wildtype LUX DBD for the PBM3 motif was ~ 43nM while same for the LUX FL was found to be ~ 177.5 nM from previous experiments.

| **LUX binding site 8-mer motifs** | **LUX DBD (nM)** | **LUX FL (nM)** | **$LUX^{R146A}$DBD** | **$LUX^{R146A}$FL** |
|---|---|---|---|---|
| **AGATTCGA (PRR9)** | 36.7 ± 2.9 | 93.3 ± 5.8 | 50±2.4 | 105±11.8 |
| **AGATACGC (crystal)** | 6.5 ± 1.4 | 98.3 ± 2.9 | 50±1.8 | 336±4.8 |
| **AGATCTTA (PBM2)** | 13.8 ± 1.8 | 92.5 ± 3.5 | 63±3.1 | 204±6.7 |
| **GGATCCGA (PBM3)** | 16.5 ± 2.1 | 117.5 ± 10.6 | 120±1.0 | 164±4.2 |
| **ATATTCGA (structure)** | 43.0 ± 4.2 | 177.5 ± 3.5 | 137±13.0 | nd* |

**Table 3-4 Comparison of Dissociation constants.**
Dissociation constants for LUX FL and LUX DBD for various binding motifs compared dissociation constants of $LUX^{R146A}$ DBD and $LUX^{R146A}$ FL for same binding motifs. *binding was too weak to measure the dissociation constant.

## 3.5 Discussion

In this chapter, the DNA binding affinity of LUX full length and the MYB like DNA binding domain of LUX for top scoring motifs from the LUX PBM was established using the electrophoretic mobility shift assay. It was found that LUX DBD was able to bind target motifs in a range of 6~50 nM range for different motifs that were tested. For LUX full length the binding affinities were much higher in the range of 100~300nM.

This behavior of the full length LUX in comparison to the DNA binding domain could be attributed to the unstructured and disordered regions of n-terminal and c-terminal part of the protein. When compared with its nearest relative, ARR10 MYB domain, the LUX MYB domain shows better binding affinity. While the dissociation constants for the ARR10 MYB domain were determined in the range of higher nano-molar to micro-molar range (Hosoda, 2002), the dissociation constants for the LUX MYB domain were in lower nano-molar range. This illustrates the difference in the two MYB domains. The high binding affinity of the LUX MYB domain can be attributed to a combination of allosteric effects in conjunction with the plasticity of N-terminal Arginine clamp that allows the protein to interact efficiently with its target DNA sequence.

From the structural studies, it was found that it was not possible to crystallise full length LUX, in consistence with the binding affinity studies, it's suggested that this might be because of the n-terminal and c-terminal parts of LUX which are predicted to be flexible and possess intrinsic disorder. Crystal structure was obtained for the MYB like DNA Binding Domain of LUX. From the crystal structure of LUX MYB domain, it was found that helix 3 was important for base read out and interaction with target DNA motif. In accordance with other MYB type domains, it was found that these interactions were stabilised with the formation of a hydrophobic core. When compared with the other MYB Domains, it was found that the Tryptophans in helix 2 and helix 3 in the LUX MYB Domain were replaced with a Proline and Leucine residues, however even with these replacements, the LUX MYB was still able to retain the hydrophobic core required for stable interaction with the DNA.

From the crystal structure of LUX MYB it was found that the Arginine present in the N-terminal played a vital role in its affinity for interaction with target DNA. This Arginine residue acted as a clamp to increase the binding affinity interaction. It can tolerate change in sequences. This suggests that its not important for specificity but rather increases the affinity of the interaction. This was further proven by site directed mutagenesis studies. It was seen that upon mutation of the Arginine 146 to Alanine, the binding affinity of the LUX MYB changed up to 6 folds in some cases. When dissociation constants for full length and MYB domain of LUX were compared with the Arginine to Alanine mutated version of the respective proteins it was found that the dissociation constants of the mutated protein version were higher than the wildtype version for all target motifs.

## 3.6 Methodology

### Electrophoretic Mobility Shift Assays (EMSAs)

The DNA binding affinities of native LUX FL and LUX DBD proteins were tested via EMSA assays. A 36bp DNA oligomer (5´-ATGATGTCTTCTCAAGATTCGATAAAAATGGTGTTG-3') from the PRR9 promoter containing a LUX DNA-binding site (in bold and underlined) was tested, as well as high scoring sequences from PBM (Helfer et al., 2011) experiments with sequences as follows:

| | |
|---|---|
| Oligo PRR9 | 5´-ATGATGTCTTCTCAAGATTCGATAAAAATGGTGTTG-3' |
| Oligo PBM1 | 5'- ATGATGTCTTCTCAAGATACGCTAAAAATGGTGTTG-3' |
| Oligo PBM2 | 5'-ATGATGTCTTCTCAAGATCTTATAAAAATGGTGTTG-3' |
| Oligo PBM3 | 5'-ATGATGTCTTCTCGGATCCGATAAAAATGGTGTTG-3' |
| Oligo PBM4 | 5'-ATGATGTCTTCTCGAATATTCGATAAAAATGGTGTTG-3' |

**Table 3-5 List of oligo-nucleotides used for EMSA.**

a) Binding Buffer used in EMSA

| | |
|---|---|
| Tris pH 7.0 | 10mM |
| NaCl | 50mM |
| $MgCl_2$ | 1mM |
| TCEP | 1mM |
| Herring sperm DNA | 28ng/mL |
| BSA | 20mg/mL |
| CHAPS | 2.5% |
| Spermidine | 1.25mM |

**Table 3-6 List of Buffers used for EMSA.**

All dsDNA oligomers tested were 35-37bp in length and Cy5 labelled (Eurofins Genomics). Protein concentration was varied from 0nM to 1000nM for LUX DBD and LUX full length (0nM, 2.5nM, 5.0nM, 15nM, 30nM, 60nM, 120nM, 250nM, 500nM and 1000nM) using a constant DNA concentration of 10nM in all reactions. The same was also done for the $LUX^{R146A}$ DBD and $LUX^{R146A}$ Full length. Protein and DNA were incubated at room temperature for 40min in binding buffer (mentioned in the table above) and protein-DNA complexes (LUX FL or LUX DBD) run on a 8% polyacrylamide gel using TBE buffer 0.5x in non-denaturing conditions at 4ºC.Gels and blots were scanned using a Chemidoc scanner (Biorad). A summary of the methodology is given in the **Figure 3-8**.

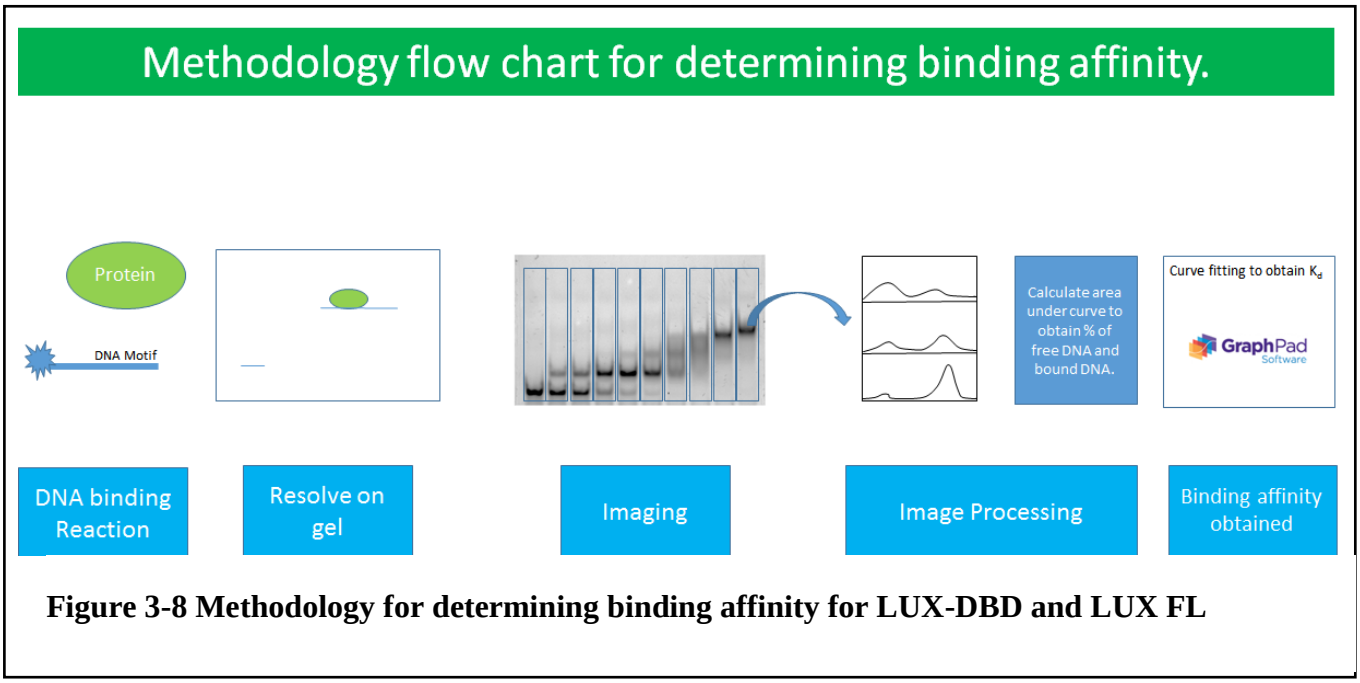



**Figure 3-8 Methodology for determining binding affinity for LUX-DBD and LUX FL**

## LUX DBD Crystallization

### Protein Purification

Protein expression and purification was performed using methods previously described in chapter 2. The methods for expression purification and crystallization of seleno-methionine version of LUX DBD was done using methods previously described (Silva et al., 2016).

### Protein Crystallisation and data collection

Crystallisation experiments were carried out by the vapour diffusion method at 293K, using sitting-drops with a 1:1 ratio of protein-DNA complex:precipitant with a protein concentration of ~6 mg/mL. Suitable well-diffracting crystals were grown after 2-4 days in 0.1 M BisTris Propane, pH 6.5, 20% PEG 3350 and 0.2 M sodium malonate. Crystals grew as needles to dimensions of 200x50x50 µm and were harvested and flash frozen in liquid nitrogen. Crystallisation was performed by the EMBL High Throughput Crystallisation Facility (HTX).

For the $LUX^{MYB}$ structure with the DNA sequence (5'-TATATTCGAA-3', reverse oligo 5'-ATTCGAATAT-3') data were collected on beamline 23-1 at 0.976 Å wavelength. Indexing was performed using EDNA (Incardona et al., 2009)and the default optimized oscillation range and collection parameters used for data collection. The data set was processed and scaled using the programs XDS and XSCALE(50). Due to anisotropy of the data, the Staraniso server was used for further data reduction and subsequent refinements. All refinements were performed with Phenix. Final Ramachandran statistics were 100% preferred region for all residues. The structure is deposited to PDB under PDB id 6QEC.

**Mutagenesis of LUX FL and LUX DBD.**

Previous constructs of LUX FL and LUX DBD were subjected to Site Directed Mutagenesis (SDM) using the QuickChange® site directed mutagenesis kit from Stratagene. The primers for SDM reactions were designed using PrimerX software from www.bioinformatics.org website. We used standard parameters provided in the software for substitution of Arg146 to Ala. The software returned the following primers for SDM reactions:

LUX R146A Forward: 5' CACTTAAACGACCGGCTTTAGTGTGGACACC 3'
LUX R146A Reverse: 5' GGTGTCCACACTAAAGCCGGTCGTTTAAGTG 3'

Standard protocols from the kit were followed to obtain the mutated constructs which were later verified by sequencing. Sequence verified constructs were used for expression of mutated $LUX^{R146A}$ FL and $LUX^{R146A}$ DBD in *E.coli*. The purification protocol used was same as the LUX FL and LUX DBD purification protocols mentioned in CHAPTER 2 section 2.7.

# CHAPTER 4. **ALTERING EC FUNCTIONS *IN PLANTA*.**

## 4.1 Introduction

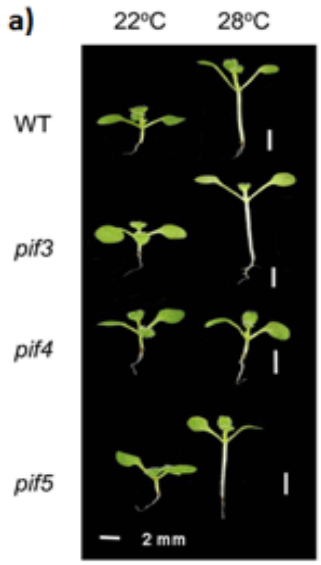
The crystal structure of the DNA binding domain of LUX provided insight into important amino acids involved in interaction with target motifs. It was found that the Arg146 of LUX acts as a clamp and facilitates LUX binding to the target motifs. EMSA assays using $LUX^{R146A}$ showed reduced binding affinity to target motifs when compared with the wt LUX. However to understand its relevance *in-planta*, it was required that the same mutation be made in plants and the effect of this mutation on thermoresponsive growth be studied. This chapter deals with the above said objective.

EC regulates thermoresponsive growth in evening through repression by direct binding to *PIF4* and *PIF5* promoters. As cellular levels of EC decreases, repression on PIF4 and PIF5 expression is removed and expression of *PIF4* and *PIF5* peaks around Zt-8 when evening complex is at its minimum levels (Nusinow et al., 2011).

PIF4 is a hub for temperature related growth by effecting multiple genes encoding for hormone signaling pathways that are integrated for growth and development. It controls the auxin, gibberellic acid and brassinosteroid signaling pathways that are involved in phototropism, cell elongation and leaf senescence. The relevance of PIF4 in temperature sensing was observed in *pif4* mutants that are insensitive to higher ambient temperature(Koini et al., 2009). It was shown that along with wt *Col-0* plants, *pif3* and *pif5* mutants responded to hypocotyl elongation at higher ambient temperature however *pif4* mutants on the other hand didn't **(Figure 4-1).**

Hypocotyl elongation phenotype was also observed for individual EC component mutant plants. The EC component mutants, *elf4*(Kim et al., 2012), *elf3* and *lux*(Box et al., 2015), have elongated hypocotyl at normal ambient temperature (22ºC). Effects of *elf3* mutation on PIF4 assisted thermoresponsive growth has been very extensively studied. These studies have shown that ELF3 binding to *PIF4* promoter is decreased at higher ambient temperature compared to normal temperature 22ºC (Box et al., 2015). Consistently, *elf3* mutant displays elongated hypocotyl at 22 ºC. This supports the fact that ELF3 inactivation at high temperature is a major mechanism to induce thermoresponsive growth. Moreover, the hypocotyl growth of *elf3* mutants is significantly enhanced at high temperature under long day conditions (Raschke et al., 2015). ELF3 is also known to interact directly with PIF4. Yeast two hybrid experiments have shown that the bHLH domain of PIF4 interacts with ELF3 suggesting that ELF3 sequester PIF4(Nieto et al., 2015). It has also been proposed at higher temperature, PIF4 sequestration by ELF3 is removed leading to availability of active PIF4. This suggests high temperature activates PIF4 both post-transcriptionally and transcriptionally. However it hasn't been shown that ELF3 has any DNA binding activity on its own. Hence if ELF3 acts through the evening complex for controlling *PIF4* expression, it must be through EC and hence LUX is an important target because of its DNA binding ability.

To test the role of R146A mutation in LUX on ambient temperature responses, *in-planta* experiments were designed to complement *lux4* mutant plants with *LUX* promoter driven wildtype *LUX* gene and *LUXR146A* mutant gene. Homozygous single insertion lines were obtained to perform phenotypic and transcriptomic analysis of these mutant complemented lines to understand the role of LUX in regulating PIF4 mediated thermoresponsive growth.



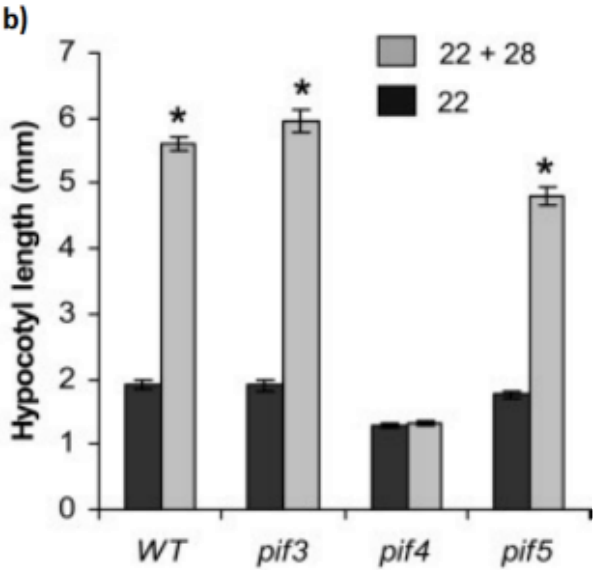




**Figure 4-1 High Temperature-Mediated Hypocotyl Elongation in *pif* Mutants.**

**Plant phenotypes (a) and hypocotyl lengths (b) of WT and *pif* mutant seedlings grown in continuous irradiation at different temperatures. Plants were grown at 22°C for 4 days before transfer to 28°C for 3 days. Control plants were maintained at 22°C. Error bars represent SE. Asterisks represent a significant difference from the 22°C sample with Student's t test ($p \leq 0.05$). (Adapted from Koini et al. 2009)**

## Results

### 4.2 Growth and development studies of LUX $^{R146A}$ plants.

As it was observed that Arginine at the 146 position of LUX is important for DNA binding through *in-vitro* assay. To investigate whether this mutation has a role to play *in-planta*, it was decided to generate *lux1-4 loss of function* mutant plants (hereafter referred to as *lux4*) complemented with the wildtype *LUX* (hereafter referred to as *lux4*-LUXwt) and $LUX^{R146A}$ (hereafter referred to as *lux4*-$LUX^{R146A}$) genes that were cloned under the wildtype *LUX* promoter. These complemented lines were used for studying flowering time, hypocotyl length measurements and gene expression studies.

