Design and Synthesis of Circadian Clock Modulators and The Study of LOV Domain Protein LKP2 in Arabidosis thaliana and Brassica rapa

Aditi Nagar
Southern Methodist University, aditin@smu.edu

Follow this and additional works at: https://scholar.smu.edu/hum_sci_chemistry_etds

Part of the Biochemistry, Biophysics, and Structural Biology Commons, and the Organic Chemistry Commons

Recommended Citation
Nagar, Aditi, "Design and Synthesis of Circadian Clock Modulators and The Study of LOV Domain Protein LKP2 in Arabidosis thaliana and Brassica rapa" (2018). Chemistry Theses and Dissertations. 7. https://scholar.smu.edu/hum_sci_chemistry_etds/7

This Dissertation is brought to you for free and open access by the Chemistry at SMU Scholar. It has been accepted for inclusion in Chemistry Theses and Dissertations by an authorized administrator of SMU Scholar. For more information, please visit http://digitalrepository.smu.edu.
DESIGN AND SYNTHESIS OF CIRCADIAN CLOCK MODULATORS AND STUDY OF LOV DOMAIN PROTEIN LKP2 IN *ARABIDOPSIS THALIANA* AND *BRASSICA RAPA*
DESIGN AND SYNTHESIS OF CIRCADIAN CLOCK MODULATORS AND STUDY OF LOV DOMAIN PROTEIN LKP2 IN ARABIDOPSIS THALIANA AND BRASSICA RAPA

A Dissertation Presented to the Graduate Faculty of

Dedman College

Southern Methodist University

in

Partial Fulfillment of the Requirements

for the degree of

Doctor of Philosophy

with a

Major in Chemistry

by

Aditi Nagar

B.S., Biology, Chemistry, University of Lucknow, Lucknow, India
M.S., Chemistry, University of Houston-Clear Lake, Houston

August 07, 2018
ACKNOWLEDGMENTS

In the past five years, innumerable people have directly and indirectly contributed towards my Ph.D. First and foremost, Dr. Brian D. Zoltowski my Ph.D. advisor without whose support and guidance I would not have reached to this stage. I always fall short of words to express my gratitude towards him. He is not only a great scientist but a nice human being and is a perfect role model. I feel truly blessed to have gotten a chance to work with him, I have learned so much just by observing him in the laboratory or the way he teaches in the class. His confidence has always inspired and motivated me to take up any challenge with my full potential. I have not only grown as a scientist but as a better professional. I thank Dr. Z, for being the best mentor and a Ph.D. advisor.

I thank my lab mates Dr. Jameela Lokhandwala and Dr. Ashutosh Pudasaini for their companionship and for making a conducive lab environment. I thank them for being patient while teaching me biochemistry lab techniques. I thank my peers Dr. Jian Cao and Rajesh Kumar, for being the best classmates, I will never be able to forget the camaraderie we shared.

I would like to extend my heartfelt thanks to Dr. Edward R. Biehl, Dr. John D. Buynak, Dr. Alex R. Lippert and Dr. Steven B. Vik, and Dr. Michael Lattman for agreeing to be in my graduate committee. I thank Dr. Biehl for allowing me to work in his lab and for being extremely
patient while teaching me complex concepts of benzyne chemistry. My heartfelt thanks to Dr. Buynak and Dr. Lippert, for always taking out time for answering my questions and for giving me good advice to carry out the organic reactions.

I thank Dr. Sukanta Kamila for getting me started in the organic lab and teaching me different analytical techniques. I thank Dr. Alan Humason for being approachable and always willing to help whenever needed.

I thank Dr. Patty Wisian-Neilson, Dr. Jennifer O’Brien and Mrs. Andrea Adams for the teaching assistant training provided by them, the guidance will help me in my chosen career. I would like to extend my heartfelt thanks to Dr. Patty for writing my recommendation letter.

I deeply thank Robert Green, Priya Patel, Tyler Rozanitis, Brandon Le, Shannon Swisher, Sydney Schmitt for their companionship and the support in the lab. I thank Samuel Weber for teaching cell culture techniques and for always getting cells ready before every drug screening experiment.

On the personal level, I thank my parents for being the constant source of encouragement, whatever am I today is all because of them. I thank my grandmother for her unconditional love, she will always be in my memory. I thank my sister Ankita for fun at the same time motivational phone calls. Last but certainly not the least, I thank my husband, Anurag for being a strong pillar of support and for encouraging me to follow my passion for getting a Ph.D. I would not be able to have reached this far without his support. Also, I thank him for analyzing all the drug screening data without any complaints.
Circadian rhythms are self-autonomous endogenous clocks synchronized with the rotation of the Earth. With the Earth’s rotation and revolution on its axis, the internal clock undergoes oscillation in the period of ~24 hour and governs day to day activities in most organisms. In humans, it regulates the day to day physiological activities. Today’s modern lifestyle has an impact on health: shift work, jet lag, and irregular eating habits contribute to the misalignment of the endogenous circadian oscillator, thereby, increasing the risk of many metabolic disorders including diabetes, irregular blood pressure, sleep disorders, obesity, depression, and cancer. The primary goal of this research is to design and synthesize therapeutic circadian clock modulators to target clock-related dysfunctions.

This dissertation focusses on designing a library of small molecules which can act on core circadian rhythm proteins. More specifically, the goal is to synthesize different analogs to target cryptochrome. Cryptochrome is an integral part of the mammalian core circadian system and regulates the process of gluconeogenesis, therefore it has been identified as a potential target for type II diabetes. A carbazole scaffold (KL001) was shown to bind at the active site of cryptochrome resulting in a shift in circadian period, and modulation of gluconeogenesis. Herein,
a library of the small molecules is synthesized and evaluated based on the structure-activity relationship of KL001.

Just like mammals, most living organisms have developed an endogenous circadian clock. In the latter part of the dissertation, the focus shifts to study of the crucial circadian rhythm proteins of plants. In plants, Light Oxygen Voltage (LOV) domain proteins play an essential role in photoperiodic growth events. LOV domains employ small molecule chemistry to sense blue light in the environment and facilitate plants acclimating their surroundings. In *Arabidopsis*, the Zeitlupe (ZTL), flavin binding kelch repeat F-box1 (FKF1) and LOV Kelch protein-2 (LKP2) work in concert to measure the day length and to determine flowering time and regulate other circadian processes. Specifically, the focus will be to study the thermal kinetics of LKP2 protein of *Arabidopsis thaliana*. Furthermore, this study is extended to examine the thermal kinetics of LKP2 protein in an agriculturally important crop *Brassica rapa*. Finally, a crystal structure of the cryptochrome-like algal protein OtCPF1 is discussed.
TABLE OF CONTENTS

LIST OF FIGURES .................................................................................................................. xiv

LIST OF TABLES ....................................................................................................................... xxiii

LIST OF SCHEMES ................................................................................................................... xxiv

LIST OF ABBREVIATIONS ........................................................................................................ xxv

LIST OF AMINO ACIDS ........................................................................................................... xxix

LIST OF THE ZOLTOWSKI LIBRARY ....................................................................................... xxxi

CHAPTER 1 ................................................................................................................................. 1

CIRCADIAN RHYTHM: NEW AVENUES FOR THE DRUG DISCOVERY ......................... 1

1.1 Introduction ......................................................................................................................... 1

1.2 Characteristic and Molecular basis of the circadian clock ............................................. 3

1.2.1 The feedback loop ......................................................................................................... 4

1.2.2 Important clock parameters .......................................................................................... 6

1.3 Discovery of circadian rhythm to chronotherapy: Historical perspective .................. 7

1.4 Mammalian circadian rhythm .......................................................................................... 8

1.5 Molecular architecture of the mammalian circadian clock .......................................... 11

1.6 Structures of the core circadian rhythm proteins ......................................................... 14

1.7 Resetting the clock: Chemical chronotherapy .............................................................. 16

1.7.1 Circadian rhythm and drug intervention ..................................................................... 17

1.7.2 Identification of the synthetic clock modulators ......................................................... 18
1.8 Cryptochrome as a target of drug discovery ................................................................. 23
  1.8.1 Cryptochrome ............................................................................................................ 23
  1.8.2 Cryptochrome regulation of gluconeogenesis ......................................................... 25
  1.8.3 Gluconeogenesis and Diabetes ................................................................................ 27
1.9 Cryptochrome modulator ................................................................................................ 28
  1.9.1 KL001 ....................................................................................................................... 29
1.10 Research Objective ...................................................................................................... 31
REFERENCES ..................................................................................................................... 32

CHAPTER 2 .......................................................................................................................... 39
SYNTHESIS OF CRYPTOCHROME MODULATOR ............................................................ 39
2.1 Introduction .................................................................................................................... 39
2.2 Structure-activity relationship of KL001 ....................................................................... 40
2.3 Synthesis of the Zoltowski Library (ZL series) ............................................................ 43
  2.3.1 Synthesis of KL001 .................................................................................................. 43
2.4 Design principle for ZL001 .......................................................................................... 44
  2.4.1 Synthesis of δ-Carbazole using Benzyne Chemistry ............................................... 48
    2.4.1a. A 2,3-pyridyne approach for delta carboline synthesis ....................................... 49
    2.4.1b. A benzyne approach for delta carboline synthesis ............................................ 50
2.5 Design Principle of ZL002 and ZL003 ......................................................................... 54
  2.5.1. Synthesis of ZL002 ................................................................................................. 54
5.2 LOV Domain................................................................................................................. 144

5.2.1 Flavin chemistry ................................................................................................. 146

5.2.3 LOV photocycle and kinetics ............................................................................ 146

5.2.4 Factors affecting LOV photocycle .................................................................... 149

5.3 ZTL Family .............................................................................................................. 151

5.3.1 Arabidopsis thaliana LKP2.................................................................................. 154

5.3.2 Brassica rapa LKP2 ........................................................................................... 155

5.4 Results and discussions ....................................................................................... 156

5.4.1 Thermal kinetics of A. thaliana LKP2................................................................. 157

5.4.2 Thermal kinetics of BrLKP2- B and C .............................................................. 159

5.4.2a Thermal kinetics of BrLKP2 B........................................................................... 159

5.4.2b Thermal kinetics of BrLKP2 C ......................................................................... 161

5.5 Materials and methods ......................................................................................... 162

REFERENCES .............................................................................................................. 165

APPENDIX I .................................................................................................................. 169

CRYSTAL STRUCTURE OF CRYPTOCHROME LIKE ALGAL PROTEIN: OTCPF1 ..... 169

A1. Introduction ............................................................................................................. 169

A1.2 Protein Crystallization ......................................................................................... 170

A2 Results and discussion ......................................................................................... 171

A3 Future direction ...................................................................................................... 173
A4 Materials and methods................................................................. 173

REFERENCES ....................................................................................... 175

APPENDIX II.......................................................................................... 176

SCANNED $^1$H AND $^{13}$C NMR SPECTRA .............................................. 176
LIST OF FIGURES

Figure 1: An abstract view of circadian rhythm: The input signal activates the pacemaker of the biological clock, which in turn, drives the expression in a rhythmic fashion....................4

Figure 2: General mechanism of autoregulatory feedback inhibition in the basic circadian clock: The positive element drives the expression of the negative element, which in-turn inhibits action of the positive element. (adapted from ref: 15)........................................5

Figure 3: Important clock parameters: Period is the time taken to complete one cycle. Phase is the time at a particular event. Amplitude is half of the distance between a peak and trough .7

Figure 4: An abstract view of the mammalian circadian rhythm: The Suprachiasmatic nucleus (SCN) acts as a central oscillator. Light, the primary source of input activates the SCN, which in turn relays the message to the peripheral organs. The clock also receives reciprocal effect from the output such as food intake, temperature, and physical activity...............................9

Figure 5: A model of mammalian circadian rhythm system: Light activates the retinal ganglion cells in eyes, the signal gets relayed to the Suprachiasmatic nucleus (SCN), which acts as a pacemaker of the circadian rhythm. The SCN then entrains signals to activate peripheral organs via neural and endocrine pathways.................................................................10

Figure 6: The molecular architecture of mammalian circadian rhythm: An autoregulatory circadian feedback loop involves: the core loop consisting of the activators: CLOCK/BMAL1 and the repressors: crys and pers. The stabilizing loop involves REV-erba and ROR that regulates the expression of BMAL1. CLOCK/BMAL1 heterodimer also regulates the downstream clock-controlled genes (ccg). The stability of CRY and PER is regulated by the F-box proteins FBXL3 and β-Trcp respectively. The kinases, Casein kinase 1(Ck1) and AMPK phosphorylates PER and CRY respectively to promote the ubiquitination via their respective F-box proteins. The F-box proteins tag PER and CRY for the degradation by the 26S proteasome complex. (adapted from ref:33)..................................................................13
Figure 7: Overall view of the circadian rhythm in mammals: Light activates the Suprachiasmatic nucleus located in the brain. The SCN then acts as a pacemaker of the circadian rhythms and relays the signal down to the peripheral organs though neural or hormonal signals. On the molecular level, the core circadian clock involves clock and bmal heterodimerize rhythmically to drive the expression of CRY and PER, which in turn repress the expression of CLOCK: BMAL1 through feedback inhibition.

Figure 8: Endogenous ligands of the important circadian proteins. The structure of the ligands is used as a reference to synthesize small drug-like molecules to target clock proteins.

Figure 9: Circadian clock modulators: Some of the circadian clock modulators and their protein targets in the different stages of the negative feedback loop. Adapted from ref:13

Figure 10: Domain architecture of the mammalian cryptochrome: Mammalian cryptochrome has a photolyase homology region along with a disordered C-terminal extension.

Figure 11: Glucose homeostasis: During daytime or the feeding cycle, the body utilizes glucose supplied by food-intake and stores excess glucose in muscles and adipose tissue. On the contrary, at night time or the fasting cycle, glucagon gets secreted by the pancreas which activates the G-protein coupled receptors and thereby camp. The camp ultimately phosphorylates CREB, and that activates gluconeogenic genes to secrete glucose from non-carbohydrate sources.