### 4.2.1 Hypocotyl elongation phenotype of LUX$^{R146A}$ plants.

Hypocotyl elongation phenotypes were studied with the complemented (*LUX and LUX*$^{R146A}$) single insertion homozygous lines in the *lux4* mutant background to understand the importance of *LUX* $^{R146A}$ mutation on growth. Hypocotyl lengths were determined for *lux4*, *lux4*-LUXwt, *lux4-LUX*$^{R146A}$ and Col-0 lines under short days at 22°C. It was found that in short day photoperiod *lux4* seedlings had longer hypocotyls compared to all the other three lines used while hypocotyl lengths of *lux4*-LUXwt seedlings and Col-0 seedlings didn't have significant difference in hypocotyl lengths. However for the *lux4-LUX*$^{R146A}$ line it was found that the hypocotyl phenotype was in between the wild type and the *lux4* mutants **(Figure 4-2).**

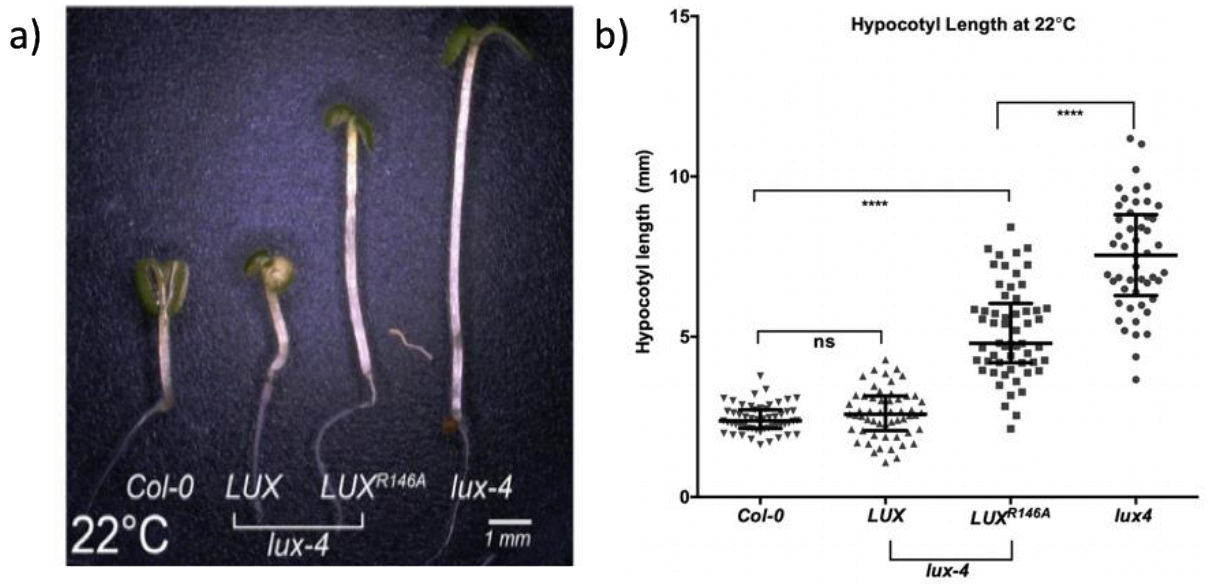



**Figure 4-2 Hypocotyl growth comparison in complemented *lux* plants.**

**a)** Representative images showing that LUX$^{R146A}$ doesn't fully suppress the long-hypocotyl phenotype of the *lux-4* while wildtype LUX protein complements the *lux* loss of function mutation. Shown are 7-days-old seedlings of lux4, *lux-4* LUX$^{R146A}$, *lux-4* LUX and Col-0 grown at 22°C short day conditions for 7 days. **b)** Hypocotyl length measurement showing that LUX$^{R146A}$ doesn't fully suppress the long-hypocotyl phenotype of the *lux-4* mutants while LUX wildtype protein complements the lux loss of function. Hypocotyl length of 7-days-old seedlings of the indicated genotypes grown at 22°C SD(8H-L/16H-D)was measured. Data shown are average±SD. One way annova test between mutant/transgenic and wild-type seedlings was performed (ns, not significant;****, P<0.001). Shown are representative data from three biological replicate using 3 independent mutant lines for *lux-4* LUX$^{R146A}$ and *lux-4* LUX, measured for 15~20 seedlings for each line.

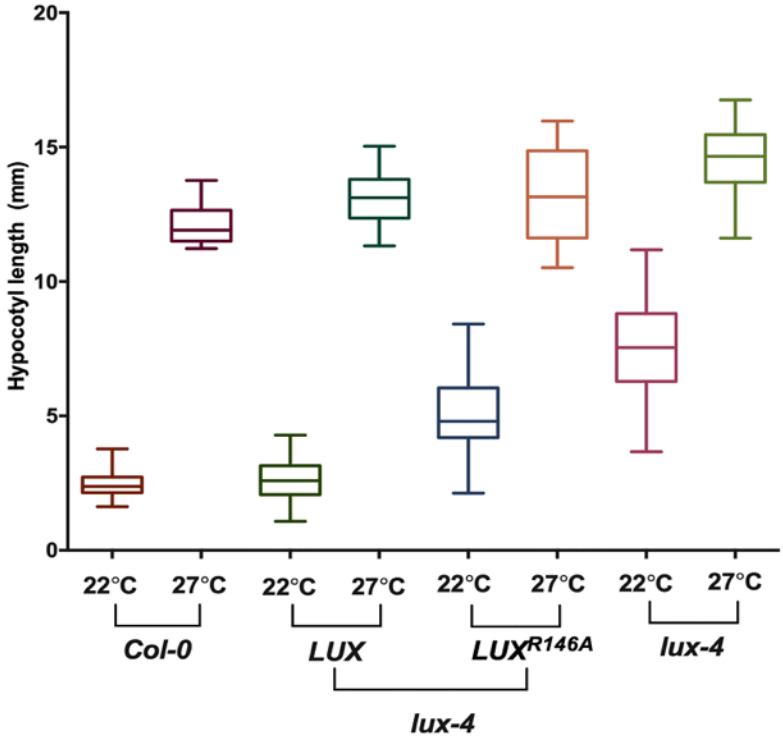



**Figure 4-3 Hypocotyl Length at 22°C compared with 27°C**

Hypocotyl lengths of wildtype plants compared with mutant plants when grown at 22°C vs 27°C in SD light conditions. The box plot presents hypocotyl lengths of ~15 plants for each line.

In contrast, hypocotyl lengths for *lux4*, *lux4*-LUX, *lux4-LUX*$^{R146A}$ and Col-0 lines under short days at 27°C were also measured. It was observed that compared to the 22°C, the *lux4*-LUXwt, *lux4-LUX*$^{R146A}$ and Col-0 lines, had longer hypocotyl lengths while No significant difference in hypocotyl lengths were observed between the *lux4*-LUX, *lux4-LUX*$^{R146A}$ and Col-0 lines at 27°C. Taken together, these data show that temperature responsive growth was still intact in the *lux4-LUX*$^{R146A}$ plants however it was more attenuated compared to the Col-0 and *lux4*-LUX plants. The increase in hypocotyl length of Col-0 and *lux4*-LUX lines were higher when temperature was changed from 22°C to 27°C when compared to the *lux-4* **(Figure 4-3).**

### 4.2.2 Flowering Phenotype of LUX$^{R146A}$ complemented plants

Flowering time was determined for the *lux4*-LUXwt, *lux4-LUX*$^{R146A}$ and were compared with the *lux4* and Col-0 plants. For the flowering time experiments, it was found that the phenotype of *lux4-LUX*$^{R146A}$ plants were intermediate between the *lux4* and Col-0 plants. The *lux4*-LUXwt plants had phenotypes similar to the Col-0 plants with no significant difference in the flowering time. From the flowering time experiments it was found that *lux4* flowered at 10-12 rosette leaves while WT flowered at 17-19 leaves.

The flowering phenotype of *lux4*-LUXwt was similar to that of WT and flowered between 17-18 rosette leave while the *lux4-LUX*$^{R146A}$ flowered between 13-15 rosette leaves **(Figure 4-4).**

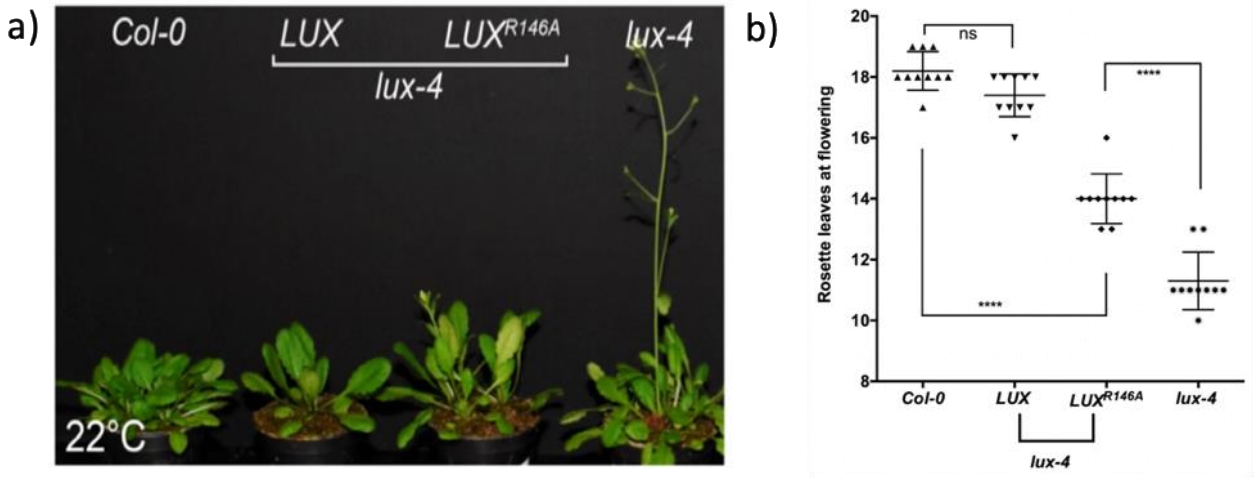



**Figure 4-4 *lux4*-LUX$^{R146A}$ plants flower later than *lux4* LUXwt plants.**

a) Representative images showing that LUX$^{R146A}$ doesn't fully suppress the early flowering phenotype of the *lux4* while wildtype LUX protein complements the *lux* loss of function mutation. Shown are plantlets of lux4, *lux4*-LUX$^{R146A}$, *lux4*LUXwt and Col-0 grown at 22°C in soil.
b) Flowering time measurement showing that LUXR146A doesn't fully suppress the early flowering phenotype of the *lux4*. LUX wildtype protein complements the *lux* loss of function. Number of rosette leaves at bolting are indicated for different lines grown at 22°C was measured. Data shown are average±SD. One way annova test between mutant/transgenic and wild-type seedlings was performed (ns, not significant;****, P<0.001). Shown are representative data from three biological replicate using 3 independent mutant lines for *lux4*LUX$^{R146A}$ and *lux4*LUXwt, measured for 10 plants for each line.

### 4.2.3 Petiole phenotype of LUX$^{R146A}$ complemented plants

Petiole phenotype was observed at 22ºC to see if the complemented lines had difference. It was seen that while both the *lux4* and the *lux4-LUX*$^{R146A}$ had accelerated growth and elongated petioles, the *lux4*-LUXwt plants had growth and petiole phenotypes similar to the wildtype plants. In comparison to Col-0 the *lux4*-LUXwt plants had longer petiole however this was still less than what was observed for the *lux4* and the *lux4-LUX*$^{R146A}$ lines **(Figure 4-5).**

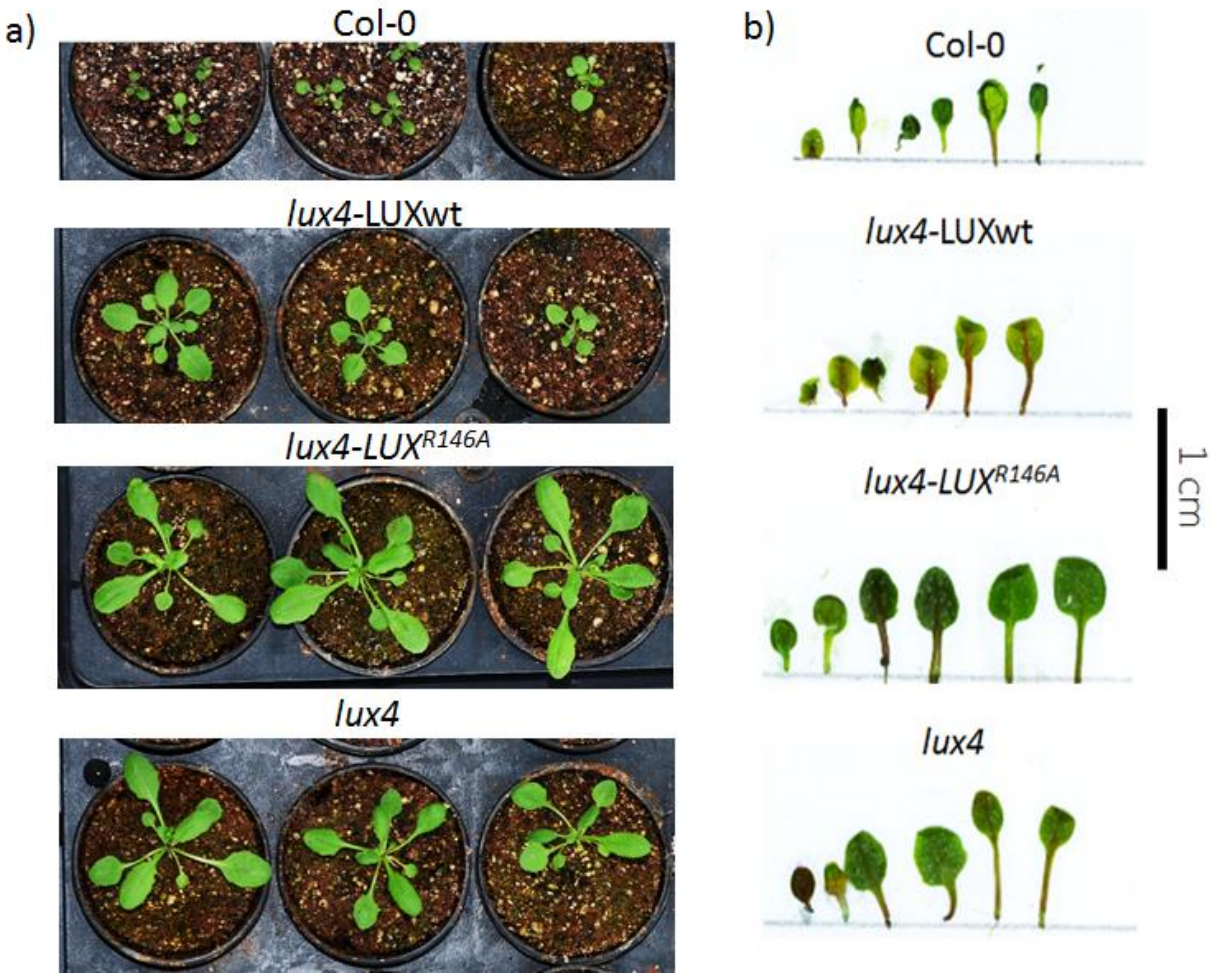



**Figure 4-5 Petiole Elongation Phenotype of *lux4* compared with other complemented lines.**

**a)** Col-0, *lux4*-LUXwt, *lux4-LUX*$^{R146A}$ and *lux4* plants grown on soil for 20 days and petiole phenotype was observed. **b)** Col-0, *lux4*-LUXwt, *lux4-LUX*$^{R146A}$ and *lux4* plants grown on MS agar media plates for 2 weeks. All leaves from the plants were imaged.

### 4.3 PIF4 expression studies in LUX$^{R146A}$ and LUXwt complemented lines.

PIF4 is a master regulator of temperature related growth by effecting multiple genes encoding for hormone signaling pathways that are integrated for growth and development. It controls the auxin, gibberellic acid and brassinosteroid signaling pathways that are involved in phototropism, cell elongation and leaf senescence. *PIF4* expression is controlled by the EC where LUX acts as the DNA binding protein. In the *lux4* mutant, the Pro171Leu substitution is responsible for the hypocotyl elongation and early flowering phenotype. In previous studies it has been shown that there are changes in *PIF4* expression that leads to the observed phenotypes of early flowering and hypocotyl elongation.(Nusinow et al., 2011).

*PIF4* expression was analyzed in the complemented lines to observe changes in the PIF4 expression profile and compare these changes with respect to the Col-0 and *lux4* plants. RNA samples were collected from Col-0, *lux4*-LUXwt, *lux4-LUX*$^{R146A}$ and *lux4* seedings grown on 0.5 MS media under short day (8L:16D) photoperiod for 7 days. A time course sample harvesting procedure was followed with samples collected every 4 hours beginning with switching on of the lights in the growth cabinets.

Samples were collected from 3 independent lines of LUX $^{R146A}$ and LUXwt and were compared with the Col-0 and *lux4* lines. For the qRT-PCR experiments PP2A gene was taken as an internal control against which *PIF4* expression was normalized.

From the expression profile studies, it was found that at the peak of expression which was observed at ZT-4 for all the lines, the *PIF4* expression for the *lux4*-LUX$^{R146A}$ lines were intermediate between Col-0 and *lux4*. At ZT-4, *lux4* mutant lines had the highest PIF4 expression while Col-0 plants had the lowest and the *PIF4* expression in the *lux4*-LUX$^{R146A}$ plants was found to be in between. Furthermore, at ZT-12 it was observed that there was a sharp decrease in PIF4 expression for Col-0 while this decrease was less for the R146A and *lux4* line. The decrease in PIF4 expression continued till ZT-16 for the *lux4* and the *lux4*-LUX$^{R146A}$ mutant. The lowest expression of PIF4 was observed at ZT-12 for the wildtype lines while the same for the *lux4* and the *lux4*-LUX$^{R146A}$ was observed at ZT-16 **(Figure 4-6).**

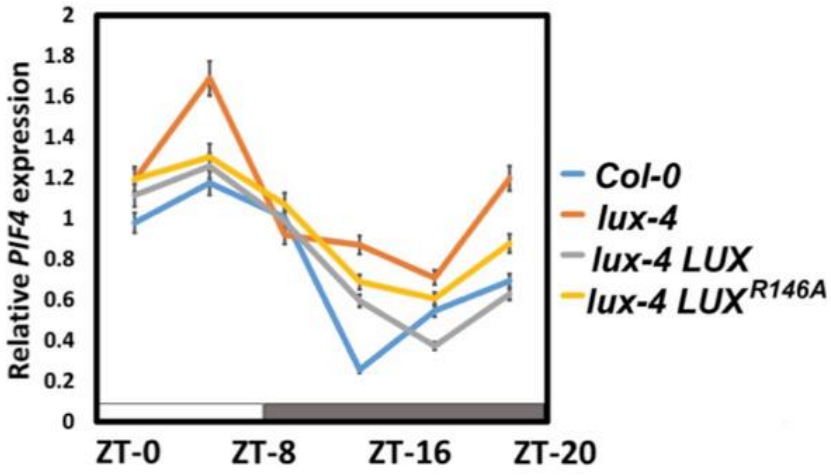



**Figure 4-6 PIF4 expression profile over a course of 24 hours.**

*PIF4* expression over a 24-hour period for seedlings grown at 22°C under short day conditions for the different genotypes. A single representative line was used for qPCR measurements. *PIF4* expression is higher in the *lux-4 pLUX::LUX*$^{R146A}$ and *lux-4* lines as compared to wild type. Day and night are indicated as a bar below the graph. Error bars represent the mean with standard deviation.

## 4.4 Discussion

In this chapter the effect of Arg146Ala mutation *in planta* was investigated. The results presented in this chapter complement the *in vitro* and structural studies presented in the previous chapter. It was observed that when $LUX^{R146A}$ was introduced in the *lux-4* mutants background, it poorly complemented the *lux* mutation. As predicted, we observed dampened EC activity arising due to weaker binding because of the R146A mutation in the LUX DNA binding domain. At 22°C, an intermediate hypocotyl elongation and early flowering phenotype between wild type and *lux-4* mutant was observed for the *lux-4* mutant complemented with $LUX^{R146A}$ . This suggests that the Arg at the 146 position of LUX is an important site that affects binding affinity of LUX *in planta*. Furthermore, to corroborate the effect of poor EC binding on *PIF4* expression, results pertaining to *PIF4* expression studies were presented in this chapter. The peak of *PIF4* expression was similar for wild type and *lux-4 pLUX::*$LUX^{R146A}$, however the characteristic strong decrease in PIF4 expression between ZT8 and ZT12, which coincides with maximum EC expression, was less apparent, likely due to the decreased affinity of the R146A mutation for its LBS.

## 4.5 Methodology

### Plant materials and generation of transgenic Arabidopsis plants.

The *lux1-4* mutant alleles (background accession Col-0) were provided Dr. Phil Wiggie (The Sainsbury Lab at Cambridge University). For obtaining mutant plants, *lux1-4* mutant was transformed with plasmid described below using electro-competent *Agrobacterium* that was prepared and transformed as described (Shen and Forde, 1989). Plants were transformed with the transgenic *Agrobacterium* harboring desired plasmid using the floral-dip method, as described (Clough and Bent, 1998).

### Plasmid construction and generation of transgenic plants

For the pLUX:: Flag::LUX$^{R146A}$ , pLUX::Flag::LUX constructs, 1.5kb upstream fragments of LUX was PCR-amplified from genomic DNA. Full length CDS of LUX and LUXR146A constructs were PCR-amplified from Expression vector containing the respective CDS with Flag tag using primers listed in the table below. NEBuilder® HiFi DNA Assembly Kit (E2621S, NEB) was used for assembling the promoter fragment with cDNA fragment and vector backbone pFP101 containing *At*Se2 promoter driven GFP for selection of transformants. (For list of primers see table 4). Transgenic plants were generated by *Agrobacterium*-mediated gene transfer with floral dip method. *Lux-4* plants were dipped with *Agrobacterium* containing pLUX::Flag::LUX$^{R146A}$ , pLUX::Flag::LUX constructs to obtain *lux-4 pLUX::LUX$^{R146A}$* and *lux-4 pLUX::LUX* plants. Mutant plants were screened for single copy transgene insertion using fluorescence screening of seeds. Phenotypes of transgenic plants were verified in at least three independent transgenic lines for each complemented line.

| Primer name | Primer DNA Sequence |
|---|---|
| **LUX Promoter FR** | GCTAAGCTTGCATGCCTGCACGTTTCGTCAGTTTGTGAAG |
| **LUX Promoter RV** | ATCGTCTTTGTAGTCCATTTCAAACTCTCTAATTTCTCG |
| **LUX cDNA FR** | AGAGTTTGAAATGGACTACAAAGACGATGACGACAAGATGGGAGAGGAAGTACAAATG |
| **LUX cDNA RV** | TCTGCAGGTCGACGGATCCTTTAATTCTCATTTGCGCTTC |

**Table 4-1 Primers used for cloning LUX wt cDNA and LUX R146A cDNA**

### Hypocotyl length measurements

For hypocotyl measurement experiments, 3 independent homozygous lines were selected from F1 generation for both the *lux4* pLUX:LUX$^{R146A}$ and *lux4* pLUX:LUX. Seeds were sterilized by placing dry seeds in -80ºC for 2 hours and then washing with 70% ethanol and 100% ethanol subsequently. Seeds were left to dry in the plant cell culture hood for 2 hours to get rid of any remaining residual ethanol. Sterilized seeds were plated on 0.5 MS agar medium and cold stratified for 3 days in dark at 4ºC. After cold stratification seeds were moved to growth chambers (FitoClima D1200, aralab,). Seeds were grown in short day condition (8L:16D) at 22ºC and 27ºC with 70% humidity. Hypocotyls were measured after 7 days of growth by imaging plates containing seedlings with a flatbed scanner. Hypocotyl length were measured from the high resolution images obtained from the flatbed scanner using ImageJ software. Hypocotyl measurement were done for 20 plants each for every mutant lines.

**Flowering time measurements**

The plants were grown in soil under Long days of 24 h (16L: 8D). Time to flowering was taken as the number of rosette leaves at time of a 1-cm-high flower bolt.

**Petiole and growth measurement.**

As a measurement of growth response, petioles from mutants, complemented and wildtype *Col-0* plants were photographed and compared. Plants were grown in MS Agar Media under long day conditions (16L:8D) and were photographed at 6 leaves stage.

**RNA isolation and quantitative PCR**

Plants were grown growth chamber in Short Day (8L:16D) for 7 days in 0.5 MS Media and samples were harvested in an interval of 4 hours starting on light switching in growth cabinets. 4-6 seedlings were harvested for each line at each time point. Total RNA was extracted using RNeasy Plant mini kit (Qiagen) according to manufacturer's instructions. Total RNA (1µg) was treated with DNaseI (Roche) qRT-PCR was doing using iTaq® Universal SYBR® Green One-Step Kit from Bio-Rad following manufacturer's protocol. One step kit allows direct quantification from RNA with a cDNA synthesis step integrated before the mRNA quantification step with gene specific primers.
Expression of *PIF4* in different plant lines were determined through qRT-PCR with *PP2A* used as a control. qRT-PCR measurements were performed with a Bio-Rad CFX connect Real-Time system. Quantification was performed with the relative –ΔCt method, using *PP2A* for normalization. All quantification and statistical analysis were performed using CFX Maestro™ software (Bio-Rad).

| No. | Name of Primer | Sequence | Product length |
|---|---|---|---|
| 1 | PP2A qPCR FR | TATCGGATGACGATTCTTCGTGCAG | 173bp |
| 2 | PP2A qPCR RV | GCTTGGTCGACTATCGGAATGAGAG | |
| 3 | PIF4 qPCR FR | GCCAAAACCCGGTACAAAACCA | 125bp |
| 4 | PIF4 qPCR RV | CGCCGGTGAACTAAATCTCAACATC | |

**Table 4-2 Primers used for PIF4 qRT-PCR.**

# CHAPTER 5. **EXAMINATION OF *PIF4* PROMOTER *CIS*-ELEMENTS USING CRISPR-CAS9**

## 5.1 Introduction

In order to better understand regulation of *PIF4* expression by the EC, the promoter region of *PIF4* was analyzed. The *PIF4* promoter contains multiple *cis* elements including an E-box (CACATG), G-box (CACGTG) and the LBS (GATWCG) that have been implicated in PIF4 regulation with respect to temperature. Chromatin immune-precipitation followed by sequencing (CHiP- Seq) experiments of phyB reveal that the G-Box element present in the *PIF4* promoter was enriched at lower ambient temperature (17°C) versus 27°C and that PhyB binding has a repressive effect on *PIF4* expression (Jung et al., 2016). The same G-box and the E-Box element were enriched at higher ambient temperature (28°C) in BRASSINAZOLE RESISTANT 1 (BZR1) CHiP qPCR experiments, suggesting that BZR1 acts as an enhancer of *PIF4* expression at higher ambient temperatures (Ibañez et al., 2018). The LBS is bound by the EC through LUX at lower ambient temperature and has a repressive effect on *PIF4* expression with this effect decreasing as temperature increases (Ezer et al., 2017a; Nieto et al., 2015). Overall, this suggests a model in which the EC and PHYB repress PIF4 expression at lower temperatures and BZR1 activates PIF4 expression at higher temperatures via binding to the G-box and/or E-box elements. At higher ambient temperatures neither PHYB nor the EC would be bound, allowing BZR1 to activate PIF4 expression.

In order to understand the effect of different cis regulating elements present in the *PIF4* promoter, promoter deletion experiments using CRISPR/Cas9 were performed. In this chapter, results pertaining to promoter deletion at the G-box and LBS are presented.

## 5.2 Results

### 5.2.1 LUX Binding Site (LBS) CRISPR mutant plants.

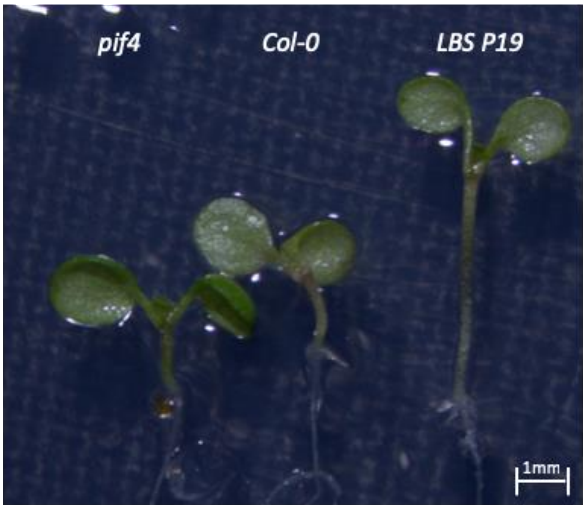
The LBS CRISPR plants were obtained using a plasmid construct that contains a codon optimised Cas9, a single guideRNA expression cassette and a seed coat GFP expression cassette for selection of positive transformants. As the CRISPR construct for the LBS CRISPR mutations were single guideRNA based, PCR samples from the LBS containing promoter region were amplified and sequenced for detecting single base pair mutations caused by CRISPR/Cas9. From the sequencing results it was found that there were mutations adjacent to the LBS in the PIF4 promoter. From sequencing results, 3 different mutations in 3 different plants were found. All these 3 plants had a similar early flowering and elongated hypocotyl phenotype. Seeds were collected from homozygous single mutant plant harboring mut2 mutation where there is a deletion of an adenine 5 base pairs upstream of the LBS **(Figure 5-1)**. This plant was chosen because the single base pair deletion is the closest to the LBS. This plant would be referred to as *LBS P19* from here on. All phenotype experiments were done on this mutant plant.


**Figure 5-1 Targeted mutagenesis of *PIF4* promoter**

Sequence alignment of PIF4 promoter fragment targeted for CRISPR mutation was done using sequencing results from different T1 plants. Three different mutants, mut, mut2 and mut3 were found. The LBS on PIF4 promoter is shown in green, base pair deletion is denoted by "_" while insertion is denoted by inserted base pair in red "N".

**Hypocotyl elongation response**

To understand the effect of the mutations close to the LBS site on the PIF4 promoter we did experiments to measure hypocotyl length of the plants harboring the mutation. Plants were grown in long day conditions (16h light and 8h dark) at 22ºC for 7 days before measuring hypocotyl lengths. Hypocotyl elongation phenotypes were studied with the *LBS P19* homozygous lines in Col-0 background to understand the effects of mut2 on growth. Hypocotyl lengths were determined for *pif4*, *LBS P19* and Col-0 lines. It was found that *LBS P19* seedlings had longer hypocotyls compared to Col-0 and *pif4* lines. The average hypocotyl length of *pif4* and Col-0 plants were found to be ~ 2mm while that of the *LBS P19* was found to be ~ 4mm. The average hypocotyl length for the *LBS P19* plants were double the wildtype plants at lower ambient temperature **(Figure 5-2)**.



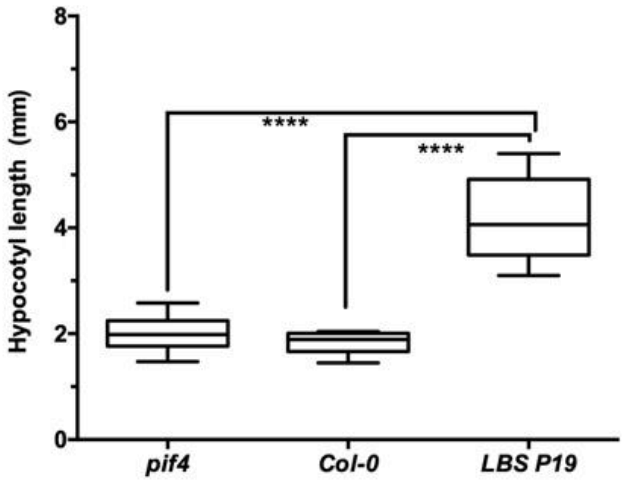




**Figure 5-2 Hypocotyl phenotype of *LBS P19* mutants**

**a)** Hypocotyl elongation Phenotypes of *pif4*, Col-0 and *LBS P19* (Scale bar is 2mm). **b)** Graph representing Hypocotyl lengths of *pif4*, Col-0 and *LBS P19* with Box plot (Median with 25 to 75 percentile) and whiskers(2.5 to 97.5 percentile) for 10-15 plants for each line. (**** represents P<0.001)

**Flowering time response.**

Flowering time measurement experiments were carried out to understand the effect of mut2 on *PIF4* promoter. Seeds from *LBS P19*, *Col-0* and *elf3-1* were planted in soil and placed at 22°C in long days to measure flowering time. For this experiment *elf3-1* plants were chosen as positive control. Previously it has been shown that *elf3-1* mutants flower early due to higher expression of *PIF4* (Mizuno et al., 2014; Nieto et al., 2015). From the flowering time experiments it was found that *LBS P19* mutants are early flowering. It was found that the wildtype *Col-0* plants flowered at approximately 17 rosette leaves while the *LBS P19* plants flowered earlier at approximately 7.2 rosette leaves which is very similar to *elf3-1* mutant plants that flowered at 7 leaves **(Figure 5-3)**.

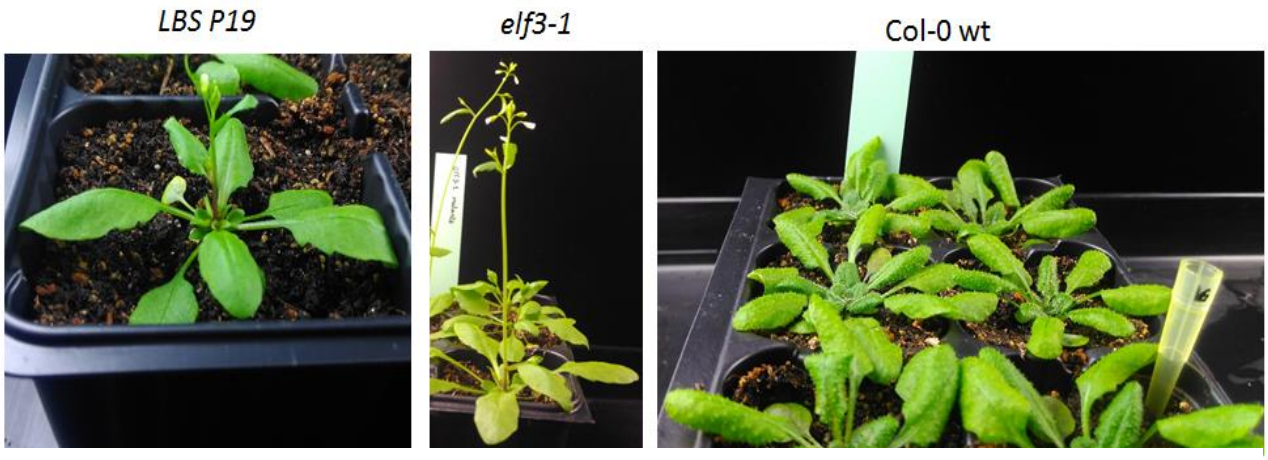



**Figure 5-3 Flowering phenotype of *LBS P19* compared with *elf3-1* and *Col-0* plants.**

**Gene Expression Studies**

From the hypocotyl elongation and flowering time experiments it was observed that the hypocotyl and flowering phenotypes of *LBS P19* plants were similar to plants grown at higher ambient temperature. It's known that at higher ambient temperature hypocotyl elongation and early flowering is promoted by higher *PIF4* gene expression (Capovilla et al., 2015; Mizuno et al., 2014). Hence it was decided to study *PIF4* gene expression along with *ELF3* and *ELF4* gene expression. Since we made changes adjacent to the EC binding site on the *PIF4* promoter, we expected that *ELF3* and *ELF4* expression would not be affected while *PIF4* expression should have a significant difference.