Figure 12: Cryptochrome acts as a gate for the gluconeogenesis process: Cryptochrome is active during the feeding cycle and inhibits glucagon-stimulated camp production which further prohibits the gluconeogenesis process. At the nighttime or at the fasting cycle, cryptochrome becomes inactive and glucagon induces the gluconeogenesis process to take place.

Figure 13: Cryptochrome as a drug discovery target: A stabilized cryptochrome would restrict gluconeogenesis process for a longer period which potentially could aid in lowering the blood glucose level and hence would help the diabetic patient in maintaining an optimal level of glucose.
Figure 14: Structure of KL001 that could stabilize cryptochrome and delay its degradation process.  ................................................................. 29

Figure 15: KL001 mode of action. KL001 competitively binds at the FBXL3 binding site of the flavin and prevents it from degradation, which in turn lengthens the period. ............................. 30

Figure 16: Comparison of KL001 and FAD: (a): KL001 has carbazole as a scaffold (b): Flavin adenine dinucleotide (FAD); the isoalloxazine ring of FAD is subdivided into xylene, pyrazine and pyrimidine rings (c) Carbazole only mimics xylene and pyrazine rings of the isoalloxazine ring and lacks substituents which can participate in hydrogen bonding with the neighboring polar residues in the FAD binding pocket. ............................... 41

Figure 17: KL001 in a FAD binding pocket: The FAD binding site are enriched with polar residues. Remarkably, a polar residue arginine 376 (R 376) is just in the vicinity of the carbazole ring. Besides, another set of polar residues asparagine (N 411) and aspartic acid (D 405) are essential for the KL001 binding. ........................................................................... 42

Figure 18: ZL001 at the binding site: In ZL001, an extra nitrogen present at the para position possibly can enhance the binding efficiency of the compound at the flavin binding site. Presumably, it can participate in H-bonding interactions with R 376, N 411. (The fig was made using pdb 4MLP of KL001 complexed with CRY, the nitrogen was introduced to promote understanding) ........................................................................... 46

Figure 19: Retrosynthetic analysis shows two possible ways Delta carbazole could be synthesized using aryne chemistry ........................................................................................................ 48

Figure 20: 2,3-Pyridyne approach for the synthesis of Delta carbolines: (a) Pyridyne precursor can be coupled with the halo benzyne, the coupled product can be cyclized to yield delta carbolines. (b): Regioselectivity of pyridyne precursor at the 2nd position. ................................. 49

Figure 21: Benzyne approach: An in-situ generated benzyne will be coupled with the the ortho-aminopyridine derivation, subsequent cyliczation yield delta carboline .................................. 50
Figure 22: A mild benzyne precursor: o-Trimethylsilylphenyl triflate: a fluoride source can abstract the trimethylsilyl group, which subsequently eliminates the triflate group leading benzyne generation.................................................................51

Figure 23: Natural metabolite: Structural similarity of pinoline with the natural metabolites, serotonin and melatonin .................................................................57

Figure 24: Structure-activity of ZL004. Pinoline as a scaffold will place two substituents with the capability of participating in hydrogen bonding interactions on both sides of the tricyclic ring system. The methoxy group, presumably fill the portion of the pyrimidine ring of the isoalloxazine ring system. While on the other side, the hydrogen of secondary nitrogen can participate in critical interactions with the neighboring residues. The figure was generated using pdb = 4MLP for understanding purposes.........................................................58

Figure 25: Structure-activity of ZL005: Benzquinoxaline will physically add one more ring in the tricyclic ring system, which might potentially eliminate the requirement of having a substituent in a tricyclic ring system. .................................................................61

Figure 26: Structure-activity of the KL001 side chain: The side chain binding site is enriched with the aromatic and the polar residues. The furan ring of the KL001, snuggles close to the tryptophan (W 310) and polar glutamine (Q 307) residues.................................................................92

Figure 27: Structure-activity relationship of KL044: In KL044 the isopropyl linker is replaced by the amide linkage. The amide linkage participates in important interactions with the polar residues present in the binding pocket. While carbonyl functionality of the amide linker forms a hydrogen bond with the histidine 2.3 Å away, the serine residue forms a critical hydrogen bonding with the NH functionality of amide.................................................................93

Figure 28: Schematic representation of the cell-based circadian study: The potency of the small compounds to bind with cryptochrome and mimic FAD was tested using BMAL1-dluc reporter. Cryptochrome is a repressor of BMAL1 thereby expression of the BMAL1 can be used to monitor the cryptochrome activity. 96 well plate plated with U2OS cells harboring BMAL1-dluc reporter and the luminescence monitoring was done in the biotek plate reader. The output was analyzed by the R computing environment to observe the period altering effects. The curve fitting program was used to plot the graph between luminescence vs. Time. The JTK cycle was used to tabulate the period shift behavior of the compounds..................116
Figure 29: Activity of KL001: A) KL001 shifts period in the dose-dependent manner, however, rhythm gets abolished at higher concentration. B) The bar-graph clearly shows a dose-dependent increase in period shift.

Figure 30: A comparison of CRY modulators: KL001 (A) and ZL001 (B): X axis: time (days), Y axis: amplitude of luciferase activity. Both the compounds have shown a period shift. At the higher concentration, KL001 abolishes the rhythm while still maintained with ZL001. ZL001 showed robust rhythm at 0.1 µm (red) concentration (Collaborator: Dr. Carrie Partch UC Santa Cruz).

Figure 31: ZL001 delays rhythm at the higher concentrations: (A) Structure of ZL001 (B) A graph of luminescence vs. time shows delayed rhythm at the highest concentration (8 µm, teal) in the experimental condition, a slight shift is also seen at 2.7 µm (light brown). (C) A bar graph plotted between period vs. concentration shows the period shift from 25.05 (zero drug) to 26.72. (D) Comparison of ZL001 with KL001: KL001 shows dose dependent increase in the period shift activity, however abolishes the rhythm at the higher concentration, while ZL001 is active at the higher concentration and maintains rhythm at all the concentrations used in the experimental condition.

Figure 32: ZL002 delays rhythm at the higher concentrations: (A) Structure of ZL002 (B) A graph of luminescence vs. time clearly shows delayed rhythm at the higher concentrations (0.9 µm pink and 2.7 µm, brown and) in the experimental condition (C) A bar graph plotted between period vs. Concentration shows the period shift of 2.7 hours at 0.9 and 2.7 µm. (D) comparison of ZL002 with KL001: KL001 has a better activity at the lower concentration, while ZL002 is active at the higher concentration.

Figure 33: ZL003 delays rhythm at the higher concentrations: (A) Structure of ZL003 (B) A graph of luminescence vs. time clearly shows delayed rhythm at the highest concentration 8 µm (teal) in the experimental condition. (C) A bar graph plotted between period vs. Concentration shows the period shift from 25.05 (zero drug) to 26.72. (D) Comparison of ZL003 with KL001: KL001 has a better activity at the lower concentration, while ZL003 is active at the higher concentration.

Figure 34: Pinoline, a natural metabolite shows period altering effects: (a) Pinoline showed period lengthening effects at higher concentrations, robust rhythm was observed at all the concentrations indicating the pinoline binding at the FAD binding site. (b) The period lengthening of about 1.7 hours were observed at 2.7 and 8 µm concentrations.
Figure 35: ZL004 maintains robust rhythm while delays it at the higher concentrations: (A) structure of ZL004 (B) The curve fitting graph depicts robust rhythm was maintained at all the concentrations with rhythm delaying affects seen at the higher concentrations. (C) The period delaying effect can be clearly visible at 0.9 µm and 8 µm. (D) Comparison between KL001 and ZL004: at the 0.9 µm concentration period altering effect of ZL004 is almost similar to that of KL001. .............................. 125

Figure 36: ZL005 sustained robust rhythm and delayed it consistently at the higher concentrations: (A) structure of ZL005 (B) A curve fitting graph shows robust rhythm which gets delayed at the higher concentrations (yellow, teal). (C) A bar graph shows 1.7 hours period shift from 0.9 µm to 8 µm concentrations (D) Comparison of ZL005 and KL001: the period delaying effect of ZL005 is almost similar to that of KL001 at the 0.9 µm concentration.......................................................... 127

Figure 37: ZL006 exhibits robust rhythm with no period alteration affects: (A) Structure of ZL006 (B) A graph of luminescence vs. time shows robust rhythm maintained at all the concentrations of the drug. (C) A bar graph plotted between period vs. concentration shows no period shifts. (D) Comparison of ZL006 with KL001: KL001 has a better activity at the lower concentration, while ZL006 has no period altering effects. .............................. 128

Figure 38: ZL007 shows robust rhythm with no period alteration: (A) Structure of ZL007 (B) A graph of luminescence vs. time shows the rhythm maintained at all the concentrations of the drug. (C) A bar graph plotted between period vs. concentration shows no period shifts. (D) Comparison of ZL007 with KL001: KL001 has a better activity at the lower concentration, while ZL006 has no period altering effects. ........................................ 129

Figure 39: ZL008 has no period altering effects but robust rhythm: (A) Structure of ZL008 (B) a graph of luminescence vs. time shows rhythm maintained at all the concentrations of the drug. (C) A bar graph plotted between period vs. concentration shows no period shifts. (D) Comparison of ZL008 with KL001: KL001 has a better activity at the lower concentration, while ZL008 has no period altering effects. ............................................. 131

Figure 40: ZL009, a novel compound shows both robust rhythm and the period lengthening effects: (A) Structure of ZL009. (b) The novel compound maintains robust rhythm at all the concentrations, period lengthening observed at 2.7 µm. (c) The bar graph shows a significant delay in the period at 2.7 µm concentration. The period delay is about 3.3 hours similar to that of KL001 (d) Comparison of ZL009 with KL001: the period lengthening effect at 2.7 µm concentration is similar to that of KL001................................................................. 133
Figure 41: ZL010 showed robust rhythm at the lower concentration, while period lengthening effects at the higher concentration: (A) Structure of ZL010 (B) A graph of luminescence vs. time shows a delayed rhythm at the higher concentrations 2.7 and 8 µm of the drug. (C) A bar graph plotted between period vs. concentration shows period shifts by 1.2 hours (D) Comparison of ZL010 with KL001: KL001 has a better activity at the lower concentration, while ZL010 showed period lengthening effects at the higher concentration.

Figure 42: Different photoreceptors in the plant sense light and regulate several physiological processes.................................................................143

Figure 43: Schematic diagram of PAS domain fold: PAS domain is composed of 5 antiparallel beta sheets flanked by on one side series of four alpha helices.................................145

Figure 44: Crystal structure of the Zeitlupe exhibiting a typical PAS fold.........................145

Figure 45: Mechanism of the C4a adduct formation: (A) Upon blue light activation, a conserved cysteine residue in GXNCRFLQ motif of LOV domain participates in the C4a adduct formation with the isovalloxazine ring of flavin (B), (C) The conversion of light and dark state can be monitored by observing the flavin spectra; in the dark state classic flavin peak can be observed at 450 nm with the two shoulder peaks at 425 and 475 nm. In light state only one peak is observed at 380 nm. (D) Blue light activation excites flavin to the singlet excited state followed by the triplet state. Single electron transfer event takes place from the conserved cysteine, which results in the formation of a radical pair. The reduced flavin generates upon radical recombination.................................................................148

Figure 46: Glutamine flip: The amino functionality of the glutamine forms a hydrogen bond with the O4. Upon blue light activation, the C4a adduct formation leads to the change in the hydrogen bonding pattern. As a result, a conserved glutamine flips and orients in such a way that the carbonyl functionality of the glutamine forms a hydrogen bonding with the N5H proton........................................................................................................150

Figure 47: the schematic representation of domains in ztl family: ZTL family consists of F-box domain flanked by the LOV domain at the N-terminal and six kelch repeats at the C-terminal. LOV domain participates in photoactivation, while both F-box and kelch repeats take part in ubiquitination process. ..................................................................................................................152
Figure 48: A schematic representation of circadian and photoperiodic clock in *a.thaliana*. ZTL forms a complex with Gigantia (Gi) during the day time, at the same time FKF1 forms a complex with Gi and degrades Cyclin Dof Factor (CDF), thereby regulate photoperiodic timing. At night, ZTL gets dissociated with Gi and degrades TOC1 and thereby represses the transcription of circadian genes. The symbol $\rightarrow$ indicates degradation pathway. (adapted from: pudasaini et al. *Biochemistry* 2013, 52, 7150-7158) .......................................................... 153

Figure 49: Sequence alignment of the ZTL family: Sequence alignment showed LKP2 differs from ZTL and FKF1 at the conserved glutamine (Q) position instead it has leucine (L)...

Figure 50: Sequence alignment of all three LOV domain harboring F-box proteins in *Brassica rapa* along with the LOV domain proteins of the ZTL family. The sequence alignment clearly reveals all the proteins are similar to LKP2 of the *A thaliana*. Like LKP2, instead of conserved glutamine (Q) all the three proteins of *B.rapa* have leucine (L) .......