From the qRT PCR experiments it was found that relative gene expression of *ELF3* and *ELF4* did not differ dramatically between the Col-0 and *LBS P19*. However when relative gene expression of *PIF4* from Col-0 was compared with *LBS P19* a difference in expression was observed **(Figure 5-4)**. Hence we conclude that by mutating the upstream base pairs adjacent to the LBS it is possible to upregulate *PIF4* which stimulates hypocotyl elongation and early flowering in *Arabidopsis thaliana*. This mutation provides a way to make *Arabidopsis thaliana* flower early compared to wildtype plants without having to mutate any genes.

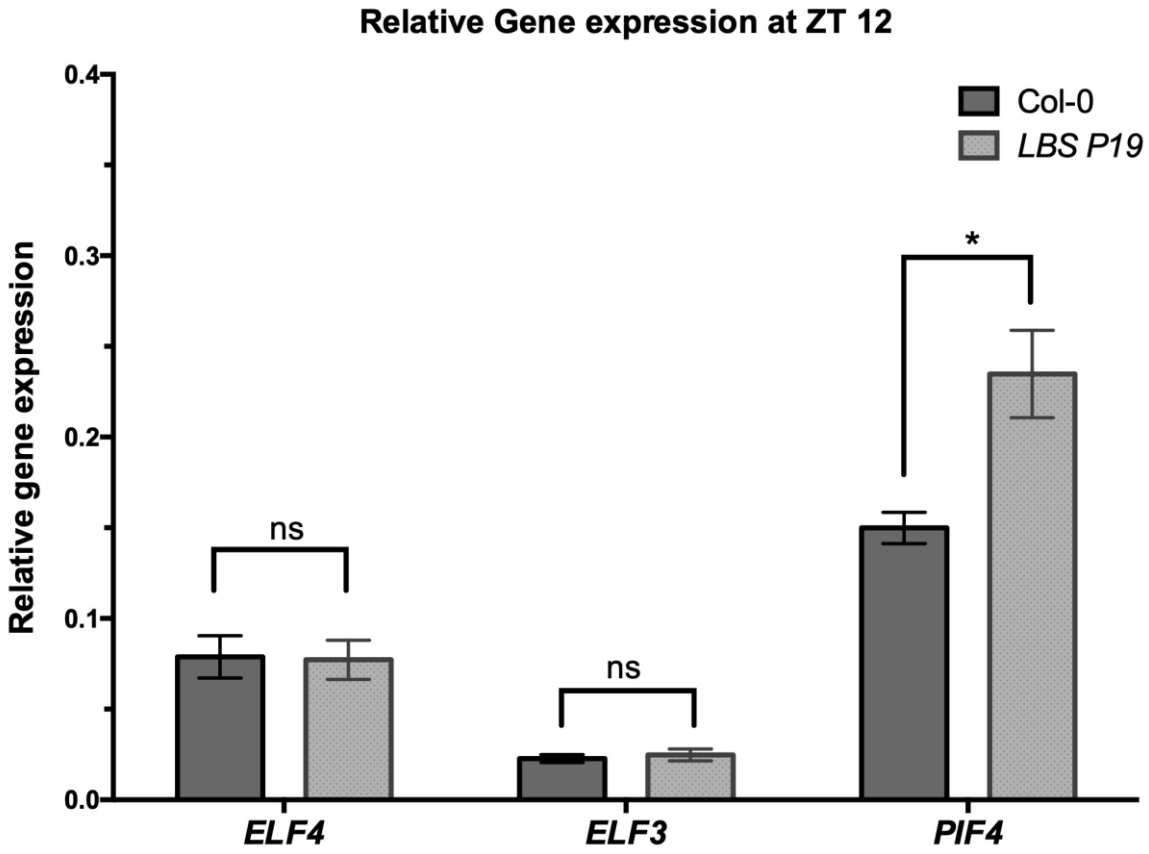



**Figure 5-4 Comparative Gene Expression of *ELF4, ELF3,* and *PIF4* in *Col-0* and *LBS P19***

Expression of *ELF3, ELF4* and *PIF4* at ZT-12 compared between *Col-0* and *LBS P19* plants. Samples were harvested from 10 days old seedling grown in long days condition at 22ºC. Expression of mentioned gene were normalized with actin expression. P values represent results from a student's t test using the Holm-Sidak method. (* $P<0.01$ ns = non-significant)

### 5.2.2 Results from G-Box CRISPR plants.

The G-box CRISPR mutant plants were obtained using a plasmid construct that was similar to the LBS CRISPR plants. Only difference being that instead of using a single guide RNA, dual guides were used to delete the G-box present in the *PIF4* promoter region. Col-0 wildtype plants were transformed with the G-Box plasmid construct to obtain G-box CRISPR plants. Stable homozygous lines were obtained and genotyped. Plants were screened through PCR for deletion mutations around the G-box and positive plants were selected for phenotype experiments. The hypocotyl, flowering time, silique length and root phenotypes were observed for the mutant lines. The results from the phenotypic observations are presented in the next sections.

**Genotyping the G-box CRISPR plants.**

The CRISPR construct for the G-box CRISPR mutations were dual guide RNA based. Hence, it was possible to screen for deletion mutants using PCR primers amplifying the G-box containing promoter region of *PIF4*.PCR primers PGP23R and PGP23F (Details in materials and methods) should amplify 586 base pairs from the PIF4 promoter in-case of a wildtype type plant. For the G-box deletion mutants, it should amplify 102 base pairs less, so the PCR product from a G-box mutant would be of 484 base pairs. From the $T_0$ transformants, $T_1$ plants were selected on the basis of seed coat GFP expression. From these $T_1$ plants genomic DNA was collected and PCR amplification was done to screen the *PIF4* promoter for G-box deletion. From genotyping it was found that 4 plants were homozygous for these deletion while two plants were heterozygous **(Figure 5-5).**These plants would be referred to as *G-Box 23* mutants from here on.

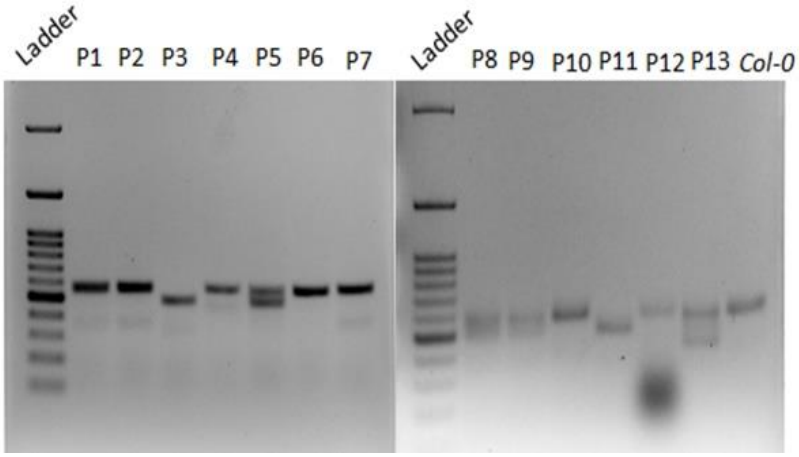



**Figure 5-5. PCR amplicon from *PIF4* promoter of different F1 plants for genotyping**

Genomic DNA from Different F1 G-Box CRISPR plants were PCR amplified with primer PGP23F and PGP23R screening for deletion of G-Box region of PIF4 Promoter. P3 and P11 are homozygous mutants for G-Box deletion while P5, P8, P9, P11 and P13 are heterozygous G-Box Crispr Mutant plants

**Hypocotyl elongation of G-Box CRISPR mutants**

Plants respond to increase in ambient temperature by hypocotyl elongation. Hence G-Box CRISPR mutant seeds were germinated at 27ºC along with Col-0 plants to check for hypocotyl elongation. From the hypocotyl elongation experiments it was found that at 27ºC the G-Box CRISPR mutant had smaller hypocotyl compared to Col-0 plants. While the average hypocotyl length of Col-0 Plants were 2.15 mm, the average hypocotyl lengths of G-Box CRISPR plants were found to be 1.08mm. Hence we conclude that hypocotyl elongation due to higher ambient temperature seen is Col-0 wild type plants is annulled with the G-Box CRISPR mutation **( Figure 5-6)**.

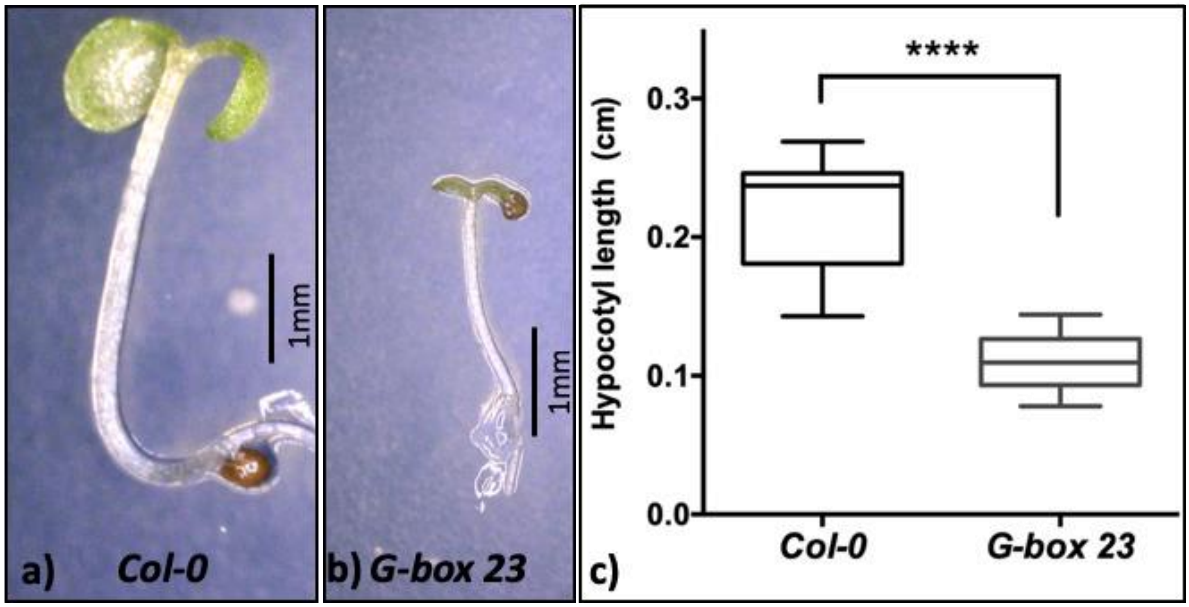



**Figure 5-7. Hypocotyl phenotype of G-Box CRISPR plants compared to Wildtype Col-0 plants.**

Representative hypocotyl images of a) Col-0 and b) G-Box CRISPR mutant plants. c) Hypocotyl length of ~15 plants each for Col-0 and G-Box CRISPR were measured and average hypocotyl length is presented in the Tukey Boxplot. (**** represents a significant difference with P<.001 )

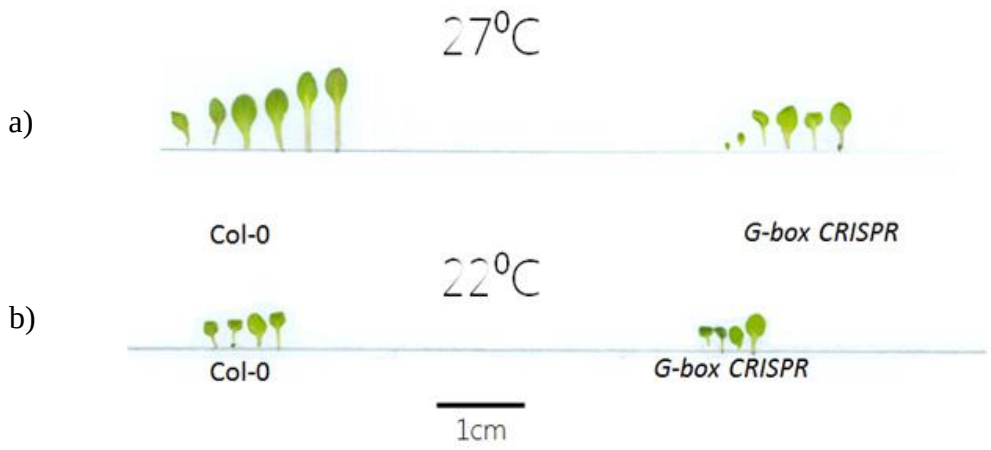



**Figure 5-6 Petiole elongation in G-Box CRISPR plants compared with *Col-0***

*Col-0* plants have bigger petioles at 27ºC compared to G-Box CRISPR Plants. However not much difference was seen between the petioles of G-Box CRISPR plants compared to *Col-0* plants when petioles phenotypes were compared at 22 ºC. (The petioles images are representative of petioles collected from whole plants in 6 leaves stage grown at 26ºC and 4 leaves stage for plants grown at 22ºC

**Petiole elongation response of G-Box CRISPR mutants**

The normal response of wildtype Col-0 plants to higher temperature is to increase petiole length. This response balances the risk of heat damage versus water shortage. The risk of heat related damage can be averted through evaporation via the stomata and requires water availability for optimum efficacy. The majority of water is lost through transpiration. Longer petioles lead to fewer stomata at higher temperature to control the rate of transpiration and minimize water loss. The architectural changes in the leaf help compensate for the reduced number of stomata to allow for cooling.

For the G-box mutants and wildtype plants grown at 22ºC little difference was observed in petiole length at four leaf stage. However for the plants grown in 27ºC keeping all conditions identical, it was observed that G-box CRISPR plants showed smaller elongation in petiole length compared to *Col-0* plants **(Figure 5-7)**. The phenotype of these plants was similar to those seen at 22°C for the wildtype plants. However there was a small amount of petiole elongation observed at 27°C for the G-Box CRISPR plants compared to those grown at 22°C.

**Flowering time response of G-Box CRISPR plants**

Flowering is also linked to temperature. It's known that at 27°C *Arabidopsis* flowers earlier than at 22°C (Capovilla et al., 2015). We hypothesized that with the G-Box mutation, the plants should be able to flower later than the wildtype type plants at 27°C. Hence *G-box* mutants and wildtype plants were grown at 27°C. It was found that indeed the wildtype plants flowered earlier than the *G-Box* mutants . It was observed that while *Col-0* plants flowered at 7~ 8 rosette leaves, the *G-Box* mutants flowered much later at 12~14 rosette leaves, similar to wildtype plants grown at 22ºC **( Figure 5-8)**.


**Figure 5-8 Flowering phenotype of *G-Box* CRISPR mutant plants grown at 27ºC.**

Col-0 plants flower earlier than the G-Box CRISPR mutant plants at 27 ºC. Col-0 and G-Box CRISPR plants were grown in Long day condition (16h light and 8h darks) at 27 ºC.

**G-Box CRISPR plant have larger seed Pods at higher ambient temperature.**

Plant productivity depends upon seeds produced hence it's linked to the size and number of siliques produced. The silique length of *G-box* mutants and wildtype plants at 27 ºC were observed to see if there was any difference in the silique lengths. It was found that silique size was altered with the G-box mutation, displaying 100% increase in the silique length when compared with the parental Col-0 silique. Mutant *G-box 23* plants had longer siliques than parental Col-0, while the *G-Box23* mutant had an average silique length of 6.7mm the parental Col-0 lines had an average silique length of 3.3mm (**Figure 5-9** ).

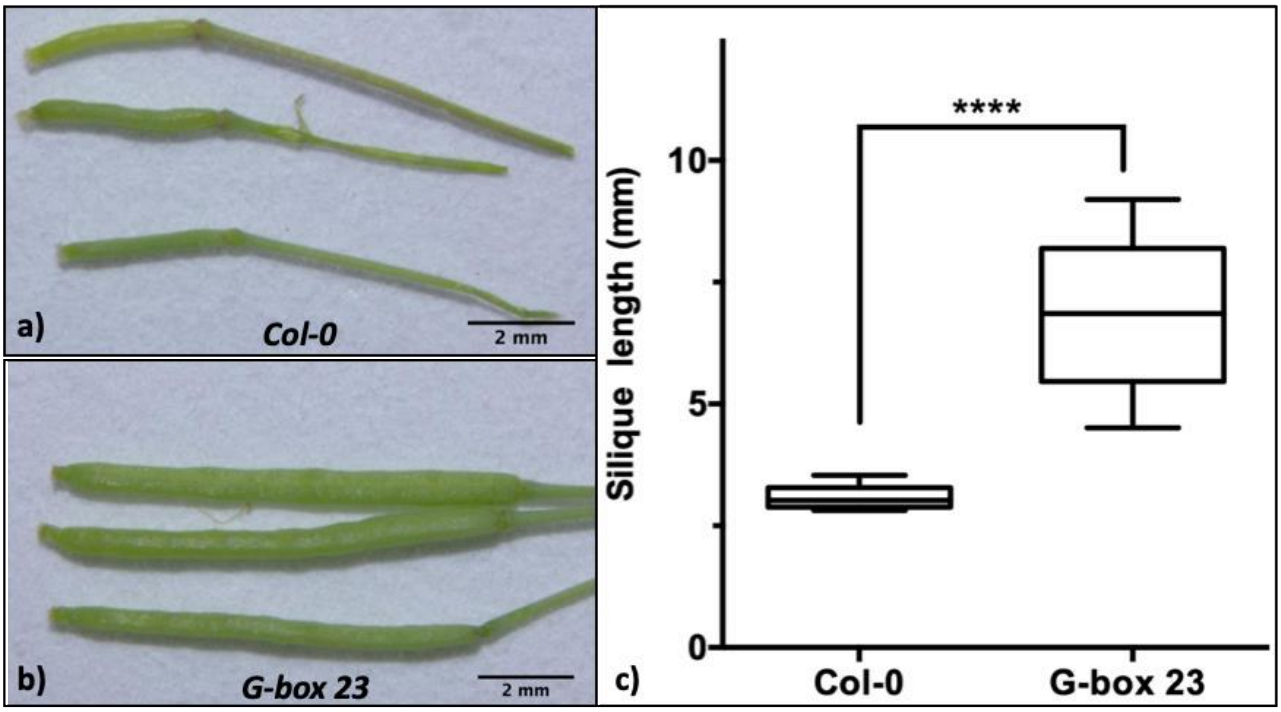



**Figure 5-9 Silique Phenotype of G-Box CRISPR Plants grown at 27ºC**

**a)** Silique from *Col-0* plants compared with **b)** siliques from *G-Box* CRISPR plants grown at 27 ºC. **c)** Box plot representing silique length observed in parental Col-0 plants compared with *G-Box* CRISPR plants (10 siliques were measured for each line). Plots represent tukey tested data from each lines.(**** represents significant difference of P <.001).

**G-Box CRISPR plant have longer roots and more secondary roots.**

For plants to survive in higher ambient temperature, it's important that they can secure water resources for cooling through transpiration. Having a longer primary root and more lateral roots helps to efficiently uptake water. This is likely to be an important trait when growing under higher ambient temperature. It was found that compared to parental Col-0 plants, the *G-Box* mutants had longer primary roots and more lateral roots emerging. Average root length of 12 days old Col-0 plants was 1.33cm

compared to 2.66 cms for the *G-box23* mutants. Also it was observed that the G-box mutants had larger number of lateral roots emerging compared to *Col-0* wildtype plants **(Figure 5-10)**.

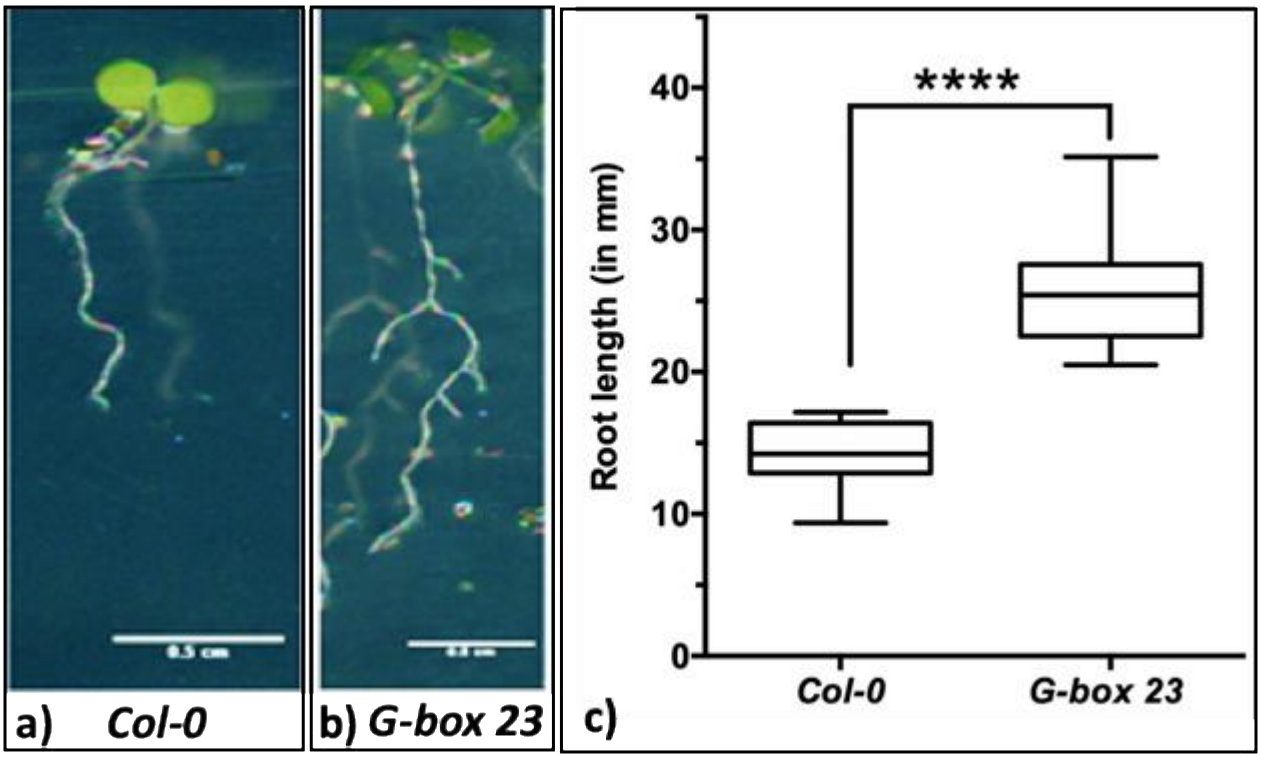



**Figure 5-10 Root phenotype of Col-0 plants compared with *G-Box* CRISPR plants.**

Root length of *Col-0* plants is smaller compared to *G-Box* CRISPR plants. Also there are more emerging lateral roots in the *G-Box* CRISPR plants compared to *Col-0* plants.

## 5.3 Conclusion

The results presented in this chapter indicate that by modifying the *PIF4* promoter, ambient temperature responses in plants can be modified. Two cases were presented in this chapter. Initially it was shown that by modifying LBS site on the *PIF4* promoter, higher ambient temperature responses could be evoked at lower ambient temperature. In the second case it was shown that by deleting the G-box containing region of *PIF4* promoter, higher ambient temperature responses could be mitigated and plants would grow normally as they would at lower ambient temperature.

These results, propose that there are multiple transcription factors that bind to the *PIF4* promoter and might be acting antagonistically. BZR1 and EC might we playing an antagonistic role in controlling temperature related growth phenotype through *PIF4* expression. Previous studies have shown that at higher ambient temperature, BZR1 is active and can upregulate *PIF4* gene expression moreover at higher ambient temperature, the EC repression on PIF4 is removed and this leads to adverse phenotypes observed. BZR1 bind to the G-box effectively to upregulate *PIF4* expression at higher ambient temperature, hence by deleting the G-box in the *PIF4* promoter, BZR1 upregulation of *PIF4* could be controlled even in the absence of EC based repression. While at lower ambient temperatures, where EC actively represses *PIF4* expression by binding to the LBS in the *PIF4* promoter to repress growth, mutations close to the LBS could perturb EC binding and hence remove the repression on growth.

These mutation could be important for engineering crop plants to cope up with the negative effect of higher ambient temperature on growth and productivity.

## 5.4 Materials and methods

### PIF4 promoter

cactaataattgttattacgacaacagaattaattgcctatttaaaacgaaatgaatcatcgaagaacatactagatctttttccaatttacaaccaaa
aaaaacattattaggtgagtagtaaaaagattaaggttatcagaaactaactgtagtagtatgttgtgtgaaattcttttgtatttcttattaagccaag
ggtgccctttcaatgcaacgtgataaacgccaaaaacgatgatatgaacaattataaactcgttgagagcattgaactcggataataaatcatctt
ttatatacatcgtagataacaaacaaacacgtaattaaatttgacgtatagcaaaagacttgaagaataaaacgtcaagttaaagataattttggta
tatatgagaaaggtatcgacaaaaaccataacgctatagatgattgtgatttgacaaaaacaccctcaaatcattgtttcagagttttttagataa
ggtacagataagaaaccacctctaaaaatcaagcaatagatctcatcgcttaaaagaagagagagatcttcacttgtatgtgtcccactgattcc
aacacaatgtcccagaacttgccacgtgtcgttcatttcaaaagattgcagtactgttgtccctagagaatcattatctcccctcgctgtaatatcttt
atgctcctgtcactttctgtctgtacccaaaagaagtaatgaacctctctcatcttcttcttctgtttctttcatgttttgtgagttgtttctcaacaattt
tctggtctcttagagtgagaggagagagatagagagttgtgttgggcgtggaacttggactagttccacatatcaggttatatagatcttctctttc
aacttctgattcgtccagaagctttcctaatctggtcagtagtactcttttatacgggttttggttttataagatgtggctatatttggaaataactattt
tgcaagctttcctagattgccagaatataaaaaaagatgtttaacaagagaacggactcatggacttgctttaaattttaattattttaaaatcattct
ataatgattagagtaaataaactattaggactctgaattataaaattcgattttatatatgctcctccttgtatctcttaatcataagttatcattagctct
gttcactagtgctataaaatatattctgaggttaataaaacttttttttctttttattttgaaatgtctccagagatctgac**ATG**

In the above sequence the spacer sequences used in the gRNA design are marked in brown and pink for the G-Box while the spacer for LUX Binding Site (LBS) is marked in red. The Genotype primers are marked in cyan. In the schematic figure explained below the G-Box is marked in green. The LBS is marked in Yellow. The schematics explaining the CRISPR target is given below **(Figure 5-11)**.

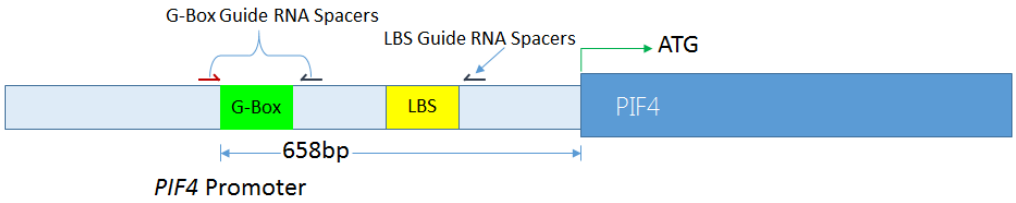



**Figure 5-11 Schematics of PIF4 promoter loci used for CRISPR/Cas9 guided mutations.**

Dual Guide RNA Spacers used for deletion of G Box are marked with split arrows in red and black. Single Guide RNA Spacer used for LBS targeting is marked in black.

### Plasmid Construct design for G-Box and Lux Binding Site

The system is designed to use CRISPR Cas9 for modifying plants and use GFP based selection for selecting positive seeds from the plants transformed with the vector. The system facilitates use of multiple gRNA's. The whole system is a two vector system comprising of a guide RNA vector for multiple guide assembly and the final Cas9 vector comprising of the eGFP and Cas9. Multiple gRNA's are assembled on the gRNA vector and then the whole GuideRNA cassette is excised and cloned into

the pGreenCRISPR vector. Later this final vector is transformed with the floral dip method. Seeds from the plants transformed with the final vector are selected for GFP signals and the positive seeds carry the CRISPR Cas9 system intended.

1000 base pairs (bp) upstream of the *PIF4* transcription start site (TSS) was selected to search for the G-Boxes (CACGTG) and LUX binding site (LBS). One G-Box was located 655 bases upstream and LBS was located 374 bases upstream of *PIF4* TSS. Sequence flanking 200bp upstream and downstream of this G-box and LBS was processed through CHOPCHOP server for locating probable GuideRNA spacers with minimal off-target effects.

Spacers with a score under 20 were selected as suggested by the program.(scores above 20 suggest that there might be off site targets apart from the desired targets) Following specific over-hangs were added on the primers for ligation:

For forward primer: 5'-GATT XXXXXXXXXXXXXXXXXXX-3'
For reverse primer: 5'-AAAC XXXXXXXXXXXXXXXXXXX-3'

Protospacer sequences used for G-box Guide RNA are as follows.

| Name | Sequence |
|---|---|
| G-box CRISPR 2 FR | GATTGCATAAAGATATTACAGCGA |
| G-box CRISPR 2 RV | AAACTCGCTGTAATATCTTTATGC |
| G-box CRISPR 3 FR | GATTCAAGTTCTGGGACATTGTGT |
| G-box CRISPR 3 RV | AAACACACAATGTCCCAGAACTTG |
| LBS CRISPR 1 FR | GATTGTCCAGAAGCTTTCCTAATC |
| LBS CRISPR 1 RV | AAACGATTAGGAAAGCTTCTGGAC |

**Cloning procedure**

**Protospacer annealing**

Protospacers mentioned were synthesized at Eurofins genomics. The forward and reverse protospacers were annealed using 10x annealing buffer whose composition is mentioned below.

10X annealing buffer:
1mL of 1M Tris pH 7.5 (100mM)
3mL of 5M NaCl (1.5M)
Water up to 10mL

Annealed protospacers were obtained using the following procedure

Mixed together:
5uL oligoFr stock (100uM)
5uL oligoRv stock (100uM)
5uL Annealing buffer
35uL water

The reaction mix was incubated 3-4 minutes at 95˚C and was let to cool down slowly in the heating block over-night.

**Ligation of spacers in pAtU6-26:gRNA plasmid.**

For obtaining G-Box and LBS specific guideRNAs, the double stranded protospacers were ligated into the pAtU6-26:gRNA plasmid. The spacers were cloned between BbsI site through restriction digestion and ligation following standard NEB protocols.

The ligated plasmids were transformed in competent bacteria to obtain the gRNA constructs.
We chose to use dual guides based gRNA constructs for targeting G-Box. To do so the plasmid containing the first guide was digested with SpeI and KpnI while the plasmid containing the second guide was digested with XbaI and KpnI. The digestion reaction was resolved on 1% agarose gel, the bands were cut out and purified using QIAquick gel extraction and PCR cleanup kit from Qiagen®. Fragments obtained were ligated overnight using T4 DNA ligase and transformed in competent bacteria.

**Final Ligation into pGreenCrispr**

The cassette containing the guide(s)RNA(s) from pBSK:AtU6-26:gRNA were finally cloned into pGreenCRISPR a vector developed in the lab containing codon optimized Cas9 under ubiquitin promoter and the AtS2:eGFP sequence. (pGreenCRISPR vector has Kanamycin Resistance for selection in bacteria)

Following protocol was followed for restriction digestion.

pBSK:AtU6-26:gRNA(s) digestion:
2ug of plasmid
2uL of buffer CutSmart
1uL of KpnI-HF
1uL of XbaI
Water up to 20uL

pGreenCRISPR digestion:
2ug of plasmid
2uL of buffer CutSmart
1uL of either KpnI-HF
1uL of XbaI
1uL of rSAP
Water up to 20uL

Reaction was incubated for 3 hours at 37˚C. The digestion products were analyzed on a 1% agarose gel, the bands were excised and purified using QIAquick gel extraction and PCR cleanup kit from Qiagen®.

Fragments obtained were ligated overnight using T4 DNA ligase and transformed in competent bacteria. Bacteria were grown in LB Agar plates containing kanamycin for selection of positive transformants.

**Genetic transformation**

Genetic transformation was performed by introducing constructs into *Agrobacterium tumefaciens* strain GV3101, which was then used to transform wildtype plants using floral dip method (Clough and Bent, 1998)

Comment: other methods of transformation can be used, such as electroporation and T-DNA insertion free transformation using CRISPR/Cas9 complex assembled in-vitro. (In this study, the floral dip method was used because this method is the most optimized method for efficiency and speed for generating mutant plants.)

**Genotyping of mutants**.

Singles leaves were harvested from 20 days old plant. Leaves were frozen in liquid nitrogen. The leaves were disrupted in a TissueRuptor II homogenizer using glass beads. DNA was extracted from these samples using CTAB DNA extraction method (Richards et al., 1994). To verify if the plants were homozygous or heterozygous for the deletion of G-Box deletion, PCR was performed using primers flanking the expected deletion sites. Primers used for this purpose are as following.

| | |
|---|---|
| PIF4 Promoter Genotype FR | 5'-TCAGAGTTTTTTTAGATAAGG-3' |
| PIF4 Promoter Genotype RV | 5'-GCAAGTCCATGAGTCCGTTC-3' |

Expected amplicon size from wildtype plants is 586 bp. The deletion results in removal of 102 bp from the amplicon hence if deletion is present the amplicon size is supposed to be 484bp. It's possible that plant might be homozygous or heterozygous for the deletion. Hence heterozygous plants should produce two amplicon, one pertaining to 586bp and the other pertaining 484bp. The homozygous mutants should produce one amplicon of 484bp.
All PCR products were resolved on a 1.5% agarose gel containing GelRed® to view amplicons.

**Phenotyping of mutants.**

*Col-0 (*wild type plants) and mutant plants were grown in soil, in long day chambers at 22˚C for 10 days and were then transferred to 27˚C to score leaf, silique size and flowering time phenotypes.

**RNA isolation and quantitative PCR**

Plants were grown in LD for 10 days in MS Media and samples were harvested in an interval of 4 hours starting on light switching in Percival cabinets. 4-6 seedlings were harvested for each line at each time point. Total RNA was extracted using RNeasy Plant mini kit (Qiagen) according to manufacturer's instructions. Total RNA (1µg) was treated with DNaseI(NEB). For qRT-PCR, cDNA was generated synthesized from 1µg of DNaseI treated RNA using iScript$^{TM}$ cDNA synthesis kit (Bio-Rad, 1708891)

using manufacturer's protocol. Expression of *PIF4* in different plant lines were determined through PCR with *ACTIN* used as a control. qRT-PCR measurements were performed in a Bio-Rad CFX384™ Real-Time system with SsoFast™ EvaGreen®Supermix (Bio-Rad). Quantification was performed with the relative –ΔΔCt method, using *ACTIN* for normalization. All quantification and statistical analysis were performed using CFX Maestro™ software (Bio-Rad).