Figure 51: (A) The spectra of LKP2 protein exemplifies the C4a adduct formation. The dark state (black) at 450 nm with the shoulder peaks at 425 and 475 nm represents oxidized flavin. Blue light bleaching results in a light state peak (red) representing a C4a adduct formation. (B) reversion of the light to the dark state demonstrate first order kinetics as shown by the absorbance band at 450 nm (red) and 478 nm (black). .......................................................... 157

Figure 52: Thermal kinetics comparison of *A. thaliana* LKP2 and LKP2 (L153Q) mutant: A and C: Arrhenius and Eyring plot of LKP2 (16-165). B and D: Arrhenius and Eyring plot of the mutant. .......................................................... 158

Figure 53: Thermal kinetics of BrLKP2-B and BrLKP2-B (L149Q) mutant: A and C: Arrhenius and eyring plot of BrLKP2-B (29-165). B and D: Arrhenius and Eyring plot of the mutant L149Q mutant .......................................................... 160

Figure 54: Thermal kinetics of BrLKP2-C and BrLKP2-C (L149Q) mutant: A and C: Arrhenius and Eyring plot of BrLKP2-C (29-165). B and D: Arrhenius and Eyring plot of the mutant (L149Q). .......................................................... 161

Figure 55: Structure of OtCPf1 (A) The cartoon structure of OtCPf1 (B) FAD bound in the FAD binding pocket (C) Residues which were found to be essential for the KL001 binding are
present and labeled. (D) Side chain of the FAD snuggle close to the tryptophan and glutamine residues.
LIST OF TABLES

Table 1: Crystal structure known for the core circadian proteins (Adapted from ref:28) ........15

Table 2: Summary of the diseases associated with the mutation of
different core clock elements........................................................................................................18

Table 3: Summary of the small molecules synthesized to mimic the endogenous ligand of the
core circadian proteins and their therapeutic actions........................................................................................................22

Table 4: Kinetic parameters of A. thaliana LKP2 and LKP2 (L153Q) mutant .................... 159

Table 5: Kinetic parameters of B.rapa LKP2-B and B.rapa LKP2-B (L149Q) mutant.......... 160

Table 6: Kinetic parameters of B. rapa LKP2-C and B. rapa LKP2-C (L153Q) mutant........ 162
LIST OF SCHEMES

Scheme 1: Synthesis of KL001 .................................................................44
Scheme 2: Synthesis of Delta-Carboline......................................................47
Scheme 3: Synthesis of benzyne precursor .................................................52
Scheme 4: Synthesis of Delta-Carboline from benzyne precursor ...............53
Scheme 5: Synthesis of ZL001 .................................................................53
Scheme 6: Synthesis of ZL002 .................................................................55
Scheme 7: Synthesis of ZL003 .................................................................56
Scheme 8: Synthesis of ZL004 .................................................................59
Scheme 9: Synthesis of ZL005 .................................................................62
Scheme 10: Synthesis of ZL006 ...............................................................95
Scheme 11: Synthesis of ZL007 ...............................................................96
Scheme 12: Synthesis of ZL008 ...............................................................97
Scheme 13: Synthesis of ZL009 ...............................................................98
Scheme 14: Synthesis of ZL010 ...............................................................99
<table>
<thead>
<tr>
<th>Abbreviations</th>
<th>Name</th>
<th>Structure</th>
</tr>
</thead>
<tbody>
<tr>
<td>CLOCK</td>
<td>Circadian locomotor output cycle kaput</td>
<td></td>
</tr>
<tr>
<td>BMAL1</td>
<td>Brain and muscle aryl hydrocarbon receptor nuclear translocator like protein-1</td>
<td></td>
</tr>
<tr>
<td>E-Box</td>
<td>Enhancer Box</td>
<td></td>
</tr>
<tr>
<td>CRY</td>
<td>Cryptochrome</td>
<td></td>
</tr>
<tr>
<td>PER</td>
<td>Period</td>
<td></td>
</tr>
<tr>
<td>RRE</td>
<td>Rev-ERB/ROR-binding element</td>
<td></td>
</tr>
<tr>
<td>CPF family</td>
<td>Cryptochrome photolyase family</td>
<td></td>
</tr>
<tr>
<td>SCN</td>
<td>Suprachiasmatic nucleus</td>
<td></td>
</tr>
<tr>
<td>PHR</td>
<td>Photolyase homology region</td>
<td></td>
</tr>
<tr>
<td>cAMP</td>
<td>Cyclic adenosine monophosphate</td>
<td></td>
</tr>
<tr>
<td>CREB</td>
<td>cAMP-response element binding</td>
<td></td>
</tr>
<tr>
<td>RORs</td>
<td>Retinoic acid receptor related orphan receptors</td>
<td></td>
</tr>
<tr>
<td>NAD</td>
<td>Nicotinamide adenine dinucleotide</td>
<td></td>
</tr>
<tr>
<td>CK1</td>
<td>Casein kinase-1</td>
<td></td>
</tr>
<tr>
<td>SCFFBXL3</td>
<td>SKP–Cullin–Rbx–F-box and leucine rich repeat</td>
<td></td>
</tr>
<tr>
<td>SIRT1</td>
<td>Sirtuin 1</td>
<td></td>
</tr>
<tr>
<td>AMPK</td>
<td>Adenosine monophosphate activated protein kinase</td>
<td></td>
</tr>
<tr>
<td>Ccg</td>
<td>Clock controlled genes</td>
<td></td>
</tr>
<tr>
<td>TTFL</td>
<td>Transcription translation feedback loop</td>
<td></td>
</tr>
<tr>
<td>DNA</td>
<td>Deoxyribose nucleic acid</td>
<td></td>
</tr>
<tr>
<td>SAD</td>
<td>Seasonal affective disorder</td>
<td></td>
</tr>
<tr>
<td>GPCR</td>
<td>G-protein coupled receptor</td>
<td></td>
</tr>
<tr>
<td>Pck1</td>
<td>Phosphoenolpyruvate carboxykinase-1</td>
<td></td>
</tr>
<tr>
<td>G6Pc</td>
<td>Glucose-6-phosphate</td>
<td></td>
</tr>
<tr>
<td>ZTL</td>
<td>Zeitlupe</td>
<td></td>
</tr>
<tr>
<td>FKF1</td>
<td>Flavin binding kelch repeat F-box-1</td>
<td></td>
</tr>
<tr>
<td>LKP2</td>
<td>LOV kelch protein-2</td>
<td></td>
</tr>
<tr>
<td>LOV</td>
<td>Light oxygen voltage</td>
<td></td>
</tr>
<tr>
<td>Acronym</td>
<td>Description</td>
<td></td>
</tr>
<tr>
<td>---------</td>
<td>-------------</td>
<td></td>
</tr>
<tr>
<td>PAS</td>
<td>Period-ARNT-Single-minded</td>
<td></td>
</tr>
<tr>
<td>ARNT</td>
<td>Aryl hydrocarbon receptor nuclear translocator</td>
<td></td>
</tr>
<tr>
<td>SCF</td>
<td>SKP–Cullin–Rbx–F-box</td>
<td></td>
</tr>
<tr>
<td>β-TrCP</td>
<td>beta-transducin repeat containing E3 ubiquitin protein ligase</td>
<td></td>
</tr>
<tr>
<td>FBS</td>
<td>Fetal Bovine Serum</td>
<td></td>
</tr>
<tr>
<td>DMEM</td>
<td>Dulbecco’s Modified Eagle Media</td>
<td></td>
</tr>
<tr>
<td>LOPAC</td>
<td>Library of pharmacologically active compounds</td>
<td></td>
</tr>
<tr>
<td>OtCPF1</td>
<td><em>Ostreococcus tauri</em> cryptochrome photolyase family I</td>
<td></td>
</tr>
<tr>
<td>SAR</td>
<td>Structure-Activity relationship</td>
<td></td>
</tr>
<tr>
<td>FDA</td>
<td>Federal Drug and Administration</td>
<td></td>
</tr>
<tr>
<td>GI</td>
<td>Gigantea</td>
<td></td>
</tr>
<tr>
<td>TOC1</td>
<td>Timing of cab I</td>
<td></td>
</tr>
<tr>
<td>CDF</td>
<td>Cyclin Dof Factor</td>
<td></td>
</tr>
<tr>
<td>TEV</td>
<td>Tobacco Etch Virus</td>
<td></td>
</tr>
<tr>
<td>DTT</td>
<td>Dithiothreitol</td>
<td></td>
</tr>
<tr>
<td>TMSCl</td>
<td>Trimethylsilylchloride</td>
<td></td>
</tr>
<tr>
<td>DBU</td>
<td>1, 8-Diazabicyclo-[5.4.0]-undec-7-ene</td>
<td></td>
</tr>
<tr>
<td>&quot;^BuLi</td>
<td>Butyllithium</td>
<td></td>
</tr>
<tr>
<td>LDA</td>
<td>Lithium diisopropylamide</td>
<td></td>
</tr>
<tr>
<td>DMF</td>
<td>N, N-Dimethylformamide</td>
<td></td>
</tr>
<tr>
<td>PhNTF&lt;sub&gt;2&lt;/sub&gt;</td>
<td>N-Phenyl-bis(trifluoromethanesulfonimide)</td>
<td></td>
</tr>
<tr>
<td>MeCN</td>
<td>Acetonitrile</td>
<td></td>
</tr>
<tr>
<td>NaH</td>
<td>Sodium hydride</td>
<td></td>
</tr>
<tr>
<td>Py</td>
<td>Pyridine</td>
<td></td>
</tr>
<tr>
<td>Abbreviation</td>
<td>Chemical Name</td>
<td>Structural Formula</td>
</tr>
<tr>
<td>--------------</td>
<td>----------------------------</td>
<td>--------------------</td>
</tr>
<tr>
<td>EtOAc</td>
<td>Ethyl acetate</td>
<td><img src="#" alt="EtOAc" /></td>
</tr>
<tr>
<td>DMA</td>
<td>N, N-Dimethylacetamide</td>
<td><img src="#" alt="DMA" /></td>
</tr>
<tr>
<td>THF</td>
<td>Tetrahydrofuran</td>
<td><img src="#" alt="THF" /></td>
</tr>
<tr>
<td>CsF</td>
<td>Cesium Fluoride</td>
<td><img src="#" alt="CsF" /></td>
</tr>
<tr>
<td>OTf</td>
<td>Triflate group</td>
<td><img src="#" alt="OTf" /></td>
</tr>
<tr>
<td>TBSOTf</td>
<td>tert-Butyldimethylsilyl trifluoromethanesulfonate</td>
<td><img src="#" alt="TBSOTf" /></td>
</tr>
<tr>
<td>AcOH</td>
<td>Acetic Acid</td>
<td><img src="#" alt="AcOH" /></td>
</tr>
<tr>
<td>MeOH</td>
<td>Methanol</td>
<td><img src="#" alt="MeOH" /></td>
</tr>
<tr>
<td>TFA</td>
<td>Trifluoroacetic acid</td>
<td><img src="#" alt="TFA" /></td>
</tr>
<tr>
<td>BoC₂O</td>
<td>Di-tert-butyldicarbonate</td>
<td><img src="#" alt="BoC₂O" /></td>
</tr>
<tr>
<td>EtOH</td>
<td>Ethanol</td>
<td><img src="#" alt="EtOH" /></td>
</tr>
<tr>
<td>TMEDA</td>
<td>Tetramethylemedamine</td>
<td><img src="#" alt="TMEDA" /></td>
</tr>
<tr>
<td>DCM</td>
<td>Dichloromethane</td>
<td><img src="#" alt="DCM" /></td>
</tr>
<tr>
<td>TEA</td>
<td>Triethylamine</td>
<td><img src="#" alt="TEA" /></td>
</tr>
</tbody>
</table>
FMN  Flavin mononucleotide

FAD  Flavin adenine dinucleotide

KL001  Cryptochrome Modulator

KL044  Cryptochrome Modulator
**LIST OF AMINO ACIDS**

<table>
<thead>
<tr>
<th>Amino acid, AA. (X)</th>
<th>Structure</th>
</tr>
</thead>
<tbody>
<tr>
<td>Glycine, Gly. (G)</td>
<td><img src="image" alt="Glycine Structure" /></td>
</tr>
<tr>
<td>Alanine, Ala. (A)</td>
<td><img src="image" alt="Alanine Structure" /></td>
</tr>
<tr>
<td>Valine, Val. (V)</td>
<td><img src="image" alt="Valine Structure" /></td>
</tr>
<tr>
<td>Leucine, Leu. (L)</td>
<td><img src="image" alt="Leucine Structure" /></td>
</tr>
<tr>
<td>Isoleucine, Ile. (I)</td>
<td><img src="image" alt="Isoleucine Structure" /></td>
</tr>
<tr>
<td>Cysteine, Cys. (C)</td>
<td><img src="image" alt="Cysteine Structure" /></td>
</tr>
<tr>
<td>Methionine, Met. (M)</td>
<td><img src="image" alt="Methionine Structure" /></td>
</tr>
<tr>
<td>Tryptophan, Trp. (W)</td>
<td><img src="image" alt="Tryptophan Structure" /></td>
</tr>
<tr>
<td>Phenylalanine, Phe. (F)</td>
<td><img src="image" alt="Phenylalanine Structure" /></td>
</tr>
<tr>
<td>Amino Acid</td>
<td>Abbreviation</td>
</tr>
<tr>
<td>------------</td>
<td>--------------</td>
</tr>
<tr>
<td>Tyrosine</td>
<td>Tyr (Y)</td>
</tr>
<tr>
<td>Serine</td>
<td>Ser (S)</td>
</tr>
<tr>
<td>Threonine</td>
<td>Thr (T)</td>
</tr>
<tr>
<td>Asparagine</td>
<td>Asp (N)</td>
</tr>
<tr>
<td>Glutamine</td>
<td>Gln (Q)</td>
</tr>
<tr>
<td>Lysine</td>
<td>Lys (K)</td>
</tr>
<tr>
<td>Arginine</td>
<td>Arg (R)</td>
</tr>
<tr>
<td>Histidine</td>
<td>His (H)</td>
</tr>
<tr>
<td>Aspartic Acid</td>
<td>Asp (D)</td>
</tr>
<tr>
<td>Glutamic Acid</td>
<td>Glu (E)</td>
</tr>
</tbody>
</table>
LIST OF THE ZOLTOWSKI LIBRARY

ZL001

ZL002

ZL003

ZL004

ZL005

ZL006

xxxii
ZL007

ZL008

ZL009

ZL010
This is dedicated to my Parents, my Grandmother, my best friend and husband Anurag
CHAPTER 1
CIRCADIAN RHYTHM: NEW AVENUES FOR THE DRUG DISCOVERY

A famous dictum by Benjamin Franklin, “Early to bed, early to rise” is based on the theory that night is for sleeping while waking up should take place with the sunrise. Indeed, time plays a vital role in day to day life. A sudden shift in the routine can make us dysfunctional. Jet-lag is inevitable while traveling to the different time-zones. It is because of an internal circadian (biological) clock which dictates the daily physiological behaviors. This chapter provides a brief overview of the mammalian circadian rhythm, and asks a fundamental question; How does it govern our daily behavior? Are there any diseases associated? And if yes, then what efforts are being made to discover clock-based therapies to improve the human physiology.