# CHAPTER 6. **CONCLUSION AND FUTURE PERSPECTIVES**

## 6.1 Conclusion

The EC plays an important role in the circadian clock by acting as a bridge to connect the evening and morning loops and forming a key circuit in the plant circadian system (Pokhilko et al., 2012). Additionally, the EC acts as a hub for integrating environmental cues and relaying this information directly to growth and developmental pathways through direct effects on target genes. The EC does this at least in part via its nighttime repression of *PIF4*, a master regulator of thermoresponsive growth, plant immunity and reproductive development (Gangappa et al., 2017; Koini et al., 2009; Kumar et al., 2012). This raises the intriguing possibility of using the EC to tune temperature sensitive growth. However, this attractive goal requires an understanding of the molecular basis of EC formation and activity. Prior to these studies, the molecular mechanisms of EC activity were poorly defined. The major objective of this thesis was to provide a molecular model of LUX, ELF3 and ELF4 interactions in EC formation and activity. The work presented here describes the structure of the DNA binding domain of LUX, the critical domain for EC targeting to its cognate DNA. The binding affinity of LUX for different target motifs was determined *in vitro* and mutagenesis was performed *in vitro* and *in vivo* with the goal of altering plant thermoresponse. Finally, different *cis*-elements including a G-box and the LBS on the *PIF4* promoter were targeted for mutagenesis using CRISPR-Cas9, resulting in plants with altered growth and development under different environmental conditions.

There were major challenges to the successful realization of the thesis project. The first difficulty was the recombinant expression and purification of EC components for *in vitro* studies. The expression of full length and domains and ELF4 was successful in *E.coli* after extensive buffer optimisation. ELF3 expression was not successful in *E.coli*. This challenge required different strategies including different expression systems and the use of a construct library for screening thousands of possible constructs for expression and solubility. Using the ESPRIT platform (Yumerefendi et al., 2010), soluble domains of ELF3 were obtained. However, the middle domain of ELF3, which is responsible for interaction with ELF4 based on yeast 2-hybrid screening (Herrero et al., 2012)could not be obtained through this method. To obtain full length ELF3, the protein was co-expressed with ELF4 and LUX in insect cells. This resulted in high levels of expression, however the protein was insoluble. To obtain soluble ELF3, the protein was extracted in a 8M urea buffer and a rapid refolding protocol was developed to obtain soluble ELF3. Upon obtaining all the EC components, *in vitro* EC reconstitution experiments were performed. While initial EC reconstitution experiments using LUX, ELF4 and soluble ELF3 fragments from the ESPRIT library did not result in a core EC complex, small scale refolding of all partners was successful. Using a denaturing-renaturing protocol, active EC was obtained and tested by band shift assays using a labeled DNA fragment containing an LBS. The EC binding resulted in a supershift of the DNA. These results are the first successful example of reconstitution of the entire EC in vitro and represent an important advance in the study of the complex. Electrophoeretic mobility shift assays were performed with the reconstituted EC to probe the function of each component. It was demonstrated that *in vitro*, ELF3 is able to attenuate the LUX-DNA interactions, at least for certain LBS sequences, suggesting that the formation of ELF3-LUX complex is not competent to bind DNA. Further, this suggests a possible sequestering role of ELF3 in the ELF3-LUX interaction which has also been observed in case of ELF3-PIF4 (Nieto et al., 2015). Further, it was demonstrated that ELF4 has a modulatory effect on the sequestration activity of ELF3 and is able to restore the DNA binding activity of the tripartite LUX-ELF3-ELF4 complex. This highlights the importance of ELF4 in EC formation. The critical importance

of ELF4 is further corroborated by previous modelling studies that show *ELF4* transcript levels are equally as powerful as using the full EC (*LUX, ELF3* and *ELF4* transcripts) to predict EC based target gene repression and are more reliable than *ELF3* transcripts alone (Ezer et al., 2017a). Hence all three components, ELF3, ELF4 and LUX are necessary for both DNA-binding and target gene regulation.

*In vitro* and structural studies presented here illustrate that LUX provides the EC with required specificity and affinity to target its cognate binding sites. The MYB domain of LUX performs direct base read out of the core LBS sequences. From the crystal structure of LUX it was observed that the plant specific signature sequences, SH(A/L)QK(F/Y) of helix 3, provides the majority of direct interaction with the DNA bases in the major grove. An N-terminal arginine, Arg146, which is part of the flexible extension of the MYB domain, is important for intercalation into the minor grove and acts as a DNA clamp. Several other TFs including homeodomain TFs and MADS TF family also possess arginine residues in the flexible extensions (Bürglin and Affolter, 2016; Käppel et al., 2018). The two LUX crystal structures presented demonstrate the flexibility of Arg146 with the DNA as this residue adopts different side chain configurations with different hydrogen bonding interaction patterns. To probe the importance of Arg146, site directed mutagenesis was carried out to change the Arginine to an Alanine. EMSAs done with this mutant protein version demonstrate poorer DNA binding for all target sequences tested *in vitro* when compared with the wild type version. However, the R146A mutants were still able to specifically bind their target sequences. This mutation was then tested *in planta* to determine whether growth and development would be affected my attenuated LUX binding affinity. It was observed that at 22°C the flowering and hypocotyl phenotype of *lux-4* plants transformed with *pLUX:LUX*$^{R146A}$ were intermediate between wildtype and *lux-4* plants. At 27°C the EC activity is greatly reduced and this was reflected in the similar phenotypes of wild type, *lux-4,* and *lux-4* transformed lines. As predicted, the R146A mutation resulted in accelerated growth but still retained thermo-responsiveness with a phenotype intermediate between wt and *lux-4*. The observed phenotypes were likely due at least in part to changes in *PIF4* expression. *PIF4* is a hub of effecting thermal responses and it is an important direct target of EC implicated in hypocotyl elongation and thermoresponsive growth (Ezer et al., 2017a; Gangappa et al., 2017; Kumar et al., 2012). Phenotypes similar to elevated *PIF4* expression was observed for the *lux-4* mutant transformed with *pLUX::LUX*$^{R146A}$. Elevated *PIF4* expression was observed in the *lux-4 pLUX::LUX*$^{R146A}$ plants as compared to wild type. The *lux-4 pLUX::LUX*$^{R146A}$ have a similar phenotype to mild *PIF4* over-expressors which have an early flowering and elongated hypocotyl phenotype at 22°C (Gangappa et al., 2017).

Since *PIF4* expression is critical for thermoresponsive growth, I wanted to understand how different *cis* elements in the *PIF4* promoter contribute to *PIF4* expression. The LBS and G-box are two important elements present in the PIF4 promoter. Multiple transcription factors such as EC, PHYB, BZR1 bind to these elements (Ezer et al., 2017a, 2017b; Nieto et al., 2015) . Hence the *cis* elements could be critical to alter thermal response in plants without targeting coding sequences. Mutations using the CRISPR-Cas9 system were made adjacent to the LBS site using a single guide and the G-box region was removed via a two-guide construct. A warm temperature phenotype was observed for the *LBS* CRISPR mutants at 22°C which is likely due to decreased EC-based repression of PIF4 activity. PhyB also has a repressive effect on *PIF4* expression 22°C by binding to the G-box element (Jung et al., 2016; Legris et al., 2016) hence with the G-box deletion a similar phenotype as the LBS mutant was expected. Surprisingly, the opposite phenotype was observed for the G-box mutants. The G-box mutants did not have a significant

growth difference at 22°C versus wt, however at 27°C they exhibited a phenotype similar to plants grown at a lower ambient temperature. Based on recent studies that show the G-box element in the PIF4 promoter is also important for BZR1 binding at higher temperatures and that this TF acts as a PIF4 activator, it is likely that mutations in the G-box result in both a lack of PHYB and BZR1 binding, accounting for the observed phenotype. Furthermore, the root phenotype observed in the G-box mutant was similar to the *bzr1-D* mutant (Singh et al., 2014). It will be necessary to determine which factors can still bind at the mutated PIF4 promoter and whether this binding is temperature dependent.

## 6.2 Future experiments

Structure –based protein engineering of LUX with increased and decreased affinity for its binding sites as well as targeted CRISPR mutations to distinct *cis*-elements in the *PIF4* promoter are potential strategies for plant modification. Engineering altered thermoresponse in crop plants based on EC activity and/or *PIF4* expression is feasible based on the initial studies presented here. The EC and PIF4 are found and active in different crop plants ranging from tomatoes to cabbage. With the changing climate due to global warming, accelerating plant development or delaying plant development with respect to changing temperatures is an important goal and necessary for long-term food security. The studies presented here provide a foundation for further investigation of the EC and/or PIF4 as targets for altering thermoresponsive growth in a desired manner for improved crop productivity under different temperature regimes.

# REFERENCES


Alabadí, D., Oyama, T., Yanovsky, M.J., Harmon, F.G., Más, P., and Kay, S.A. (2001). Reciprocal regulation between TOC1 and LHY/CCA1 within the Arabidopsis circadian clock. Science *293*, 880–883.

Allada, R., White, N.E., So, W.V., Hall, J.C., and Rosbash, M. (1998). A mutant Drosophila homolog of mammalian Clock disrupts circadian rhythms and transcription of period and timeless. Cell *93*, 791–804.

Bae, K., Lee, C., Hardin, P.E., and Edery, I. (2000). dCLOCK is present in limiting amounts and likely mediates daily interactions between the dCLOCK-CYC transcription factor and the PER-TIM complex. J. Neurosci. Off. J. Soc. Neurosci. *20*, 1746–1753.

Bieniossek, C., Richmond, T.J., and Berger, I. (2008). MultiBac: Multigene Baculovirus-Based Eukaryotic Protein Complex Production. Curr. Protoc. Protein Sci. *51*, 5.20.1-5.20.26.

Bognár, L.K., Hall, A., Adám, E., Thain, S.C., Nagy, F., and Millar, A.J. (1999). The circadian clock controls the expression pattern of the circadian input photoreceptor, phytochrome B. Proc. Natl. Acad. Sci. U. S. A. *96*, 14652–14657.

Box, M.S., Huang, B.E., Domijan, M., Jaeger, K.E., Khattak, A.K., Yoo, S.J., Sedivy, E.L., Jones, D.M., Hearn, T.J., Webb, A.A.R., et al. (2015). ELF3 Controls Thermoresponsive Growth in Arabidopsis. Curr. Biol. *25*, 194–199.

Brown, S.A., Kowalska, E., and Dallmann, R. (2012). (Re)inventing the Circadian Feedback Loop. Dev. Cell *22*, 477–487.

Buhr, E.D., and Takahashi, J.S. (2013). Molecular components of the mammalian circadian clock. Handb. Exp. Pharmacol. 3–27.

Bürglin, T.R., and Affolter, M. (2016). Homeodomain proteins: an update. Chromosoma *125*, 497–521.

Candolle, A.P. de (1832). Physiologie végétale, ou Exposition des forces et des fonctions vitales des végétaux ... (Béchet jeune).

Capovilla, G., Schmid, M., and Posé, D. (2015). Control of flowering by ambient temperature. J. Exp. Bot. *66*, 59–69.

Chow, B.Y., Sanchez, S.E., Breton, G., Pruneda-Paz, J.L., Krogan, N.T., and Kay, S.A. (2014). Transcriptional regulation of LUX by CBF1 mediates cold input to the circadian clock in Arabidopsis. Curr. Biol. CB *24*, 1518–1524.

Clough, S.J., and Bent, A.F. (1998). Floral dip: a simplified method for Agrobacterium-mediated transformation of Arabidopsis thaliana. Plant J. Cell Mol. Biol. *16*, 735–743.

Cohen, S.E., and Golden, S.S. (2015). Circadian Rhythms in Cyanobacteria. Microbiol Mol Biol Rev *79*, 373–385.

Covington, M.F., Maloof, J.N., Straume, M., Kay, S.A., and Harmer, S.L. (2008). Global transcriptome analysis reveals circadian regulation of key pathways in plant growth and development. Genome Biol. *9*, R130.

Cumming, B.G., and Wagner, E. (1968). Rhythmic processes in plants. Annu. Rev. Plant Physiol. *19*, 381–416.

Cyran, S.A., Buchsbaum, A.M., Reddy, K.L., Lin, M.-C., Glossop, N.R.J., Hardin, P.E., Young, M.W., Storti, R.V., and Blau, J. (2003). vrille, Pdp1, and dClock form a second feedback loop in the Drosophila circadian clock. Cell *112*, 329–341.

Damiola, F., Le Minh, N., Preitner, N., Kornmann, B., Fleury-Olela, F., and Schibler, U. (2000). Restricted feeding uncouples circadian oscillators in peripheral tissues from the central pacemaker in the suprachiasmatic nucleus. Genes Dev. *14*, 2950–2961.

Darlington, T.K., Wager-Smith, K., Ceriani, M.F., Staknis, D., Gekakis, N., Steeves, T.D., Weitz, C.J., Takahashi, J.S., and Kay, S.A. (1998). Closing the circadian loop: CLOCK-induced transcription of its own inhibitors per and tim. Science *280*, 1599–1603.

De Caluwé, J., Xiao, Q., Hermans, C., Verbruggen, N., Leloup, J.-C., and Gonze, D. (2016). A Compact Model for the Complex Plant Circadian Clock. Front. Plant Sci. *7*.

Dixon, L.E., Knox, K., Kozma-Bognar, L., Southern, M.M., Pokhilko, A., and Millar, A.J. (2011). Temporal Repression of Core Circadian Genes Is Mediated through EARLY FLOWERING 3 in Arabidopsis. Curr. Biol. *21*, 120–125.

Doyle, M.R., Davis, S.J., Bastow, R.M., McWatters, H.G., Kozma-Bognár, L., Nagy, F., Millar, A.J., and Amasino, R.M. (2002). The ELF4 gene controls circadian rhythms and flowering time in Arabidopsis thaliana. Nature *419*, 74–77.

Dubos, C., Stracke, R., Grotewold, E., Weisshaar, B., Martin, C., and Lepiniec, L. (2010). MYB transcription factors in Arabidopsis. Trends Plant Sci. *15*, 573–581.

Dubowy, C., and Sehgal, A. (2017). Circadian Rhythms and Sleep in Drosophila melanogaster. Genetics *205*, 1373–1397.

Dunlap, J.C. (1999). Molecular bases for circadian clocks. Cell *96*, 271–290.

Elowitz, M.B., and Leibler, S. (2000). A synthetic oscillatory network of transcriptional regulators. Nature *403*, 335–338.

Emery, P., So, W.V., Kaneko, M., Hall, J.C., and Rosbash, M. (1998). CRY, a Drosophila clock and light-regulated cryptochrome, is a major contributor to circadian rhythm resetting and photosensitivity. Cell *95*, 669–679.

Ezer, D., Jung, J.-H., Lan, H., Biswas, S., Gregoire, L., Box, M.S., Charoensawan, V., Cortijo, S., Lai, X., Stöckle, D., et al. (2017a). The Evening Complex coordinates environmental and endogenous signals in Arabidopsis. Nat. Plants *3*, 17087.

Ezer, D., Shepherd, S.J.K., Brestovitsky, A., Dickinson, P., Cortijo, S., Charoensawan, V., Box, M.S., Biswas, S., Jaeger, K.E., and Wigge, P.A. (2017b). The G-Box Transcriptional Regulatory Code in Arabidopsis. Plant Physiol. *175*, 628–640.

Fankhauser, C., and Staiger, D. (2002). Photoreceptors in Arabidopsis thaliana: light perception, signal transduction and entrainment of the endogenous clock. Planta *216*, 1–16.

Farré, E.M., Harmer, S.L., Harmon, F.G., Yanovsky, M.J., and Kay, S.A. (2005). Overlapping and distinct roles of PRR7 and PRR9 in the Arabidopsis circadian clock. Curr. Biol. CB *15*, 47–54.

Field, M.D., Maywood, E.S., O'Brien, J.A., Weaver, D.R., Reppert, S.M., and Hastings, M.H. (2000). Analysis of clock proteins in mouse SCN demonstrates phylogenetic divergence of the circadian clockwork and resetting mechanisms. Neuron *25*, 437–447.

Fowler, S., Lee, K., Onouchi, H., Samach, A., Richardson, K., Morris, B., Coupland, G., and Putterill, J. (1999). GIGANTEA: a circadian clock-controlled gene that regulates photoperiodic flowering in Arabidopsis and encodes a protein with several possible membrane-spanning domains. EMBO J. *18*, 4679–4688.

Gangappa, S.N., Berriri, S., and Kumar, S.V. (2017). PIF4 Coordinates Thermosensory Growth and Immunity in Arabidopsis. Curr. Biol. *27*, 243–249.

Gekakis, N., Saez, L., Delahaye-Brown, A.M., Myers, M.P., Sehgal, A., Young, M.W., and Weitz, C.J. (1995). Isolation of timeless by PER protein interaction: defective interaction between timeless protein and long-period mutant PERL. Science *270*, 811–815.

Gekakis, N., Staknis, D., Nguyen, H.B., Davis, F.C., Wilsbacher, L.D., King, D.P., Takahashi, J.S., and Weitz, C.J. (1998). Role of the CLOCK protein in the mammalian circadian mechanism. Science *280*, 1564–1569.

Gendron, J.M., Pruneda-Paz, J.L., Doherty, C.J., Gross, A.M., Kang, S.E., and Kay, S.A. (2012). Arabidopsis circadian clock protein, TOC1, is a DNA-binding transcription factor. Proc. Natl. Acad. Sci. *109*, 3167–3172.

Glossop, N.R.J., Houl, J.H., Zheng, H., Ng, F.S., Dudek, S.M., and Hardin, P.E. (2003). VRILLE feeds back to control circadian transcription of Clock in the Drosophila circadian oscillator. Neuron *37*, 249–261.

Golden, S.S., Ishiura, M., Johnson, C.H., and Kondo, T. (1997). Cyanobacterial Circadian Rhythms. Annu. Rev. Plant Physiol. Plant Mol. Biol. *48*, 327–354.

Golombek, D.A., and Rosenstein, R.E. (2010). Physiology of circadian entrainment. Physiol. Rev. *90*, 1063–1102.

Green, R.M., and Tobin, E.M. (2002). The Role of CCA1 and LHY in the Plant Circadian Clock. Dev. Cell *2*, 516–518.

Griffin, E.A., Staknis, D., and Weitz, C.J. (1999). Light-Independent Role of CRY1 and CRY2 in the Mammalian Circadian Clock. Science *286*, 768–771.

Hall, A., Bastow, R.M., Davis, S.J., Hanano, S., McWatters, H.G., Hibberd, V., Doyle, M.R., Sung, S., Halliday, K.J., Amasino, R.M., et al. (2003). The TIME FOR COFFEE Gene Maintains the Amplitude and Timing of Arabidopsis Circadian Clocks. Plant Cell *15*, 2719–2729.

Hall, L.N., Rossini, L., Cribb, L., and Langdale, J.A. GOLDEN 2: A Novel Transcriptional Regulator of Cellular Differentiation in the Maize Leaf. 13.

Hanson, J., Yang, Y., Paliwal, K., Zhou, Y., and Tramontano, A. (2017). Improving protein disorder prediction by deep bidirectional long short-term memory recurrent neural networks. Bioinformatics *33*, 685–692.

Hardin, P.E., Hall, J.C., and Rosbash, M. (1990). Feedback of the Drosophila period gene product on circadian cycling of its messenger RNA levels. Nature *343*, 536–540.

Harmer, S.L., Hogenesch, J.B., Straume, M., Chang, H.-S., Han, B., Zhu, T., Wang, X., Kreps, J.A., and Kay, S.A. (2000). Orchestrated Transcription of Key Pathways in Arabidopsis by the Circadian Clock. Science *290*, 2110–2113.

Harmer, S.L., Panda, S., and Kay, S.A. (2001). Molecular Bases of Circadian Rhythms. Annu. Rev. Cell Dev. Biol. *17*, 215–253.

Hazen, S.P., Schultz, T.F., Pruneda-Paz, J.L., Borevitz, J.O., Ecker, J.R., and Kay, S.A. (2005). LUX ARRHYTHMO encodes a Myb domain protein essential for circadian rhythms. Proc. Natl. Acad. Sci. *102*, 10387–10392.

Helfer, A., Nusinow, D.A., Chow, B.Y., Gehrke, A.R., Bulyk, M.L., and Kay, S.A. (2011). LUX ARRHYTHMO Encodes a Nighttime Repressor of Circadian Gene Expression in the Arabidopsis Core Clock. Curr. Biol. *21*, 126–133.

Henikoff, S. (1984). Unidirectional digestion with exonuclease III creates targeted breakpoints for DNA sequencing. Gene *28*, 351–359.

Herrero, E., Kolmos, E., Bujdoso, N., Yuan, Y., Wang, M., Berns, M.C., Uhlworm, H., Coupland, G., Saini, R., Jaskolski, M., et al. (2012). EARLY FLOWERING4 Recruitment of EARLY FLOWERING3 in the Nucleus Sustains the Arabidopsis Circadian Clock. Plant Cell *24*, 428–443.

Hicks, K.A., Albertson, T.M., and Wagner, D.R. (2001). EARLY FLOWERING3 Encodes a Novel Protein That Regulates Circadian Clock Function and Flowering in Arabidopsis. Plant Cell *13*, 1281–1292.

Hogenesch, J.B., Gu, Y.Z., Jain, S., and Bradfield, C.A. (1998). The basic-helix-loop-helix-PAS orphan MOP3 forms transcriptionally active complexes with circadian and hypoxia factors. Proc. Natl. Acad. Sci. U. S. A. *95*, 5474–5479.

Hong, S., Song, H.-R., Lutz, K., Kerstetter, R.A., Michael, T.P., and McClung, C.R. (2010). Type II protein arginine methyltransferase 5 (PRMT5) is required for circadian period determination in Arabidopsis thaliana. Proc. Natl. Acad. Sci. *107*, 21211–21216.

Honma, S. (2018). The mammalian circadian system: a hierarchical multi-oscillator structure for generating circadian rhythm. J. Physiol. Sci. *68*, 207–219.

van der Horst, G.T., Muijtjens, M., Kobayashi, K., Takano, R., Kanno, S., Takao, M., de Wit, J., Verkerk, A., Eker, A.P., van Leenen, D., et al. (1999). Mammalian Cry1 and Cry2 are essential for maintenance of circadian rhythms. Nature *398*, 627–630.

Hosoda, K. (2002). Molecular Structure of the GARP Family of Plant Myb-Related DNA Binding Motifs of the Arabidopsis Response Regulators. PLANT CELL ONLINE *14*, 2015–2029.

Houl, J.H., Yu, W., Dudek, S.M., and Hardin, P.E. (2006). Drosophila CLOCK is constitutively expressed in circadian oscillator and non-oscillator cells. J. Biol. Rhythms *21*, 93–103.

Hsu, P.Y., and Harmer, S.L. (2014). Wheels within wheels: the plant circadian system. Trends Plant Sci. *19*, 240–249.

Hsu, P.Y., Devisetty, U.K., and Harmer, S.L. (2013). Accurate timekeeping is controlled by a cycling activator in Arabidopsis. ELife *2*, e00473.

Huang, R.-C. (2018). The discoveries of molecular mechanisms for the circadian rhythm: The 2017 Nobel Prize in Physiology or Medicine. Biomed. J. *41*, 5–8.

Huang, H., Alvarez, S., Bindbeutel, R., Shen, Z., Naldrett, M.J., Evans, B.S., Briggs, S.P., Hicks, L.M., Kay, S.A., and Nusinow, D.A. (2016). Identification of Evening Complex Associated Proteins in Arabidopsis by Affinity Purification and Mass Spectrometry. Mol. Cell. Proteomics *15*, 201–217.

Huang, W., Pérez-García, P., Pokhilko, A., Millar, A.J., Antoshechkin, I., Riechmann, J.L., and Mas, P. (2012). Mapping the Core of the Arabidopsis Circadian Clock Defines the Network Structure of the Oscillator. Science *336*, 75–79.

Hunter-Ensor, M., Ousley, A., and Sehgal, A. (1996). Regulation of the Drosophila protein timeless suggests a mechanism for resetting the circadian clock by light. Cell *84*, 677–685.

Hurley, J.M., Loros, J.J., and Dunlap, J.C. (2016). Circadian Oscillators: Around the Transcription-Translation Feedback Loop and on to Output. Trends Biochem. Sci. *41*, 834–846.

Ibañez, C., Delker, C., Martinez, C., Bürstenbinder, K., Janitza, P., Lippmann, R., Ludwig, W., Sun, H., James, G.V., Klecker, M., et al. (2018). Brassinosteroids Dominate Hormonal Regulation of Plant Thermomorphogenesis via BZR1. Curr. Biol. *28*, 303-310.e3.

Imamura, A., Hanaki, N., Nakamura, A., Suzuki, T., Taniguchi, M., Kiba, T., Ueguchi, C., Sugiyama, T., and Mizuno, T. (1999). Compilation and Characterization of Arabiopsis thaliana Response Regulators Implicated in His-Asp Phosphorelay Signal Transduction. Plant Cell Physiol. *40*, 733–742.

Incardona, M.-F., Bourenkov, G.P., Levik, K., Pieritz, R.A., Popov, A.N., and Svensson, O. (2009). EDNA: a framework for plugin-based applications applied to X-ray experiment online data analysis. J. Synchrotron Radiat. *16*, 872–879.

Inoue, K., Araki, T., and Endo, M. (2018). Circadian clock during plant development. J. Plant Res. *131*, 59–66.

Ishiura, M., Kutsuna, S., Aoki, S., Iwasaki, H., Andersson, C.R., Tanabe, A., Golden, S.S., Johnson, C.H., and Kondo, T. (1998). Expression of a Gene Cluster kaiABC as a Circadian Feedback Process in Cyanobacteria. Science *281*, 1519–1523.

Jung, J.-H., Domijan, M., Klose, C., Biswas, S., Ezer, D., Gao, M., Khattak, A.K., Box, M.S., Charoensawan, V., Cortijo, S., et al. (2016). Phytochromes function as thermosensors in *Arabidopsis*. Science *354*, 886–889.

Kamioka, M., Takao, S., Suzuki, T., Taki, K., Higashiyama, T., Kinoshita, T., and Nakamichi, N. (2016). Direct Repression of Evening Genes by CIRCADIAN CLOCK-ASSOCIATED1 in the Arabidopsis Circadian Clock. Plant Cell *28*, 696–711.

Kanei-Ishii, C., Sarai, A., Sawazaki, T., Nakagoshi, H., He, D.N., Ogata, K., Nishimura, Y., and Ishii, S. (1990). The tryptophan cluster: a hypothetical structure of the DNA-binding domain of the myb protooncogene product. J. Biol. Chem. *265*, 19990–19995.

Käppel, S., Melzer, R., Rümpler, F., Gafert, C., and Theißen, G. (2018). The floral homeotic protein SEPALLATA3 recognizes target DNA sequences by shape readout involving a conserved arginine residue in the MADS-domain. Plant J. Cell Mol. Biol. *95*, 341–357.

Kidd, P.B., Young, M.W., and Siggia, E.D. (2015). Temperature compensation and temperature sensation in the circadian clock. Proc. Natl. Acad. Sci. U. S. A. *112*, E6284-6292.

Kikis, E.A., Khanna, R., and Quail, P.H. (2005). ELF4 is a phytochrome-regulated component of a negative-feedback loop involving the central oscillator components CCA1 and LHY. Plant J. *44*, 300–313.

Kim, J., Kim, Y., Yeom, M., Kim, J.-H., and Nam, H.G. (2008). FIONA1 Is Essential for Regulating Period Length in the Arabidopsis Circadian Clock. Plant Cell *20*, 307–319.

Kim, W.-Y., Fujiwara, S., Suh, S.-S., Kim, J., Kim, Y., Han, L., David, K., Putterill, J., Nam, H.G., and Somers, D.E. (2007). ZEITLUPE is a circadian photoreceptor stabilized by GIGANTEA in blue light. Nature *449*, 356–360.

Kim, Y., Yeom, M., Kim, H., Lim, J., Koo, H.J., Hwang, D., Somers, D., and Nam, H.G. (2012). GIGANTEA and EARLY FLOWERING 4 in Arabidopsis Exhibit Differential Phase-Specific Genetic Influences over a Diurnal Cycle. Mol. Plant *5*, 152–161.

Kim, Y., Lim, J., Yeom, M., Kim, H., Kim, J., Wang, L., Kim, W.Y., Somers, D.E., and Nam, H.G. (2013). ELF4 Regulates GIGANTEA Chromatin Access through Subnuclear Sequestration. Cell Rep. *3*, 671–677.

Kloppstech, K. (1985). Diurnal and circadian rhythmicity in the expression of light-induced plant nuclear messenger RNAs. Planta *165*, 502–506.

Koini, M.A., Alvey, L., Allen, T., Tilley, C.A., Harberd, N.P., Whitelam, G.C., and Franklin, K.A. (2009). High Temperature-Mediated Adaptations in Plant Architecture Require the bHLH Transcription Factor PIF4. Curr. Biol. *19*, 408–413.

Kolmos, E., Nowak, M., Werner, M., Fischer, K., Schwarz, G., Mathews, S., Schoof, H., Nagy, F., Bujnicki, J.M., and Davis, S.J. (2009). Integrating ELF4 into the circadian system through combined structural and functional studies. HFSP J. *3*, 350–366.

Kondo, T., Tsinoremas, N.F., Golden, S.S., Johnson, C.H., Kutsuna, S., and Ishiura, M. (1994). Circadian clock mutants of cyanobacteria. Science *266*, 1233–1236.

Konopka, R.J., and Benzer, S. (1971). Clock mutants of Drosophila melanogaster. Proc. Natl. Acad. Sci. U. S. A. *68*, 2112–2116.

Kumar, S.V., Lucyshyn, D., Jaeger, K.E., Alós, E., Alvey, E., Harberd, N.P., and Wigge, P.A. (2012). Transcription factor PIF4 controls the thermosensory activation of flowering. Nature *484*, 242–245.

Kume, K., Zylka, M.J., Sriram, S., Shearman, L.P., Weaver, D.R., Jin, X., Maywood, E.S., Hastings, M.H., and Reppert, S.M. (1999). mCRY1 and mCRY2 Are Essential Components of the Negative Limb of the Circadian Clock Feedback Loop. Cell *98*, 193–205.

Lee, C., Bae, K., and Edery, I. (1998). The Drosophila CLOCK protein undergoes daily rhythms in abundance, phosphorylation, and interactions with the PER-TIM complex. Neuron *21*, 857–867.

Lee, C., Bae, K., and Edery, I. (1999). PER and TIM inhibit the DNA binding activity of a Drosophila CLOCK-CYC/dBMAL1 heterodimer without disrupting formation of the heterodimer: a basis for circadian transcription. Mol. Cell. Biol. *19*, 5316–5325.

Legris, M., Klose, C., Burgie, E.S., Rojas, C.C.R., Neme, M., Hiltbrunner, A., Wigge, P.A., Schäfer, E., Vierstra, R.D., and Casal, J.J. (2016). Phytochrome B integrates light and temperature signals in. Science *354*, 897–900.

Legris, M., Klose, C., Burgie, E.S., Rojas, C.C., Neme, M., Hiltbrunner, A., Wigge, P.A., Schäfer, E., Vierstra, R.D., and Casal, J.J. Phytochrome B integrates light and temperature signals in Arabidopsis. 5.

Liu, A.C., Welsh, D.K., Ko, C.H., Tran, H.G., Zhang, E.E., Priest, A.A., Buhr, E.D., Singer, O., Meeker, K., Verma, I.M., et al. (2007). Intercellular coupling confers robustness against mutations in the SCN circadian clock network. Cell *129*, 605–616.

Liu, X.L., Covington, M.F., Fankhauser, C., Chory, J., and Wagner, D.R. (2001). ELF3 Encodes a Circadian Clock–Regulated Nuclear Protein That Functions in an Arabidopsis PHYB Signal Transduction Pathway. Plant Cell *13*, 1293–1304.

Locke, J.C.W., Kozma-Bognár, L., Gould, P.D., Fehér, B., Kevei, É., Nagy, F., Turner, M.S., Hall, A., and Millar, A.J. (2006). Experimental validation of a predicted feedback loop in the multi-oscillator clock of Arabidopsis thaliana. Mol. Syst. Biol. *2*, 59.

Lopez-Molina, L., Conquet, F., Dubois-Dauphin, M., and Schibler, U. (1997). The DBP gene is expressed according to a circadian rhythm in the suprachiasmatic nucleus and influences circadian behavior. EMBO J. *16*, 6762–6771.

Lorrain, S., Trevisan, M., Pradervand, S., and Fankhauser, C. (2009). Phytochrome interacting factors 4 and 5 redundantly limit seedling de-etiolation in continuous far-red light. Plant J. *60*, 449–461.

de Mairan, J.-J. (1729). Observation botanique.

Makino, S., Kiba, T., Imamura, A., Hanaki, N., Nakamura, A., Suzuki, T., Taniguchi, M., Ueguchi, C., Sugiyama, T., and Mizuno, T. (2000). Genes encoding pseudo-response regulators: insight into His-to-Asp phosphorelay and circadian rhythm in Arabidopsis thaliana. Plant Cell Physiol. *41*, 791–803.

Más, P., Kim, W.-Y., Somers, D.E., and Kay, S.A. (2003). Targeted degradation of TOC1 by ZTL modulates circadian function in Arabidopsis thaliana. Nature *426*, 567–570.

Matsushika, A., Makino, S., Kojima, M., and Mizuno, T. (2000). Circadian waves of expression of the APRR1/TOC1 family of pseudo-response regulators in Arabidopsis thaliana: insight into the plant circadian clock. Plant Cell Physiol. *41*, 1002–1012.

McClung, C.R., and Kay, S.A. (1994). Circadian rhythms in Arabidopsis thaliana. Cold Spring Harb. Monogr. Ser. *27*, 615–615.

McWatters, H.G., Kolmos, E., Hall, A., Doyle, M.R., Amasino, R.M., Gyula, P., Nagy, F., Millar, A.J., and Davis, S.J. (2007). ELF4 Is Required for Oscillatory Properties of the Circadian Clock. Plant Physiol. *144*, 391–401.

Michael, T.P., Mockler, T.C., Breton, G., McEntee, C., Byer, A., Trout, J.D., Hazen, S.P., Shen, R., Priest, H.D., Sullivan, C.M., et al. (2008). Network Discovery Pipeline Elucidates Conserved Time-of-Day–Specific cis-Regulatory Modules. PLOS Genet. *4*, e14.

Millar, A.J., and Kay, S.A. (1991). Circadian control of cab gene transcription and mRNA accumulation in Arabidopsis. Plant Cell *3*, 541–550.

Millar, A., Carre, I., Strayer, C., Chua, N., and Kay, S. (1995). Circadian clock mutants in Arabidopsis identified by luciferase imaging. Science *267*, 1161–1163.

Millar, A.J., Short, S.R., Chua, N.-H., and Kay, S.A. (1992). A novel circadian phenotype based on firefly luciferase expression in transgenic plants. Plant Cell *4*, 1075–1087.

Mitsui, A., Kumazawa, S., Takahashi, A., Ikemoto, H., Cao, S., and Arai, T. (1986). Strategy by which nitrogen-fixing unicellular cyanobacteria grow photoautotrophically. Nature *323*, 720–722.

Mizoguchi, T., Wheatley, K., Hanzawa, Y., Wright, L., Mizoguchi, M., Song, H.-R., Carré, I.A., and Coupland, G. (2002). LHY and CCA1 Are Partially Redundant Genes Required to Maintain Circadian Rhythms in Arabidopsis. Dev. Cell *2*, 629–641.

Mizuno, T., Nomoto, Y., Oka, H., Kitayama, M., Takeuchi, A., Tsubouchi, M., and Yamashino, T. (2014). Ambient Temperature Signal Feeds into the Circadian Clock Transcriptional Circuitry Through the EC Night-Time Repressor in Arabidopsis thaliana. Plant Cell Physiol. *55*, 958–976.