1.1 Introduction

The key to life existing on the Earth lies in its ability to adapt to the environment. Organisms living in the different geographical conditions have adapted to distinct environment necessary for the survival. Essential to such adaptation and survival is the ability to synchronize daily events to the diurnal solar cycle. To adapt to such changes, most living organisms on the
Earth have evolved an internal biological clock that anticipates the fluctuation in day (light) and night (dark) pattern and optimizes the organisms’ physiology and behavior\(^1\). This intrinsically generated biological clock is known as a circadian rhythm.

Circadian rhythms are self-autonomous endogenous clocks synchronized with the rotation of the Earth. Circadian clocks are defined by three requirements 1) They must demonstrate ~24-hour period that is consistent with the Earth’s rotation and revolution on its axis and governs day to day activities in the organisms\(^1\)–\(^5\). 2) The circadian period must be temperature independent (temperature compensated). 3) Despite the rhythm being synchronized with the Earth rotation, it must not depend on the sunlight (or any other exogenous cue) for its sustainability; indeed circadian rhythms still persist even in the dark hence they are endogenous free running rhythms of approximately 24-hours\(^5\)–\(^7\).

The internal biological clock anticipates dusk and dawn and prepares the organism to harness the maximum benefits of each aspect of the day/night cycle, while mitigating dangers associated with the damaging effects of the sun’s rays. Circadian rhythms are ubiquitous and are found both in prokaryotes and eukaryotes. Amongst prokaryotes, only cyanobacteria have been shown to have a robust circadian rhythm; which is likely associated with their photosynthetic abilities that necessitate temporal regulation of energy collection and use\(^8\)–\(^9\). In eukaryotes such as plants, their sessile nature requires coordination of their growth and development with diurnal changes in environmental conditions. Hence, plants have developed an intricate array of several photoreceptors which work in a circadian fashion to assist plants in adjusting to the surroundings. These photoreceptors sense a broad spectrum of light which facilitates acclimation to the surrounding environment and regulates their growth and development\(^10\)–\(^11\). Although mammals are not directly dependent on light for metabolic needs or growth and development, mammals retain
analogous circadian networks to temporally segregate metabolic pathways. Hence in mammals the circadian clock plays a vital role in governing daily metabolic activities like a sleep-wake cycle, hormone secretion and diverse aspects of human physiology.

Hence, it is not wrong to say circadian rhythm is a “timekeeping” mechanism present in nearly all the organisms residing on the Earth. Therefore, it is imperative to study the molecular basis of circadian clocks to provide a deeper insight into the nature of biological timekeeping and the intricate mechanisms necessary to coordinate and synchronize such a clock with environmental factors. In this chapter, we go in depth to understand how the clock works through the implementation of a negative feed-back loop. Further, we will develop a brief history of the circadian rhythm research. Finally, this chapter will delve into the specific biochemical pathways that maintain mammalian circadian rhythm and discuss how the core circadian protein cryptochrome plays a vital role in human disease including in regulating diabetes.

1.2 Characteristic and Molecular basis of the circadian clock

The core component of a basic circadian clock consists of an input pathway, which perceives the stimulus from the environment and transmits the environmental signal to a central oscillator. The central oscillator, in turn, entrains the signal and maintains ~24-hour endogenous rhythm that modulates diurnal regulation of desired outputs such as gene expression. Notably, some output feedback into the core oscillator rhythmically suppresses gene expression. As a result, the next cycle starts, and the process continues. This process is called feedback inhibition and is a central component of all clock networks5,6,12-13.
Figure 1: An abstract view of circadian rhythm: The input signal activates the pacemaker of the biological clock, which in turn, drives the expression in a rhythmic fashion.

All circadian processes follow the same pathway in all the organisms residing on the Earth, e.g. plants, fungi and animals. However, the number of proteins participating differs, depending upon the complexity of an organism. For these reasons, much of the initial circadian rhythm research focused on understanding how the central oscillator is able to maintain a temperature compensated 24-hour period through regulation of a negative feedback loop.

1.2.1 The feedback loop

“What goes up, must come down, spinning wheel got to go around.”

At the core molecular level, circadian rhythm involves successive gene expression and repression through a transcription and translation feedback loop (TTFL). The process is well-coordinated and orchestrated like a musical rhythm. Just like a wall clock, the feedback loop consists of two elements: a positive limb (or element) which drives the expression of all genes under control of the circadian clock and the negative limb (or element) that inhibits gene expression leading the path for a new cycle to begin. Importantly, the TTFL and feedback loops
do not operate in isolation. Coincident with transcription/translation numerous post translational regulation events modulate the stability and function of key clock proteins. For instance, a set of ubiquitin ligase protein maintains the periodicity by subsequently degrading the repressor proteins, such that a new cycle can get started with the 24 hours periodicity and thereby improving the robustness of the circadian network and facilitating adaptation in the presence of diverse environmental signals. By this process, the internal clock is synchronized with environmental signals allowing an organism to both maintain a 24-hour period (TTFL), but also adapt to environmental change or pressures through peripheral loops that involve post-translational pathways. Without this complicated interplay between the core TTFL and peripheral loops, we would not be able to adapt to different time zones and overcome jet lag\textsuperscript{14-17}.

\textbf{Figure 2: General mechanism of autoregulatory feedback inhibition in the basic circadian clock}: The positive element drives the expression of the negative element, which in-turn inhibits action of the positive element. (Adapted from ref: 15).
1.2.2 Important clock parameters

The autoregulatory feedback inhibition (circadian rhythm) of gene regulation can be tested in a laboratory. Experiments are typically done by measuring gene expression, behavior cues or epigenetic modification over a multiple-day time course. These time courses are conducted on samples after they have been synchronized (entrained) by the environmental stimuli such as: alternating light and dark cycle or drug or small molecule introduction. In conducting such studies, there are specific terms used during these experimental approaches which are unique to the field. We explain these terms below:

Zeitgeber time (ZT): is the German term which means “time giver.” It is a standard notation of 24-hour phase, where ZT = 0 means dawn (or beginning of the experiment) and ZT = 12 is termed as the beginning of the dark phase or night.

Period: Period is the time taken to finish one cycle. It can be measured as the distance between two consecutive peaks or troughs (Fig 3). To be considered circadian, the period must remain approximately 24-hours regardless of temperature.

Phase: Phase is the time of day at which a particular event occurs. For instance, if an element within the circadian rhythm demonstrates a peak at 12 hours after ZT = 0 (dawn), then the phase will be 12 hours\(^10\).

Amplitude: Amplitude of the rhythm is the one half of the peak to trough distance\(^10\).
Figure 3: Important clock parameters: Period is the time taken to complete one cycle. Phase is the time at a particular event. Amplitude is half of the distance between a peak and trough.

1.3 Discovery of circadian rhythm to chronotherapy: Historical perspective

Circadian rhythms were said to be first observed by the Greek geographer Androsthenes at the time of Alexander the Great in 4th century B.C when he noticed daily leaf movements of the tamarind tree, *Tamarindus indicus* on the islands of Tylos (now Bahrain) in the Persian Gulf. Despite such observations, it was another two millennia until the first experimental studies on circadian rhythms were conducted. They were conducted by the French astronomer Jean-Jacques d’Ortous de Mairan, who reported leaf movements of the sensitive *Mimosa pudica* plant (commonly known as “touch me not” plant) were persistent even in the constant darkness, which led him to conclude the rhythm to be endogenous and self-sustaining\(^\text{16}\). He later extended his research to the human patients with sleep defects. After that, in 1729 he published the first paper in circadian rhythm research. These experiments were not without controversy, and numerous researchers tried to debunk or better understand how non-human species could measure time. Research on circadian rhythms became more formalized in the 20th century. In 1950s Franz
Halberg coined the word “circadian” from the Latin word circa diem (about a day)\textsuperscript{1,19}. Over the last 100 years, identification of the molecular components of circadian clocks has attracted the attention of both biochemists and chemists, who aimed to identify fundamental chemical and biochemical components necessary to maintain biological clocks. These efforts have led to substantial research aimed at unraveling how circadian rhythms dictate growth and development in plants, to aid in crop production, and regulation of diverse aspects of human physiology as it relates to disease phenotypes. Today it is a fast-growing field in biomedical research. Delightfully, last year in 2017, the Nobel prize in physiology or medicine was awarded to Jeffery C. Hall, Michael Rosbash and Michael W. Young for their discoveries of the molecular mechanism that controls circadian rhythms.

1.4 Mammalian circadian rhythm

In mammals, the circadian clock plays pivotal roles in governing physiological and metabolic processes like the sleep-wake cycle, hormone secretion, bowel movement, body temperature, heart rate, and numerous others\textsuperscript{14,15}. Sudden changes in environmental inputs or behavior thus leads to misalignment between the environment and physiological processes. Diverse studies indicate that misalignment of the clock contributes to diverse disease phenotypes and disorders. Importantly, circadian rhythms maintain the ability to reset the clock to re-entrain endogenous rhythms in response to environmental inputs. For instance, traveling in another part of the globe can lead to jet lag as the sleep-wake cycle gets disturbed. After some period, however, we will naturally adapt to the new environment in response to metabolic cues (feeding) and environmental light input. That is why one of the primary mechanisms recommended to deal with the jet lag is to eat according the time of the day, and to look at the morning sun at the destination and sleep according to the local day-night cycle.
The circadian system of the mammals includes the input pathways, the central clock and, the output. Light is the principal input; however, many output activities like food, temperature, and physical activity reciprocally regulate the clock\textsuperscript{10,13}.

\begin{figure}[h]
\centering
\includegraphics[width=\textwidth]{circadian_system_diagram}
\caption{An abstract view of the mammalian circadian rhythm: The suprachiasmatic nucleus (SCN) acts as a central oscillator. Light, the primary source of input activates the SCN, which in turn relays the message to the peripheral organs. The clock also receives reciprocal effect from the output such as food intake, temperature, and physical activity.}
\end{figure}

Light is the primary synchronizer of the biological clock. Light enters through the eyes, gets perceived by the photoreceptor cells retinal ganglion located in the retina which then relays the signal down through the optic nerve into the suprachiasmatic nucleus located in the hypothalamus. The suprachiasmatic nucleus is the master regulator, also known as the pacemaker, of the circadian clock in mammals\textsuperscript{20,22}. Within each nucleus, there are about 20,000 neurons capable of generating circadian rhythm in a self-autonomous manner. Subsequently, neural and hormonal signals from the SCN entrains peripheral organs and tissues, synchronizing clocks in nearly all cell types to the daily light-dark cycle. These peripheral organs and tissues then behave
in a circadian fashion and participate in hormone secretion or other metabolic processes to maintain homeostasis.\textsuperscript{14, 16,18, 21-25}

**Figure 5: A model of mammalian circadian rhythm system:** Light activates the retinal ganglion cells in eyes, the signal gets relayed to the suprachiasmatic nucleus (SCN), which acts as a pacemaker of the circadian rhythm. The SCN then entrains signals to activate peripheral organs via neural and endocrine pathways.

Because of the biological clock, the human body has a natural tendency to maintain homeostasis. Disruption or misalignment of this behavior due to genetic manipulation or environmental factors can lead to many physiological defects, such as sleep disorders, cardiovascular disease, diabetes, obesity, cancer and other metabolic disorders. Consequently, understanding the biochemical mechanisms dictating circadian clock function can open new avenues in treating metabolic diseases or disorders at the genomic level and could have a
substantial impact on human health\textsuperscript{13}. This has paved the path of chronobiology or chronotherapy, which means treating disorders by modulating the natural rhythm.

1.5 Molecular architecture of the mammalian circadian clock

In mammals, the canonical auto-regulatory transcription-translation feedback loop (TTFL) consists of four core components: two transcription activators: CLOCK (Circadian Locomotor output cycles kaput) and BMAL1 (Brain and Muscle Aryl-hydrocarbon receptor nuclear translocator Like protein-1), and the transcriptional repressors, Periods (PERs) and Cryptochromes (CRYs). In the core clock network transcription activators CLOCK and BMAL1 act as a positive arm, whereas two of their target genes CRYs and PERs act as the repressor proteins to form the negative limb of the TTFL. (Fig. 6 and 7)

Mechanistically, this is achieved by the transcription activator proteins CLOCK and BMAL1, which heterodimerize to dictate the expression of clock-controlled genes (ccgs) via E-box promoter elements. Upon accumulation, CRY and PER, in turn, heterodimerize in the cytoplasm, translocate back to the nucleus to restrict their transcription by retarding CLOCK:BMAL1 transcription. Currently, the mechanism by which CRY and PER alter CLOCK:BMAL1 expression is poorly understood, but it is believed to involve a large (megadalton) protein complex recruiting numerous chromatin remodeling agents and transcriptional repressors to the promoter elements. In conjunction with CLOCK:BMAL1 repression, CRY and PER are targeted for degradation by ubiquitin-dependent proteolysis pathways. Degradation of CRY and PER relieves the repression on CLOCK and BMAL1 gets relieved, and the process continues again with approximately 24-hour periodicity\textsuperscript{1-5,15-18,21}.

The ubiquitination mediated proteasomal degradation pathway in the negative feedback loop plays a crucial role in maintaining a robust circadian oscillator. The ubiquitination process
involves targeting protein for the 26S proteasomal degradation. At the post-transcriptional level, SCF (Skp1-Cullin-F-box protein) complexes such as β-TrCP (beta-transducin repeat containing E3 ubiquitin protein ligase) and FBXL3 (F-box and leucine-rich repeat protein 3) regulate the stability of PER and CRY respectively and maintain the circadian rhythmicity\textsuperscript{24,25}. More specifically, kinases like casein kinase (CK1)\textsuperscript{26} and AMPK (adenosine monophosphate-activated protein kinase) phosphorylatePER and CRY respectively. The respective F-box proteins then act upon the phosphorylated proteins and form an E3 ubiquitin ligase complex. The tagged protein, subsequently, undergoes degradation by the 26S proteasome complex. Another homolog F-box protein of FBXL3, FBXL21 antagonizes the CRY degradation in the nucleus; thereby facilitates the CRY turnover rate in the cytoplasm\textsuperscript{17, 21, 28, 31}. Such competitive processes are common in circadian networks and are believed to contribute to the robustness of the circadian oscillator.