Moore, R.Y. (1997). Circadian rhythms: basic neurobiology and clinical applications. Annu. Rev. Med. *48*, 253–266.

Moore, F.C., and Lobell, D.B. (2015). The fingerprint of climate trends on European crop yields. Proc. Natl. Acad. Sci. *112*, 2670–2675.

Myers, M.P., Wager-Smith, K., Rothenfluh-Hilfiker, A., and Young, M.W. (1996). Light-induced degradation of TIMELESS and entrainment of the Drosophila circadian clock. Science *271*, 1736–1740.

Nagel, D.H., Doherty, C.J., Pruneda-Paz, J.L., Schmitz, R.J., Ecker, J.R., and Kay, S.A. (2015). Genome-wide identification of CCA1 targets uncovers an expanded clock network in Arabidopsis. Proc. Natl. Acad. Sci. *112*, E4802–E4810.

Nagy, F., Kay, S.A., and Chua, N.-H. (1988). A circadian clock regulates transcription of the wheat Cab-1 gene. Genes Dev *2*, 376–382.

Nakajima, M., Imai, K., Ito, H., Nishiwaki, T., Murayama, Y., Iwasaki, H., Oyama, T., and Kondo, T. (2005). Reconstitution of circadian oscillation of cyanobacterial KaiC phosphorylation in vitro. Science *308*, 414–415.

Nakamichi, N., Kita, M., Niinuma, K., Ito, S., Yamashino, T., Mizoguchi, T., and Mizuno, T. (2007). Arabidopsis clock-associated pseudo-response regulators PRR9, PRR7 and PRR5 coordinately and

positively regulate flowering time through the canonical CONSTANS-dependent photoperiodic pathway. Plant Cell Physiol. *48*, 822–832.

Nakamichi, N., Kiba, T., Henriques, R., Mizuno, T., Chua, N.-H., and Sakakibara, H. (2010). PSEUDO-RESPONSE REGULATORS 9, 7, and 5 are transcriptional repressors in the Arabidopsis circadian clock. Plant Cell *22*, 594–605.

Nieto, C., López-Salmerón, V., Davière, J.-M., and Prat, S. (2015). ELF3-PIF4 Interaction Regulates Plant Growth Independently of the Evening Complex. Curr. Biol. *25*, 187–193.

Nozue, K., Covington, M.F., Duek, P.D., Lorrain, S., Fankhauser, C., Harmer, S.L., and Maloof, J.N. (2007). Rhythmic growth explained by coincidence between internal and external cues. Nature *448*, 358–361.

Nusinow, D.A., Helfer, A., Hamilton, E.E., King, J.J., Imaizumi, T., Schultz, T.F., Farré, E.M., and Kay, S.A. (2011). The ELF4–ELF3–LUX complex links the circadian clock to diurnal control of hypocotyl growth. Nature *475*, 398–402.

Oakenfull, R.J., and Davis, S.J. (2017). Shining a light on the Arabidopsis circadian clock. Plant Cell Environ. *40*, 2571–2585.

Okamura, H., Miyake, S., Sumi, Y., Yamaguchi, S., Yasui, A., Muijtjens, M., Hoeijmakers, J.H., and van der Horst, G.T. (1999). Photic induction of mPer1 and mPer2 in cry-deficient mice lacking a biological clock. Science *286*, 2531–2534.

Onai, K., and Ishiura, M. (2005). PHYTOCLOCK 1 encoding a novel GARP protein essential for the Arabidopsis circadian clock. Genes Cells *10*, 963–972.

Ostermeier, M., and Lutz, S. (2003). The creation of ITCHY hybrid protein libraries. Methods Mol Biol *231*, 129–141.

Ouyang, Y., Andersson, C.R., Kondo, T., Golden, S.S., and Johnson, C.H. (1998). Resonating circadian clocks enhance fitness in cyanobacteria. Proc. Natl. Acad. Sci. *95*, 8660–8664.

Panda, S., Poirier, G.G., and Kay, S.A. (2002). tej Defines a Role for Poly(ADP-Ribosyl)ation in Establishing Period Length of the Arabidopsis Circadian Oscillator. Dev. Cell *3*, 51–61.

Para, A., Farré, E.M., Imaizumi, T., Pruneda-Paz, J.L., Harmon, F.G., and Kay, S.A. (2007). PRR3 Is a Vascular Regulator of TOC1 Stability in the Arabidopsis Circadian Clock. Plant Cell *19*, 3462–3473.

Partch, C.L., Green, C.B., and Takahashi, J.S. (2014). Molecular Architecture of the Mammalian Circadian Clock. Trends Cell Biol. *24*, 90–99.

Peng, S., Huang, J., Sheehy, J.E., Laza, R.C., Visperas, R.M., Zhong, X., Centeno, G.S., Khush, G.S., and Cassman, K.G. (2004). Rice yields decline with higher night temperature from global warming. Proc. Natl. Acad. Sci. *101*, 9971–9975.

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

Pokhilko, A., Hodge, S.K., Stratford, K., Knox, K., Edwards, K.D., Thomson, A.W., Mizuno, T., and Millar, A.J. (2010). Data assimilation constrains new connections and components in a complex, eukaryotic circadian clock model. Mol. Syst. Biol. *6*, 416.

Pokhilko, A., Fernández, A.P., Edwards, K.D., Southern, M.M., Halliday, K.J., and Millar, A.J. (2012). The clock gene circuit in Arabidopsis includes a repressilator with additional feedback loops. Mol. Syst. Biol. *8*, 574.

Preitner, N., Damiola, F., Lopez-Molina, L., Zakany, J., Duboule, D., Albrecht, U., and Schibler, U. (2002). The orphan nuclear receptor REV-ERBalpha controls circadian transcription within the positive limb of the mammalian circadian oscillator. Cell *110*, 251–260.

Pruneda-Paz, J.L., Breton, G., Para, A., and Kay, S.A. (2009). A Functional Genomics Approach Reveals CHE as a Component of the Arabidopsis Circadian Clock. Science *323*, 1481–1485.

Quint, M., Delker, C., Franklin, K.A., Wigge, P.A., Halliday, K.J., and van Zanten, M. (2016). Molecular and genetic control of plant thermomorphogenesis. Nat. Plants *2*, 15190.

Raschke, A., Ibañez, C., Ullrich, K.K., Anwer, M.U., Becker, S., Glöckner, A., Trenner, J., Denk, K., Saal, B., Sun, X., et al. (2015). Natural variants of ELF3 affect thermomorphogenesis by transcriptionally modulating PIF4-dependent auxin response genes. BMC Plant Biol. *15*, 197.

Reddy, P., Zehring, W.A., Wheeler, D.A., Pirrotta, V., Hadfield, C., Hall, J.C., and Rosbash, M. (1984). Molecular analysis of the period locus in Drosophila melanogaster and identification of a transcript involved in biological rhythms. Cell *38*, 701–710.

Reppert, S.M., and Weaver, D.R. (2001). Molecular analysis of mammalian circadian rhythms. Annu. Rev. Physiol. *63*, 647–676.

Richards, E., Reichardt, M., and Rogers, S. (1994). Preparation of Genomic DNA from Plant Tissue. Curr. Protoc. Mol. Biol. *27*, 2.3.1-2.3.7.

Riechmann, J.L., and Ratcliffe, O.J. (2000). A genomic perspective on plant transcription factors. Curr. Opin. Plant Biol. *3*, 423–434.

Riechmann, J.L., Heard, J., Martin, G., Reuber, L., Jiang, C.-Z., Keddie, J., Adam, L., Pineda, O., Ratcliffe, O.J., Samaha, R.R., et al. (2000). Arabidopsis Transcription Factors: Genome-Wide Comparative Analysis Among Eukaryotes. Science *290*, 2105–2110.

Robinson, I., and Reddy, A.B. (2014). Molecular mechanisms of the circadian clockwork in mammals. FEBS Lett. *588*, 2477–2483.

Roenneberg, T., and Merrow, M. (1998). Molecular circadian oscillators: an alternative hypothesis. J. Biol. Rhythms *13*, 167–179.

Rost, B., Yachdav, G., and Liu, J. (2004). The PredictProtein server. Nucleic Acids Res. *32*, W321–W326.

Rugnone, M.L., Soverna, A.F., Sanchez, S.E., Schlaen, R.G., Hernando, C.E., Seymour, D.K., Mancini, E., Chernomoretz, A., Weigel, D., Más, P., et al. (2013). LNK genes integrate light and clock signaling networks at the core of the Arabidopsis oscillator. Proc. Natl. Acad. Sci. *110*, 12120–12125.

Rutila, J.E., Suri, V., Le, M., So, W.V., Rosbash, M., and Hall, J.C. (1998). CYCLE is a second bHLH-PAS clock protein essential for circadian rhythmicity and transcription of Drosophila period and timeless. Cell *93*, 805–814.

Sathyanarayanan, S., Zheng, X., Xiao, R., and Sehgal, A. (2004). Posttranslational Regulation of Drosophila PERIOD Protein by Protein Phosphatase 2A. Cell *116*, 603–615.

Schaffer, R., Ramsay, N., Samach, A., Corden, S., Putterill, J., Carré, I.A., and Coupland, G. (1998). The late elongated hypocotyl mutation of Arabidopsis disrupts circadian rhythms and the photoperiodic control of flowering. Cell *93*, 1219–1229.

Sehgal, A., Rothenfluh-Hilfiker, A., Hunter-Ensor, M., Chen, Y., Myers, M.P., and Young, M.W. (1995). Rhythmic expression of timeless: a basis for promoting circadian cycles in period gene autoregulation. Science *270*, 808–810.

Shalit-Kaneh, A., Kumimoto, R.W., Filkov, V., and Harmer, S.L. (2018). Multiple feedback loops of the Arabidopsis circadian clock provide rhythmic robustness across environmental conditions. Proc. Natl. Acad. Sci. *115*, 7147–7152.

Shearman, L.P., Sriram, S., Weaver, D.R., Maywood, E.S., Chaves, I., Zheng, B., Kume, K., Lee, C.C., van der Horst, G.T., Hastings, M.H., et al. (2000). Interacting molecular loops in the mammalian circadian clock. Science *288*, 1013–1019.

Shen, W.J., and Forde, B.G. (1989). Efficient transformation of Agrobacterium spp. by high voltage electroporation. Nucleic Acids Res. *17*, 8385.

Silva, C.S., Lai, X., Nanao, M., and Zubieta, C. (2016). The Myb domain of LUX ARRHYTHMO in complex with DNA: expression, purification and crystallization. Acta Crystallogr. Sect. F *72*, 356–361.

Singh, A.P., Fridman, Y., Friedlander-Shani, L., Tarkowska, D., Strnad, M., and Savaldi-Goldstein, S. (2014). Activity of the Brassinosteroid Transcription Factors BRASSINAZOLE RESISTANT1 and BRASSINOSTEROID INSENSITIVE1-ETHYL METHANESULFONATE-SUPPRESSOR1/BRASSINAZOLE RESISTANT2 Blocks Developmental Reprogramming in Response to Low Phosphate Availability. Plant Physiol. *166*, 678–688.

Somers, D.E., Devlin, P.F., and Kay, S.A. (1998). Phytochromes and cryptochromes in the entrainment of the Arabidopsis circadian clock. Science *282*, 1488–1490.

Somers, D.E., Schultz, T.F., Milnamow, M., and Kay, S.A. (2000). ZEITLUPE Encodes a Novel Clock-Associated PAS Protein from Arabidopsis. Cell *101*, 319–329.

Stokkan, K.-A., Yamazaki, S., Tei, H., Sakaki, Y., and Menaker, M. (2001). Entrainment of the circadian clock in the liver by feeding. Science *291*, 490–493.

Strayer, C., Oyama, T., Schultz, T.F., Raman, R., Somers, D.E., Más, P., Panda, S., Kreps, J.A., and Kay, S.A. (2000). Cloning of the Arabidopsis clock gene TOC1, an autoregulatory response regulator homolog. Science *289*, 768–771.

Swan, J.A., Golden, S., LiWang, A., and Partch, C.L. (2018). Structure, function, and mechanism of the core circadian clock in cyanobacteria. J. Biol. Chem. jbc.TM117.001433.

Tarendeau, F., Boudet, J., Guilligay, D., Mas, P.J., Bougault, C.M., Boulo, S., Baudin, F., Ruigrok, R.W.H., Daigle, N., Ellenberg, J., et al. (2007). Structure and nuclear import function of the C-terminal domain of influenza virus polymerase PB2 subunit. Nat. Struct. Mol. Biol. *14*, 229–233.

Thines, B.C., Youn, Y., Duarte, M.I., and Harmon, F.G. (2014). The time of day effects of warm temperature on flowering time involve PIF4 and PIF5. J. Exp. Bot. *65*, 1141–1151.

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

Trowitzsch, S., Bieniossek, C., Nie, Y., Garzoni, F., and Berger, I. (2010). New baculovirus expression tools for recombinant protein complex production. J. Struct. Biol. *172*, 45–54.

Ventrella, D., Charfeddine, M., Moriondo, M., Rinaldi, M., and Bindi, M. (2012). Agronomic adaptation strategies under climate change for winter durum wheat and tomato in southern Italy: irrigation and nitrogen fertilization. Reg. Environ. Change *12*, 407–419.

Vijayachandran, L.S., Viola, C., Garzoni, F., Trowitzsch, S., Bieniossek, C., Chaillet, M., Schaffitzel, C., Busso, D., Romier, C., Poterszman, A., et al. (2011). Robots, pipelines, polyproteins: Enabling multiprotein expression in prokaryotic and eukaryotic cells. J. Struct. Biol. *175*, 198–208.

Vitaterna, M.H., Selby, C.P., Todo, T., Niwa, H., Thompson, C., Fruechte, E.M., Hitomi, K., Thresher, R.J., Ishikawa, T., Miyazaki, J., et al. (1999). Differential regulation of mammalian Period genes and circadian rhythmicity by cryptochromes 1 and 2. Proc. Natl. Acad. Sci. U. S. A. *96*, 12114–12119.

Wang, Z.Y., and Tobin, E.M. (1998). Constitutive expression of the CIRCADIAN CLOCK ASSOCIATED 1 (CCA1) gene disrupts circadian rhythms and suppresses its own expression. Cell *93*, 1207–1217.

Wang, L., Kim, J., and Somers, D.E. (2013). Transcriptional corepressor TOPLESS complexes with pseudoresponse regulator proteins and histone deacetylases to regulate circadian transcription. Proc. Natl. Acad. Sci. U. S. A. *110*, 761–766.

Welsh, D.K., Imaizumi, T., and Kay, S.A. (2005). Real-time reporting of circadian-regulated gene expression by luciferase imaging in plants and mammalian cells. In Methods in Enzymology, (Elsevier), pp. 269–288.

Wenkel, S., Turck, F., Singer, K., Gissot, L., Le Gourrierec, J., Samach, A., and Coupland, G. (2006). CONSTANS and the CCAAT Box Binding Complex Share a Functionally Important Domain and Interact to Regulate Flowering of Arabidopsis. PLANT CELL ONLINE *18*, 2971–2984.

Wykoff, D.D., Grossman, A.R., Weeks, D.P., Usuda, H., and Shimogawara, K. (1999). Psr1, a nuclear localized protein that regulates phosphorus metabolism in Chlamydomonas. Proc. Natl. Acad. Sci. *96*, 15336–15341.

Yeom, M., Kim, H., Lim, J., Shin, A.-Y., Hong, S., Kim, J.-I., and Nam, H.G. (2014). How Do Phytochromes Transmit the Light Quality Information to the Circadian Clock in Arabidopsis ? Mol. Plant *7*, 1701–1704.

Yin, R., Arongaus, A.B., Binkert, M., and Ulm, R. (2015). Two Distinct Domains of the UVR8 Photoreceptor Interact with COP1 to Initiate UV-B Signaling in Arabidopsis. Plant Cell *27*, 202–213.

Yoshida, R., Fekih, R., Fujiwara, S., Oda, A., Miyata, K., Tomozoe, Y., Nakagawa, M., Niinuma, K., Hayashi, K., Ezura, H., et al. (2009). Possible role of EARLY FLOWERING 3 (ELF3) in clock-dependent floral regulation by SHORT VEGETATIVE PHASE (SVP) in Arabidopsis thaliana. New Phytol. *182*, 838–850.

Yu, J.-W., Rubio, V., Lee, N.-Y., Bai, S., Lee, S.-Y., Kim, S.-S., Liu, L., Zhang, Y., Irigoyen, M.L., Sullivan, J.A., et al. (2008). COP1 and ELF3 Control Circadian Function and Photoperiodic Flowering by Regulating GI Stability. Mol. Cell *32*, 617–630.

Yumerefendi, H., Tarendeau, F., Mas, P.J., and Hart, D.J. (2010). ESPRIT: An automated, library-based method for mapping and soluble expression of protein domains from challenging targets. J. Struct. Biol. *172*, 66–74.

Zagotta, M.T., Hicks, K.A., Jacobs, C.I., Young, J.C., Hangarter, R.P., and Meeks-Wagner, D.R. (1996). The Arabidopsis ELF3 gene regulates vegetative photomorphogenesis and the photoperiodic induction of flowering. Plant J. *10*, 691–702.

Zaitseva, J., Holland, I.B., and Schmitt, L. (2004). The role of CAPS buffer in expanding the crystallization space of the nucleotide-binding domain of the ABC transporter haemolysin B from *Escherichia coli*. Acta Crystallogr. D Biol. Crystallogr. *60*, 1076–1084.

Zeng, H., Qian, Z., Myers, M.P., and Rosbash, M. (1996). A light-entrainment mechanism for the Drosophila circadian clock. Nature *380*, 129–135.

Zubieta, C., Silva, C.S., Lai, X., Wigge, P., Nanao, M.H., and Nayak, A. Structure of the DNA-binding domain of LUX ARRHYTHMO. BE Publ.