Besides the core circadian oscillator, numerous feedback loops also participate in modulating circadian function. One such secondary stabilization loop consists of nuclear hormone receptors REV-ERB α/β and the antagonist-receptor ROR α/β/γ\textsuperscript{29,30}. Both REV-ERBs and ROR (retinoic acid receptor related orphan receptors) work antagonistically and dictate the expression of the core transcription factor BMAL1\textsuperscript{31, 32}. CLOCK/BMAL1 heterodimer, in turn, activates the ROR and Rev-erb gene expression\textsuperscript{21,33,34}. In this manner, peripheral loops can integrate diverse pathways such as stress response, metabolism and cell-cycle repair into the core clock network through secondary feedback loops. These secondary loops both facilitate cross-talk to diverse pathways, but also contribute to the stability and robustness of the core oscillator.
Figure 6: The molecular architecture of mammalian circadian rhythm: An autoregulatory circadian feedback loop involves: the core loop consisting of the activators: CLOCK/BMAL1 and the repressors: CRYs and PERs. The stabilizing loop involves Rev-erba and ROR that regulates the expression of BMAL1. CLOCK/BMAL1 heterodimer also regulates the downstream clock-controlled genes (ccg). The stability of CRY and PER is regulated by the F-box proteins FBXL3 and β-TrCP respectively. The kinases, casein kinase 1(CK1) and AMPK phosphorylates PER and CRY respectively to promote the ubiquitination via their respective F-Box proteins. The F-box proteins tag PER and CRY for the degradation by the 26S proteasome complex. (Adapted from Ref:33)

Importantly, all peripheral tissues and organs follow the same mechanism to couple physiological and metabolic pathways to the circadian clock. All the core proteins involved in the TTFL are crucial for the proper metabolic functioning. Indeed, several studies have shown complete loss of metabolic pathways in mice lacking CLOCK and BMAL1, and CRY functionalities34.
To summarize, the circadian rhythm in mammals is generated by a set of hierarchal interlocking genes that undergo self-autonomous transcription/translation feedback loops, resulting in cascades of gene expression. To promote understanding, Fig 7 shows the simplified version of the circadian loop.

![Circadian Rhythm Diagram]

**Figure 7: Overall view of the circadian rhythm in mammals:** Light activates the suprachiasmatic nucleus located in the brain. The SCN then acts as a pacemaker of the circadian rhythms and relays the signal down to the peripheral organs though neural or hormonal signals. On the molecular level, the core circadian clock involves CLOCK and BMAL heterodimerize rhythmically to drive the expression of CRY and PER, which in turn repress the expression of CLOCK: BMAL through feedback inhibition.

### 1.6 Structures of the core circadian rhythm proteins

The circadian clock genes in mammals show rhythmic expression not only in the suprachiasmatic nucleus (SCN) which is the pacemaker of the circadian clock generation but also in the peripheral organs. The core components of circadian clocks involve the hierarchy of genes working in a rhythmic fashion. Central to this function are numerous competing and/or cooperative protein:protein interaction networks that are modulated by post-translational modifications. To
unravel the mechanism by which these protein:protein interaction networks modulate clock function, and to delineate their signaling mechanisms, significant effort has been placed in obtaining high-resolution structures and/or low resolution biophysical information on clock proteins and functional protein complexes\textsuperscript{21,28}. Recently, the high-resolution crystal structures of several critical interacting proteins and proteins bound to the small molecule have become available. These include a crystal structure of the CLOCK: BMAL1 complex\textsuperscript{35}, the crystal structures of isolated domains of all three PER proteins\textsuperscript{36,37}, and crystal structures of multiple functional forms of CRY proteins\textsuperscript{38-41}. Table 1 summarizes the crystal structures of core clock components known so far along with their PDB codes and the ligand bound.

Table 1: Crystal structure known for the core circadian proteins (adapted from Ref:28)

<table>
<thead>
<tr>
<th>Protein</th>
<th>Domain</th>
<th>Ligand</th>
<th>PDB-ID</th>
</tr>
</thead>
<tbody>
<tr>
<td>mCLOCK: mBMAL1</td>
<td>bHLH</td>
<td>DNA</td>
<td>4H10\textsuperscript{35}</td>
</tr>
<tr>
<td>mPER1</td>
<td>PAS AB</td>
<td>DNA</td>
<td>4DJ2\textsuperscript{36}</td>
</tr>
<tr>
<td>mPER2</td>
<td>PAS AB</td>
<td>3GDI\textsuperscript{37}</td>
<td></td>
</tr>
<tr>
<td>mPER3</td>
<td>PAS AB</td>
<td>4DJ3\textsuperscript{37}</td>
<td></td>
</tr>
<tr>
<td>mCry apo form</td>
<td>PHR-CC</td>
<td>4K0R\textsuperscript{38}</td>
<td></td>
</tr>
<tr>
<td>FAD bound mCRY</td>
<td>PHR-CC</td>
<td>FAD</td>
<td>416G\textsuperscript{40}</td>
</tr>
<tr>
<td>mCRY with FBXL3</td>
<td>PHR-CC</td>
<td>FBXL3</td>
<td>4I6J\textsuperscript{40}</td>
</tr>
<tr>
<td>mCRY with KL001</td>
<td>PHR-CC</td>
<td>KL001</td>
<td>4MLP\textsuperscript{41}</td>
</tr>
<tr>
<td>CRY2 in complex with per</td>
<td>PHR-CC</td>
<td>4CT0\textsuperscript{59}</td>
<td></td>
</tr>
</tbody>
</table>

These crystal structures provide a snapshot of a protein/interacting proteins and hence reflect better insight into protein function and signaling mechanisms. For instance, the crystal structure of CRY complexed with FBXL3 has revealed the mechanism by which the C-terminal tail of the FBXL3 interacts with the CRY to initiate the degradation process\textsuperscript{40}. In combination with a crystal structure of the mCRY with its small molecule stabilizer (KL001), researchers were able to delineate the mechanism by which small molecules can modulate clock function by attenuating
the stability of CRY proteins. Namely, the carbazole scaffold mimics the isoalloxazine ring of the FAD to bind to the vacant FAD binding cleft in mammalian CRYs. In doing so, KL001 impersonates the FBXL3 tail thereby competitively binds to the CRY:FBXL3 interaction site and delays the degradation process\textsuperscript{41}. Thus, obtaining a high-resolution crystal structure not only assist in understanding the signaling mechanism but also aid in designing small therapeutic molecules\textsuperscript{8, 9, 21, 44}.

1.7 Resetting the clock: Chemical chronotherapy

Today’s modern lifestyle is impacting health and that ultimately is contributing to the misalignment of the circadian oscillator with our normal day-night cycle. Shift work, frequent travels to different time zones, and irregular eating habits contribute to misalignment of the internal clock with the normal day night-cycle and are increasing the risk of many physiological disorders like diabetes, sleep disorders, obesity, depression, and cancer\textsuperscript{6, 23, 50, 51}.

Recent research identifies significant health benefits that result from robust circadian rhythms. As a result, medical researchers now focus on how circadian rhythms may modulate our health or treatment practices. Initially, chronotherapy focused on identifying whether there were specific times of day where drug or stimulus administration achieved the optimal therapeutic index and pharmacokinetic profile. Interestingly, such approaches clearly indicated time-of-day specific administration of elements as simple as blue (morning) and red (evening) light positively impacted patients suffering from insomnia or depression. As a matter of fact, time-of-day specific light treatments have been used as a drug-free remedy for the seasonal affective disorder (SAD)\textsuperscript{42, 43}. Furthermore, extension of such time-of-day specific approaches to therapeutic intervention has had significant impact on cancer patients, as they showed fewer side effects when the anti-cancer drug gemcitabine was administered at 9 AM daily instead of 3 PM\textsuperscript{44}. Delightfully, identification
of the molecular architecture of the circadian rhythm proteins and their role in physiology has emerged a new era in the field of chronobiology. In particular, specific focus has been on targeting core clock proteins directly, to deal with the negative consequences of misalignment between endogenous circadian rhythms and the diurnal cycle.

Of late, small molecules are being exploited to manipulate the misaligned clock to prevent or alleviate the clock related diseased conditions. Additionally, small synthetic drug-like molecules bestow great potential to not only understand the function of the molecular clock system but also can be used to target the diseases associated with the clock dysfunction. The approach anticipates modulating the endogenous biological rhythm which can potentially “reset the clock” and hopefully can treat physiological or metabolic pathologies.

1.7.1 Circadian rhythm and drug intervention

Different clock elements (discussed earlier in the section 1.5) have been shown to impart pathological disorders in mutant mice; implying, each component of the mammalian circadian clock system has a crucial function. For instance, CLOCK mutant mice were found to be hyperglycemic and obese whereas BMAL1 deficit mice showed decreased body weight and infertile. Additional physiological defects resulting from clock genotypes are outlined in the table below.
Table 2: Summary of the diseases associated with the mutation of different core clock elements

<table>
<thead>
<tr>
<th>Mutant</th>
<th>Physiological Disorder</th>
</tr>
</thead>
<tbody>
<tr>
<td>CLOCK</td>
<td>Altered sleep pattern, hyperglycemia, diet induced obesity[6,49]</td>
</tr>
<tr>
<td>BMAL1</td>
<td>Decreased body weight, infertility, premature death[16]</td>
</tr>
<tr>
<td>CRY I and II</td>
<td>Alterations in liver regeneration, hyperglycemic conditions due to gluconeogenesis, altered sleep pattern, glucose intolerance[8,50,51,52]</td>
</tr>
<tr>
<td>PER 1 and 2</td>
<td>Cancer development, improper cell division, abnormal apoptosis[16]</td>
</tr>
<tr>
<td>REV ERB</td>
<td>Dyslipidemia[53]</td>
</tr>
<tr>
<td>CK1ε/δ</td>
<td>Familial advanced sleep phase syndrome[43,54]</td>
</tr>
</tbody>
</table>

1.7.2 Identification of the synthetic clock modulators

The circadian rhythm is highly pliable and can reset by external cues such as environmental change or metabolism; therefore, small molecules can be used to modulate the rhythm[13,44-48]. The purpose of drug discovery is to target either the input pathway or the feedback loop. Moreover, the primary goal is to target the core clock component instead of components far away from the central oscillator to avoid pleiotropic effects[13]. Incidentally, modern research has determined that all core proteins are druggable, and most of them have an endogenous ligand that drives their actions. For instance, REV-ERBα/β employs a Heme cofactor, while RORα has been found to be regulated by cholesterol. Alternatively, important dinucleotides such as flavin adenine dinucleotide (FAD) and nicotinamide adenine dinucleotide (NAD) have been found to be signaling molecules for core circadian proteins like cryptochrome (in plants and insects) and BMAL1[48] respectively. In plants
and insects, FAD governs the signaling mechanism of the CRY in a light-dependent manner. Despite retaining the FAD binding pocket, vertebrate CRYs are not photo regulated and do not bind FAD, but the binding pocket can be targeted to regulate clock function (e.g. KL001)\textsuperscript{48}. Additionally, NAD regulates the activity of the CLOCK BMAL1 complex through a feedback loop involving SIRT1. Fig 8 shows the structures of the endogenous ligand of the core clock proteins.

**Figure 8: Endogenous ligands of the important circadian proteins.** The structure of the ligands is used as a reference to synthesize small drug-like molecules to target clock proteins
Assuming that most of the clock components are under control of endogenous ligands, that provides the lead to synthesize small drug-like molecules which can mimic the natural metabolite and modulate the clock as well as treat diseases associated with the clock dysfunction. In particular small-molecule screens have been assayed for their effect on circadian period and amplitude, yielding identification of a host of clock-affecting small molecules targeting nearly all components of the core and peripheral clock elements. In addition to testing small molecule libraries, efforts are also being made to test FDA approved drugs that might target circadian proteins to reduce the cost of identifying new bioactive molecules. For instance, Tamai et al. tested approximately 1000 molecules from the FDA approved library and found approximately 5% of them had period shortening abilities\textsuperscript{56}.

Nevertheless, the burgeoning circadian modulator research has identified several new small molecules which act on different clock proteins. Fig 9 shows some of the potent circadian clock modulators synthesized and the part of the cycle/protein they act on. Additionally, Table 3 summarizes some of the drugs synthesized based on the structure of the endogenous ligand.
Figure 9: Circadian clock modulators: Some of the circadian clock modulators and their protein targets in the different stages of the negative feedback loop. Adapted from Ref:13
Table 3: Summary of the small molecules synthesized to mimic the endogenous ligand of the core circadian proteins and their therapeutic actions

<table>
<thead>
<tr>
<th>Clock proteins</th>
<th>Endogenous ligand</th>
<th>Small molecule modulators</th>
<th>Physiological effects</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cryptochrome (Cry 1/2)</td>
<td>FAD</td>
<td>KL001 and derivatives, KL044</td>
<td>Glucose tolerance in obese mice, Period-lengthening effects</td>
</tr>
<tr>
<td>REV-ERB</td>
<td>Heme</td>
<td>GSK4112</td>
<td>Inhibit gluconeogenesis</td>
</tr>
<tr>
<td>RORs</td>
<td>Cholesterol</td>
<td>Nobiletin</td>
<td>Improve homeostasis in diabetic mice</td>
</tr>
</tbody>
</table>

![Chemical structures](attachment:chemical_structures.png)
However, for the purpose of this dissertation, we will narrow down our focus to the core clock component CRY. Henceforth, we will discuss cryptochrome and its function in mammalian circadian rhythm, followed by drug intervention.

1.8 Cryptochrome as a target of drug discovery

1.8.1 Cryptochrome

Cryptochromes were first identified in *Arabidopsis thaliana* and were named “cryptochrome” as their function was “cryptic” and for their widespread abundance in cryptogamic plants\(^\text{61-63}\). They exhibit sequence homology with DNA photolyase and hence belong to the CPF family (Cryptochrome photolyase family). The founding member of the family, photolyase is found in most organisms, where it functions as a light-driven enzyme that repairs UV damaged DNA\(^\text{64}\). Initially, widespread efforts failed to identify a photolyase in placental mammals, the closest homologs were identified as the Cryptochrome family. Domain structures of CRYs demonstrates that it retains a photolyase homology region characteristic of the CPF family. In addition, it has a disordered cryptochrome C-terminal (CCT) or extension (CCE) equivalently\(^\text{26,40,61,64,65}\). Despite high sequence homology, cryptochromes do not participate in DNA repair activity. Instead, they sense blue light in plants and contribute to growth and development, whereas in mammals they work in a light-independent manner and are the part of the core circadian clock system\(^\text{64,66}\).
**Figure 10: Domain architecture of the mammalian cryptochrome:** Mammalian cryptochrome has a photolyase homology region along with a disordered C-terminal extension.

The animal cryptochrome was first discovered as a photolyase like protein with no DNA repair activity in animals in 1995\(^6\). The animal cryptochromes can be classified as CRY type I (CRY1) and CRY type II (CRY2), which are differentiated based on the presence and function of the FAD cofactor. While CRY1 proteins readily bind FAD and demonstrate blue light sensitivity to participate in light-activated circadian behavior in Drosophila, CRY2 proteins are currently believed to not bind FAD, but rather work in a light-independent manner in the mammalian circadian clock\(^6\). As discussed above, CRY along with period (PER) acts as a repressor of BMAL1 and CLOCK. However, the mechanism by which it interacts with the complex is still elusive. As noted, unlike other cryptochromes, which have been crystallized and the role of cofactor the flavin adenine dinucleotide (FAD) in the signaling mechanisms are well known\(^{67,68}\), the mammalian cryptochrome weakly binds FAD with an open cofactor pocket and is believed to function in an FAD independent manner, where the FAD binding pocket has been co-opted to function as a protein:protein interaction site\(^{21,28,40,72}\). Moreover, recent biochemical studies have shown *CRY* regulates gluconeogenesis in mammals and therefore, could be a potential target for treating type
II diabetes. Hence the development of a small molecule that can mimic FAD might not only assist in understanding the signaling mechanisms of CRY but could also can set the stage for drug development.

1.8.2 Cryptochrome regulation of gluconeogenesis

Gluconeogenesis:

Gluconeogenesis, by definition, is “generation of glucose from non-carbohydrate precursors.” It is a metabolic process and is stimulated by glucagon and epinephrine in the liver. It is the process by which human body maintains optimal blood glucose level and hence glucose homeostasis. During the daytime, or when the body is in its feeding cycle, glucose requirements are fulfilled by food intake. The body utilizes the required amount of sugar and converts the rest for storage in the form of fat. At night, or during fasting cycles, the pancreas secretes glucagon to activate hepatic gluconeogenesis process and avoid hypoglycemic condition.

**Figure 11: Glucose homeostasis:** During daytime or the feeding cycle, the body utilizes glucose supplied by food-intake and stores excess glucose in muscles and adipose tissue. On the contrary, at night time or the fasting cycle, glucagon gets secreted by the pancreas which activates the G-protein coupled receptors and thereby cAMP. The cAMP ultimately phosphorylates CREB, and that activates gluconeogenic genes to secrete glucose from non-carbohydrate sources.
At the fasting cycle, glucagon secreted by the pancreas binds to the G-protein coupled receptor (GPCR) or glucagon receptors in hepatocytes (liver cells), which in turn triggers intracellular cyclic AMP (cAMP) resulting in CREB (cAMP response element binding) phosphorylation and subsequent expression of gluconeogenic genes like pyruvate kinase (*Pck1*) and glucose-6-phosphatase (*G6pc*). Importantly, gluconeogenic genes are expressed rhythmically and are controlled in a circadian fashion\(^{51,69-71}\).

“High Glucose, No CRY”

Recent studies by Zhang et al. demonstrated that cryptochrome controls gluconeogenesis process rhythmically. By luciferase in-vivo imaging technique they demonstrated that when CRY is active it inhibits cyclic AMP production and hence modulates gluconeogenesis at the genomic level\(^{51,69-71}\).

To summarize, cryptochrome forms a “gate” for gluconeogenesis in humans. In the morning time (or during the active feeding cycle), high CRY levels inhibit the gluconeogenesis process. On the contrary, during nighttime (or fasting cycle), CRY levels are low, and gluconeogenesis proceeds normally, hence it maintains an optimal level of glucose in a healthy individual\(^{71}\).
Figure 12: The cryptochrome acts as a gate for the gluconeogenesis process: Cryptochrome is active during the feeding cycle and inhibits glucagon-stimulated cAMP production which further prohibits the gluconeogenesis process. At the nighttime or at the fasting cycle, cryptochrome becomes inactive and glucagon induces the gluconeogenesis process to take place.

1.8.3 Gluconeogenesis and Diabetes

Diabetes mellitus type 2 is a metabolic disorder characterized by hyperglycemia or high blood sugars as cells become insulin-insensitive and fail to take up glucose which leads to dysregulation of glucose homeostasis. Diabetes II is the seventh leading cause of death in the United States. According to the US department of health and human services (2014), approximately 29.1 million Americans have diabetes with 1.7 million diagnosed every year\textsuperscript{70}.

Hyperglycemic conditions can cause several other diseases such as heart failure, obesity, strokes, kidney failure, and blindness. Moreover, gluconeogenesis is a regular metabolic process
which can lead to persistent hyperglycemic conditions in diabetic patient. Therefore, designing drugs which can target the hepatic gluconeogenesis process can facilitate in maintenance of optimal blood glucose levels\textsuperscript{73}.

As discussed earlier, recent research affirmed that cryptochrome when rhythmically expressed could inhibit the hepatic gluconeogenesis process. As such, cryptochrome was identified as a suitable drug target. Specifically, if cryptochrome levels are modulated or stabilized through a small molecule, then production of glucose by the process of gluconeogenesis can be suppressed.

1.9 Cryptochrome modulator

Cryptochrome is a core protein of the mammalian circadian rhythm that controls many physiological processes. As discussed earlier in the chapter, gluconeogenesis is modulated when cryptochrome is expressed. The big vision is to synthesize a small molecule which can stabilize or delay the degradation process of the cryptochrome. A stabilized cryptochrome can delay the gluconeogenesis for long and thereby help to alleviate hyperglycemic conditions in the diabetic patient\textsuperscript{48,57,58}.

\textbf{Figure 13: Cryptochrome as a drug discovery target:} A stabilized cryptochrome would restrict gluconeogenesis process for a longer period which potentially aids in lowering the blood glucose level and hence would help the diabetic patient in maintaining an optimal level of glucose.
Hence, the goal of the small molecule intervention is to synthesize a scaffold which can stabilize the cryptochrome and delay its degradation process. However, to date, its action in the regulation of cryptochrome (CRY) is elusive. Notably, vertebrate CRYs (type 2 CRYs) are unique in that they do not bind FAD, but rather employ the FAD binding site to interact with FBXL3, to target the protein for degradation. Thus, it is plausible that identification of a small-molecule that could compete at the FAD binding site would allow both interrogation of the signaling pathway of CRYs, but also result in CRY stabilization through competition with FBXL3. Furthermore, such molecules could unravel cryptochrome’s mode of action in the sleep-wake cycle, diabetes and other related metabolic disorders. Remarkably, high throughput analysis of the effect of ~60,000 compounds on the circadian period in human osteosarcoma U2OS cell lines harboring Bmal-dLuc reporter showed three carbazole-containing scaffolds with dose-dependent period lengthening effects, consistent with what would be expected for a CRY-stabilizer. Among the three, KL001 showed pronounced period lengthening effect in a dose-dependent manner. KL001’s action in the modulation of the clock could suggest potential therapeutic effect towards diabetes and other metabolic diseases.48

1.9.1 KL001

![KL001](image)

**Figure 14: Structure of KL001** that could stabilize cryptochrome and delay its degradation process.
KL001 was found to be first of a kind of small-molecule that could stabilize CRY. Incidentally, KL001 indeed competitively bound to the SCF^{FBXL3} binding site (an F-box protein which tags cryptochrome for proteasomal degradation), hence prevents its ubiquitination process and stabilizes cryptochrome. Cryptochrome stabilization by KL001 ultimately lengthens the overall period of the cryptochrome, enabling the therapeutic potential^{40,41,57}.

**Figure 15: KL001 mode of action.** KL001 competitively binds at the FBXL3 binding site of the flavin and prevents it from degradation, which in turn lengthens the period.

Previous studies have shown that CRY2-FBXL3 complex involves insertion of the C-terminal tail of FBXL3 into the FAD binding pocket of the cryptochrome^{21,40,41}, hence competitively binding of KL001 in the FBXL3 pocket was postulated as the mechanism of KL001 induced stabilization of CRY. The postulation was further confirmed by the crystal structure of CRY2 PHR bound by KL001. Nangle et al. compared the crystal structure with that of FBXL3 bound CRY and confirmed that KL001 and FBXL3 compete for the same FAD binding pocket. Moreover, high resolution (1.94 Å) structures have unraveled the binding ability of the KL001 in the FAD-binding pocket, giving better insights for drug discovery^{41}.
1.10 Research Objective

This dissertation mainly focuses on designing the small molecules to target the core circadian protein cryptochrome. In this work, we have deeply investigated the crystal structure of the cryptochrome in complex with KL001 using modeling software such as Coot and Pymol. Insights gained from the structure-activity relationship, we have designed the “Zoltowski Library” consisting of 10 different small molecules. In the initial part of the research, we modified the scaffold and replaced the carbazole with Delta-carbolines, Pinoline, and Quinoxaline derivative to target the core scaffold for improved binding to the CRY FAD binding site.

Motivated by our previous results, we further investigated the crystal structure of KL001 bound cryptochrome and designed numerous side chains for better recognition and binding. Furthermore, we also focused on changing the linker region. All resulting compounds were tested for activity by interrogating their effect on circadian period in U2OS cells harboring the BMAL1-dLuc reporter. Although the research is in the primary stages of the circadian modulators development, we can still get better insights to proceed with the synthesis of more compounds.

In the latter part of this dissertation, we will steer the focus to study about LOV domain proteins of plant circadian clock. Specifically, we will focus on the study of the thermal kinetics of LKP2 protein of Arabidopsis thaliana. We also have extended our research to examine the thermal kinetics of LKP2 protein in an agriculturally important crop Brassica rapa.
REFERENCES


CHAPTER 2
SYNTHESIS OF CRYPTOCHROME MODULATOR

2.1 Introduction

Cryptochrome (CRY) is an integral part of the mammalian clock system participating in the autoregulatory core transcription-translation feedback loop\textsuperscript{1-3}. Recent studies have shown that CRY regulates the gluconeogenesis process and hence maintains glucose homeostasis\textsuperscript{4-6}. Cryptochrome is a circadian protein that is expressed rhythmically for its function. The anticipation was to synthesize a synthetic ligand which stabilizes CRY and prevents it from degradation and thereby modulates gluconeogenesis. Modulation of gluconeogenesis will help a diabetic patient with a sudden spike in the blood glucose level overnight.

Recently, a cell-based study showed a carbazole scaffold (KL001) was able to stabilize CRY proteins resulting in period lengthening effects. KL001’s action in stabilizing CRY suggested potential therapeutic effect towards diabetes and other metabolic diseases\textsuperscript{7}. Notably, the carbazole moiety was found to be essential for that period-changing activity, as it was reported that replacing it with a phenyl indole or a non-planar moiety resulted in complete loss.
of activity. Furthermore, its crystal structure indicated that KL001 is bound to the FAD binding pocket in the precise location of the isoalloxazine ring, supporting the requirement of a planar heterocycle at the core of CRY stabilizing drugs. Despite the fact that carbazole can mimic the isoalloxazine ring of the FAD, it lacks critical H-bonding elements that can facilitate efficient recognition and binding that could explain the loss of rhythm at higher KL001 concentrations\textsuperscript{8,9}.

Structure-activity relationships revealed the FAD-binding site is enriched with polar residues that bind to polar elements within the isoalloxazine ring of FAD. Therefore, we postulated that incorporating heteroatoms into the carbazole scaffold could increase the propensity of hydrogen bonding and facilitate recognition and binding of a small molecule\textsuperscript{9,11}.

The primary purpose of this research is to introduce H-bonding characteristics to the carbazole scaffold for efficient binding at the FAD-binding pocket\textsuperscript{7}. In this research, to better understand the functional differences of carbazole and other scaffolds, we sought to prepare analogs varying in the nature of core scaffold and developed the “Zoltowski library.” In this process, we have identified a simple two steps reaction that allows robust formation of delta-carbolines via benzyne synthesis.

2.2 Structure-activity relationship of KL001

The crystal structure of CRY with bound KL001 demonstrated that the carbazole ring of KL001 buries itself in the inner cleft of the flavin adenine dinucleotide binding pocket in a manner similar to the isoalloxazine ring of FAD. Notably, the isoalloxazine of FAD is subdivided into three types of rings: xylene, pyrazine and pyrimidine rings. The pyrimidine ring of the FAD participates in the critical hydrogen bonding with the polar residues present in the binding pocket. A detailed examination of the crystal structure revealed, carbazole only mimics outer xylene and central pyrazine moieties of the isoalloxazine ring and not the complete tricyclic system. Most
importantly, carbazole lacks the moieties which can mimic the pyrimidine ring that participate in critical hydrogen bonding\textsuperscript{9,11}. (Fig 16)

\textbf{Figure 16: Comparison of KL001 and FAD:} (A): KL001 have carbazole as a scaffold (B): Flavin adenine dinucleotide (FAD); the isoalloxazine ring of FAD is subdivided into xylene, pyrazine and pyrimidine rings (C) Carbazole only mimics xylene and pyrazine rings of the isoalloxazine ring and lacks substituents which can participate in hydrogen bonding with the neighboring polar residues in the FAD binding pocket.

Moreover, a careful examination of the flavin binding site shows the presence of polar residues like arginine (R 376) aspartic acid (D 405) and asparagine (N 411), which snuggle close to the isoalloxazine ring and participate in significant interactions required for efficient binding.
Notably, previous studies have shown that point mutations of aspartic acid (D405) and asparagine (N411) are capable of abolishing the interactions between KL001 and CRY2 binding pocket\textsuperscript{7,9,10}. (Fig 17, PDB: 4mlp).

**Figure 17: KL001 in a FAD binding pocket:** The FAD binding site are enriched with polar residues. Remarkably, a polar residue arginine 376 (R 376) is just in the vicinity of the carbazole ring. Besides, another set of polar residues asparagine (N 411) and aspartic acid (D 405) are essential for the KL001 binding.

All these observations inspired us into incorporating either an extra ring or a polar substituent to the scaffold. That might better mimic the size of FAD, filling the FAD binding pocket and potentially increasing binding efficiency. Furthermore, introducing substituents with
hydrogen bonding capabilities might increase the binding efficiency through interactions with essential polar residues in the binding site.

2.3 Synthesis of the Zoltowski Library (ZL series)

Based on insights gained from the structure-activity relationship of KL001, compounds ZL001 to ZL005 were synthesized using different scaffolds while keeping the side chain the same as KL001. Our overall design principle of the scaffolds was to either attach a polar side chain or an extra ring so that the complete isoalloxazine ring can be mimicked.

We first synthesized KL001 side chain in bulk to both synthesize KL001 in the lab as a reference molecule for biological evaluation and for further attachment of the ZL scaffolds.

2.3.1 Synthesis of KL001

We synthesized KL001 based on the published synthesis. The side chain was synthesized by mesylating the furfural amine 1 using mesitylene chloride and pyridine, where pyridine acted both as solvent and a base. The mesylated product 2 was then N-alkylated with epibromohydrin in the presence of sodium hydride to get compound 3. The coupling reaction of the side chain with the secondary amine of carbazole in the presence of sodium hydride yielded KL001 4 (yield 60%). The same procedure of side chain synthesis is used for most of the ZL library.
2.4. Design principle for ZL001

For the first compound in the library, we did not want to change the scaffold drastically, but wanted to introduce an H-bonding moiety. To achieve this, we focused on the carboline scaffold. Carbolines are the N-heterocyclic compounds found in many biologically active natural and synthetic compounds. Most importantly, β-carbolines are widely distributed in nature and have shown DNA intercalation properties to inhibit CDK (cyclin-dependent kinase), topoisomerase and monoamine oxidase. Moreover, it interacts with 5-hydroxyserotonin receptors\textsuperscript{12}.

Essentially, carbolines are pyrido-indoles, and based on the position of nitrogen with respect to the indole ring they are named as alpha (α), beta (β), gamma (γ) and delta (δ) carbolines.
Amongst all carbolines, \( \delta \)-carboline is extremely rare in nature. Interestingly, benzo fused \( \delta \)-Carboline like cryptolepine and quindoline are found to be pharmaceutically active compounds and have shown antimalarial, antiplasmodial and antibacterial activity. Cryptolepine is an indoloquinoline natural product extracted from the roost of *Cryptolepis triangularis* used as a traditional antimalarial drug in Central and West Africa for ages\(^{13-15}\).

We envisioned, \( \delta \)-carboline being similar to carbazoles will bind in a similar fashion to KL001, with the introduction of an H-bonding atom due to a nitrogen at the para position. An extra nitrogen can potentially increase the scaffold’s propensity to participate in critical hydrogen bonding with the polar residues in the vicinity such as arginine (R 376) (fig 18). Furthermore, it can also interact with other polar “essential” residue asparagine N 411.
**Figure 18: ZL001 at the binding site:** In ZL001, an extra nitrogen present at the para position possibly can enhance the binding efficiency of the compound at the flavin binding site. Presumably, it can participate in H-bonding interactions with R 376, N 411. (The fig was made using PDB = 4MLP of KL001 complexed with CRY, the nitrogen was introduced to promote understanding)

Delta carboline was primarily synthesized following a known synthesis using aniline as a starting material. At first, the amino functionality of aniline was protected with pivaloyl chloride in the presence of triethylamine as a base yielding pivaloyl amino benzene 5. Amidation served two purposes, it not only acted as a protecting group but also provided neighboring group assistance during borination at the next step. Subsequent metalation of the compound 5 using ⁴BuLi in anhydrous THF followed by the treatment with trimethylborate and subsequent hydrolysis
afforded the corresponding boronic acid derivative 5. The product 6 was coupled with 3-fluoro-2-iodobenzene 7 using Suzuki coupling conditions for 48 hours. The coupled product 8 was deprotected in 20% of sulfuric acid solution under refluxing conditions to afford the coupled product 9. Cyclization of the resulted compound using pyridine-hydrochloride at 215 °C followed by the treatment with ammonium hydroxide afforded δ-carboline 10 in moderate yield\textsuperscript{16,17}.

3-Fluoro-2-iodopyridine used to couple with the compound 6 was also synthesized based on a patented synthesis. Iodination of 3-fluoro-2-chloropyridine was achieved using sodium iodide in the presence of trimethylsilane under refluxing conditions\textsuperscript{18}.

Scheme 2: Synthesis of Delta-carboline (45%)
The above synthesis of delta-carboline involved several steps under stringent conditions demanding longer reaction times with a moderate yield requiring purification in each step. Delta carbazoles are essential bioactive molecules. So far, the development of an efficient cyclization step was the focus of attention, and not much had been done to modify the coupling reaction. Consequently, an efficient approach was needed to synthesize δ-carboline efficiently.

2.4.1 Synthesis of δ-Carbazole using Benzyne Chemistry

We were intrigued to explore the use of aryne chemistry in the preparation of carbazoles due to its feasibility in incorporating heterocyclic rings in various natural product syntheses. Aryne chemistry has proved useful and efficient in the synthesis of complex bioactive heterocyclic molecules. Most importantly, aryne chemistry provides a strategic advantage in functionalizing an aromatic ring system, because of its feasibility to form multiple carbon-carbon or a carbon-heteroatom bond in a single step in a regioselective manner.19

To synthesize delta carboline through aryne chemistry, we did a retrosynthetic analysis and found two plausible approaches: the benzyne approach or the pyridyne approach. We possibly can synthesize either benzyne or 2,3-pyridyne precursors which could be coupled with halo-substituted aminopyridine or phenylamine respectively to yield the desired product.

![Pyridyne Approach](image1.png) ![Benzyne Approach](image2.png)

**Figure 19: Retrosynthetic analysis** shows two possible ways delta carbazole could be synthesized using aryne chemistry
2.4.1a. A 2,3-pyridyne approach for delta carboline synthesis

Pyridyne approach was captivating as the precursor can be coupled with the substituted phenylamine, giving an opportunity to synthesize a library of substituted delta carboline with just one precursor.

![Diagram of pyridyne approach for delta carboline synthesis](image)

**Figure 20: 2,3-Pyridyne approach for the synthesis of delta carbolines:** (A) pyridyne precursor can be coupled with the halo benzyne, the coupled product can be cyclized to yield delta carbolines. (B): Regioselectivity of pyridyne precursor at the 2nd position.

Regrettably, 2,3-pyridynes are highly regioselective; therefore, nucleophilic substitution exclusively takes place at the 2nd position. Incidentally, nucleophilic substitution at the 3rd position builds up negative charge on the 2nd position that creates repulsion due to the lone pair of electrons on nitrogen\textsuperscript{20,21}. Delightfully, the approach gave the alpha carboline precursor instead, which will be discussed in the later part of this research (ZL002).
2.4.1b. A benzyne approach for delta carbol ine synthesis

It occurred to us that the reaction of o-benzyne with 2-amino-3-halo-pyridine derivative should give a (2-halo-pyridin-3yl)-phenyl-amine intermediate that upon intramolecular cyclization would give δ-carboline directly. Most importantly, substituted δ-carbazole could also be synthesized by using same benzyne precursor with substituted pyridines. On the contrary, a different substituent benzyne precursor can also be generated.

![Image of benzyne approach]

**Figure 21: Benzyne approach:** An in-situ generated benzyne will be coupled with the the ortho-aminopyridine derivation, subsequent cyclization yield delta carbol ine

Similar to pyridynes, benzyne reaction also pose some restrictions due to the regioselectivity effect of different substituents present in the benzyne ring system. Incidentally, substituents present in the benzyne dictate the fate of the product; for instance, benzyne precursor with the electron donating substituent directs nucleophilic substitution regioselectivity at the meta position. Whereas, an electron withdrawing substituents facilitate it at the ortho position. On the contrary, a methyl substituent would yield a mixture of products. This restricted us to synthesize a library of benzyne precursors in the lab\textsuperscript{20,21}. 
Several methods for generating o-benzyne were explored that would react with an amino-pyridine derivative to give a coupled product. Many of the benzyne generation procedures required harsh conditions, use of strong bases such as butyllithium ("BuLi) or high temperature. Kobayashi’s method of benzyne generation is a milder approach known to generate benzyne in-situ, and remarkably, the method had been used to generate several natural products. Ortho-trimethylsilylphenyl triflate is a milder approach to generate benzyne which requires a fluoride source which acts as a desilylating agent, which subsequently eliminates the bulky triflate group leading to the formation of benzyne in-situ. Any simple fluoride source like cesium fluoride, tetra-n-butylammonium fluoride can be used at room temperature under an argon atmosphere with acetonitrile as a solvent.

Figure 22: A mild benzyne precursor: o-Trimethylsilylphenyl triflate: A fluoride source can abstract the trimethylsilyl group, which subsequently eliminates the triflate group leading benzyne generation.
The benzyne precursor (o-trimethylsilylphenyl triflate, 13) was synthesized using recently reported synthesis without further optimization. The procedure was highly efficient requiring only one purification step with approximately 66% of overall yield.

Phenol was treated with triisopropyl isocyanate in the presence of catalytic amount of triethylamine (Et$_3$N) to yield the corresponding carbamate 11. Ortho-metalation of the carbamate 11 using $^n$BuLi in anhydrous diethyl ether followed by its silylation and acidic work up afforded trimethylsilyl compound 12. Silyl triflate (benzyne precursor 13) was then conveniently formed in a one pot carbamate cleavage/ triflation of 12. The carbamate cleavage was achieved by heating a solution of 12 in anhydrous acetonitrile with 1-8-diazabicyclo [5.4.0] undec-7-ene (DBU) and diethyl amine. Thereafter, N-phenyl-bis-(trifluoromethanesulfonimide) (PhNTf$_2$) was used as a triflate source. The compound was then purified using flash chromatography$^{23}$. 

Scheme 3: Synthesis of benzyne precursor (66%)

The coupled product 14 was obtained by allowing benzyne precursor to react with a halo substituted amino pyridine in anhydrous acetonitrile under argon atmosphere at room temperature.
Cesium fluoride was used as a fluoride source to abstract the silyl group to generate benzyne in situ\textsuperscript{24}. The coupled product obtained was conveniently microwaved for 10 min at 180 °C with a palladium catalyst and sodium acetate in N, N-dimethyl acetamide to obtain the δ-carboline 10 in approximately 70% yield\textsuperscript{25}.

![Scheme 4: Synthesis of Delta-Carboline from benzyne precursor (70%)](image_url)

The δ-carboline 10 was coupled with the side chain 3 in presence of sodium hydride and DMF to achieve ZL001.

![Scheme 5: Synthesis of ZL001 (57%)](image_url)

Motivated with our success with the aryne chemistry, we used pyridyne precursor to synthesize the second compound in the library, while the coupled product of the delta carboline
was used for the third compound in the library. ZL002 and ZL003 were designed to gauge if planarity of a scaffold is required for the efficient binding in the FAD binding pocket.

2.5 Design Principle of ZL002 and ZL003

The design principle for the next two compounds involves the use of alpha carboline and delta carboline precursors in place of carbazole.

![ZL002](image)

![ZL003](image)

We anticipated that breaking the central pyrrole ring would allow free rotations of the rings that could potentially improve binding capacity by enhancing interactions with other residues. Furthermore, it would allow us to understand if planarity is one of the criteria essential for the activity of small molecules.

2.5.1. Synthesis of ZL002

As mentioned earlier in the chapter, 2,3-pyridynes are highly regioselective, and all the nucleophilic reactions are directed towards the 2\textsuperscript{nd} position (ortho-position). By taking advantage of this behavior, we synthesized the coupled product and then linked it with the KL001 side chain. We later planned to cyclize the coupled product to generate an alpha carboline, which could potentially be used as another scaffold for the library. Regrettably, all our efforts to cyclize the coupled product were futile.
Just like benzyne precursor, a mild pyridyne precursor was synthesized based on the published synthesis with minor modifications.

![Scheme 6: Synthesis of ZL002 (25%)](image)

The pyridyne precursor 17 was synthesized by silylating the 2-hydroxy pyridine using lithium diisopropyl amide as a base to yield silylated derivative 16. The compound 16 was then treated with triflic anhydride in the presence of 2,4-lutidine in DCM to achieve pyridyne precursor 17\(^{26}\). Pyridyne was then generated in situ using cesium fluoride as a fluoride source in anhydrous acetonitrile, and then allowed to react with aniline to yield the coupled product 18, which was then attached with the side chain to achieve ZL002 19.

2.5.2. Synthesis of ZL003

For the synthesis of the third compound in the library we used delta carboline precursor as a scaffold. Here, along with the effect of ring movement, we anticipated observing the effect of
halogen in the binding site. We expected the chloro group might help in facilitating binding by van der waals interactions in the flavin binding pocket.

The benzyne precursor 13 was coupled with 2-chloro-3-amino pyridine in presence of cesium fluoride. Side chain 3 was attached to the coupled product 14 by generating a nitrogen anion in presence of sodium hydride, then reaction undergoes ring opening step when heated with side chain 3 at 50 °C.

So far, subtle changes were made in the carbazole ring system for the synthesis of a novel KL001 analog. It intrigued us to observe the effect if the carbazole is replaced by another known tricyclic ring system. Particularly, if the tricyclic system could mimic the structure of natural endogenous metabolite which are known to work in the circadian fashion. The design principle of the fourth compound in the library relies on using a scaffold similar in the structure of the neurotransmitters like serotonin and melatonin. Pinoline is tetrahydro-β-carboline as a scaffold and is structurally similar to both serotonin and melatonin and is a natural biological metabolite of melatonin. Thus, if it demonstrated physiological effects on circadian period, it could be an endogenous ligand of CRY that modulates activity30.
Design Principle of ZL004

The design principle for ZL004 involved substitution of carbazole with pinoline, which is a natural metabolite. Pinoline is tetrahydro-β-carboline and is suggested to be produced in the pineal gland during the production of melatonin and is believed to follow the circadian rhythm\textsuperscript{27-30}. Hence, the anticipation was the scaffold would potentially assist in maintaining the circadian rhythm. Besides this, we anticipated that the methoxy group present in the structure could potentially participate in much stronger hydrogen bonding with the polar residues in the vicinity. Furthermore, the NH group present in the cyclohexane ring could also enhance the binding
capability of the molecule. Pinoline was synthesized by the cyclization of 5-methoxy tryptamine by reported synthesis\textsuperscript{30,31}.

![Figure 24: Structure-activity of ZL004.](image)

**Figure 24: Structure-activity of ZL004.** Pinoline as a scaffold will place two substituents with the capability of participating in hydrogen bonding interactions on both sides of the tricyclic ring system. The methoxy group, presumably fill the portion of the pyrimidine ring of the isoalloxazine ring system. While on the other side, the hydrogen of secondary nitrogen can participate in critical interactions with the neighboring residues. The figure was generated using PDB = 4MLP for understanding purposes.
5-Methoxy tryptamine was heated with paraformaldehyde in 10:1 mixture of acetic acid and methanol to get pinoline 21. The primary amine of the pinoline was then selectively protected with Boc anhydride in THF, the protected scaffold 22 was attached to the side chain 3 of KL001 using sodium hydride in DMF. The Boc-protected amine 23 was then deprotected using trifluoroacetic acid in DCM to get ZL004 2430,31.
2.7 Design Principle of ZL005

For the fifth compound in the library, benz-quinoxaline derivative was chosen for substituting carbazole. The quinoxaline derivative was anticipated to mimic the FAD efficiently than carbazole in KL001. We envisioned, presence of one extra ring could potentially enhance pi-pi stacking and non-covalent interactions between the quinoxaline rings with histidine H 373 present in the vicinity. Additionally, it can participate in salt bridge formation with asp (D 405) and arginine (R 376). Furthermore, just like isoalloxazine ring system, quinoxaline has a pyrazine like moiety that could potentially participate in hydrogen bonding with adjacent arginine R 376.
The scaffold was conveniently prepared by the condensation reaction of isatin and o-phenylenediamine. A solution of isatin and ortho-phenylenediamine in acetic acid was microwaved for 10 min at 150 °C to get the desired compound $25^{32-33}$. The scaffold such obtained was then coupled with the side chain in the presence of sodium hydride in DMF at 60 °C to obtain ZL005 $26$. 

Figure 25: Structure-activity of ZL005: Benzquinoxaline will physically add one more ring in the tricyclic ring system, which might potentially eliminate the requirement of having a substituent in a tricyclic ring system.
2.8 Materials and methods

All the reactions were performed in dried glassware under an argon atmosphere. All reagents were purchased from Sigma Aldrich (St. Louis, MO), Fisher scientific (Waltham, MA), Alfa Aesar (Ward Hill, MA), Chem Impex (Wood Dale, IL) and used as it was received. The deuterated solvents (CDCl$_3$, DMSO-d$_6$) were purchased from Cambridge Isotope Laboratories (Cambridge, MA). $^1$H-NMR and $^{13}$C NMR spectra were recorded using 500 MHz Jeol multinuclear NMR spectrometer in the department of chemistry at Southern Methodist University. All chemical shifts are reported in the standard notation of parts per million using the peak of residual proton signals of the deuterated solvent as an internal reference. The coupling constant units are in Hertz (Hz) splitting patterns are indicated as follows: br, broad; s, singlet; d, doublet; t, triplet; q, quartet; m, multiplet; dd, doublet of doublets; dt, doublet of triplets. Microwave reactions were carried out in CEM Discover$^{\text{TM}}$. Column chromatography was carried out on Combi-Flash instrument by using pre-packed silica gel columns.
N-(Furan-2-ylmethyl)-methane sulfonamide (2)

Furfuryl amine (20 g, 20.5 mmol, 1 eq.) was dissolved in 150 mL of pyridine under an argon atmosphere at room temperature. The reaction mixture was cooled to 0 °C before methanesulfonyl chloride (24 mL, 30.8 mmol, 1.5 eq.) was introduced dropwise. The reaction was then allowed to stir at the room temperature overnight. The reaction mixture was quenched with water, the aqueous layer was extracted with DCM thrice. The combined organic layer was washed successively with 5% hydrochloric acid followed by water, brine, dried over sodium sulfate, concentrated under vacuum to yield dark brown liquid (32 g, 99%). The compound was confirmed by $^1$H-NMR and used for the next step without further purification. $^1$H NMR (500 MHz, CDCl$_3$) δ 7.31 (s, 1H), δ 6.32 (d, $J$ =1.75 Hz, 1H), δ 5.03 (br, 1H), δ 4.31(d, $J$ = 5.7 Hz, 2H), δ 2.82 (s, 3H).

N-(Furan-2-ylmethyl)-N-(oxiran-2-ylmethyl)-methane sulfonamide (3)

N-(Furan-2-ylmethyl)-methane sulfonamide (20 g, 114 mmol, 1 eq.) was dissolved in N, N- formamide (150 mL) at the room temperature under an argon atmosphere. The reaction mixture was cooled down to 0°C -5 °C followed by addition of sodium hydride (60% dispersion in mineral oil, 6.80 g, 171 mmol, 1.5 eq.) in portions over the period of 15 min, while it was still stirring over
an ice bath. The reaction mixture was allowed to warm up to the room temperature for 30 - 45 min. It was again cooled down to 0 °C - 5 °C and added epibromohydrin (12 mL, 171 mmol, 1 eq.) dropwise. The reaction flask was then equipped with condenser and the reaction was heated at 50 °C overnight under an argon atmosphere. The reaction was quenched by adding water and extracted with ethyl acetate thrice. Combined organic layer was washed with brine, dried over sodium sulfate. Concentration under vacuum afforded 45 g (quant.) brown liquid.

\[
\text{KL001: (4)}
\]

\[
\text{N-(3-(9H-carbazol-9-yl)-2-hydroxypropyl)-N-(furan-2-ylmethyl)-methanesulfonamide}
\]

The carbazole (20 g, 86 mmol, 1 eq.) was dissolved in anhydrous 150 mL of N, N-dimethylformamide under an argon atmosphere at the room temperature. The reaction mixture was further cooled down to 0 °C - 5 °C for 15 min. To the cooled reaction mixture, added sodium hydride (60% dispersion in mineral oil, 5.80 g, 129 mmol, 1.5 eq.) over the period of 15 min. After complete addition, the reaction mixture was allowed to stir at room temperature for 30 min before cooling it down again to 0 -5 °C. N-(furan-2-ylmethyl)-N-(oxiran-2-ylmethyl)-methanesulfonamide dropwise was then introduced dropwise while the reaction mixture was stirring over an ice bath. After complete addition, the reaction flask was equipped with a condenser and heated at 50 °C overnight under argon atmosphere. The work up was done by quenching the reaction with water and extracting it with ethyl acetate thrice. The combined organic layer was washed with
water followed by the brine solution, dried over sodium sulfate and concentrated under vacuum. The crude obtained was then purified by flash chromatography and sticky solid was obtained. The sticky solid was crystalized by dissolving in dichloromethane and then adding n-hexane drop wise till white crystalline solid settled down in the round bottom flask. The solid was then filtered, washed with hexane and dried under vacuum. 7.0 gm of white solid obtained was confirmed by $^1$H and $^{13}$C-NMR. $^1$H-NMR (DMSO-d$_6$) δ 8.11 (d, $J$ = 7.45 Hz, 1H) δ 7.52 (d, $J$ = 8.6 Hz, 3H), δ 7.43 (t, $J$ = 1.15 Hz, 2H) δ 7.19 (t, $J$ = 8 Hz, 2H) δ 6.28 (t, $J$ = 1.7, 1 H) δ 5.38 (1 H, br, OH, s) δ 4.48 (m, 2H) δ 4.43 (m, 1H), δ 4.19 (m, 1H) δ 4.10 (s, br., 1H) δ 3.35 (m, 1H) δ 3.20 (m, 1H) δ 2.88 (s, 3H); $^{13}$C NMR (DMSO-d$_6$, 500 MHz) δ 150, 143.75, 140.97, 126.12, 122.65, 120.6, 119.2, 110.5, 110.0, 68.5, 60.4, 47.4, 45.0, 39.9, 38.2.

![N-phenylpivalamide (5)](image-url)

**N-phenylpivalamide (5)**

The aniline (5 g, 53.7 mmol, 1 eq.) was dissolved in 50 mL dichloromethane at the room temperature under an argon atmosphere. The reaction mixture was then cooled down to 0 °C-5 °C for 10 min. To the cooled reaction mixture added triethylamine (30 mL, 215 mmol, 4 eq.) followed by the addition of pivaloyl chloride (9.9 mL, 805 mmol, 15 eq.) drop wise. The reaction was then continued to stir at 0 °C-5 °C for 15-20 min, then allowed to stir at the room temperature overnight. The workup was done by concentrating the reaction mixture under vacuum followed by addition of 200ml of water. The aqueous layer was then extracted with dichloromethane thrice. The combined organic layer was washed with 5% hydrochloric acid followed by the saturated solution
of sodium bicarbonate and the brine solution. It was then concentrated under vacuum to yield reddish brown solid which was purified by the flash chromatography to get 9.1 g of white solid. 

\(^1\)H NMR (CDCl\(_3\), 500 MHz) \(\delta\) 7.52 (d, \(J = 7.4\) Hz, 2H) \(\delta\) 7.30 (t, \(J = 2.3\) Hz, 2H) \(\delta\) 7.2 (t, \(J = 1.15\) Hz, 1H) \(\delta\) 1.31 (s, 9H).

\(\text{6} \)

2-(pivaloylamino)-benzenboronic acid (6)

N-phenylpivalamide (5 g, 25 mmol, 1 eq.) was dissolved in 50 mL of anhydrous THF under an argon atmosphere at the room temperature. The reaction mixture was then allowed to stir over an ice bath followed by dropwise addition of butyl lithium (15 mL, 75 mmol, 3 eq.). After complete addition, the reaction mixture was allowed to warm up to the room temperature gradually. The reaction mixture was continued to stir at room temperature for 5 hours. After 5 hours, trimethylborate was introduced to the reaction mixture drop wise at 0 - 5 °C. After complete addition, the reaction mixture was continued to stir at 10 °C for 2 hours. The reaction mixture was then quenched by adding 5% hydrochloric acid till its pH became acidic. The aqueous layer was then extracted with dichloromethane thrice. The combined organic layer was then washed with water, dried over sodium sulfate and concentrated under vacuum. The coffee brown fluffy solid obtained was dissolved into the minimum quantity of acetone followed by the addition of water dropwise which had crystallized the white solid. The solid such obtained was washed with water and filtered. 0.7 gm of pure white solid was obtained, verified by \(^1\)H NMR. \(^1\)H NMR (DMSO-d\(_6\)
Alternative synthesis of borate derivative using 2-Bromo-benzylpivaloyl amide (97 mmol) was dissolved in 250 mL of anhydrous tetrahydrofuran under argon atmosphere. n-butyl lithium (234 mmol) was added drop wise after cooling reaction mixture to -65 to -70 °C (temperature was maintained using acetone dry ice mixture). Reaction was continued to stir at -65 to -70 °C for 1 hour. After 1-hour, triisopropyl borate (250 mmol) was introduced to the reaction mixture drop wise over the period of 30 minutes, while it was still stirring at -65 to -70 °C. The reaction mixture was then continued to stir at -70 °C for 4 hours. After 4 hours reaction was quenched by adding hexane and allowed to warm up to 0-5 °C. Continued stirring for 2 hours lead to generation of solid. The solid such obtained was filtered and washed with saturated solution of ammonium chloride. The aqueous layer was then extracted with DCM. Combined organic layer was dried over sodium sulfate and concentrated. The coffee brown fluffy solid obtained was dissolved in a minimum volume of acetone followed by water, which crystallized the white solid. The solid such obtained was washed filtered and washed with water. 4 gm. of pure white solid was obtained, verified by $^1$H NMR. $^1$H NMR (DMSO-d$_6$ 500 MHz) δ 11.14 (s, br, OH), δ 7.67 (d, $J = 8.25$ Hz, 1H) δ 7.57 (d, $J = 7.45$ Hz, 2H) δ 7.20 (t, $J = 6.9$ Hz).
2-Iodo-3-fluoropyridine (7)

3-Fluoro-pyridine (5 g, 38.01 mmol, 1 eq.) was dissolved in 75 mL of 1, 4-dioxane at the room temperature under an argon atmosphere. To the solution added chlorotrimethylsilane (9.5 mL 76.0 mmol, 2 eq.) followed by the addition of sodium iodide (27.5 gm, 190 mmol, 5 eq.). After addition, the reaction mixture was allowed to reflux at 80 °C overnight under an argon atmosphere. The work up was done by quenching the reaction mixture with water and extracting the aqueous layer with ethyl acetate. The combined organic layer was washed with water followed by the brine solution, dried over sodium sulfate and concentrated under vacuum. From this approximately 7 gm of crude compound was obtained. Confirmed by $^1$H NMR. It was used for further reaction without any further purification. $^1$HNMR (CDCl$_3$, 500 MHz) $\delta$ (ppm) $\delta$ 8.21 (m, 1H) $\delta$ 7.30 (m, 1H) $\delta$ 7.24 (m, 1H).
2’-Fluoro[1,1’-biphenyl]-2-pivaloylamine (8)

2-(pivaloylamino-benzeneboronic acid (0.442 g, 20 mmol, 1 eq.) was added to the mixture of 2 M potassium carbonate and ethanol at the room temperature. To this mixture 2-Iodo-3-fluoropyridine (0.445 g, 20 mmol, 1 eq.) dissolved in deoxygenated toluene was introduced. The resulting mixture was allowed to stir at room temperature under argon atmosphere for about 1 hour. Tetrakis-triphenyl phosphorus palladium (0.07 g, 30 mmol, 1.5 eq.) was added to the reaction mixture after one hour of stirring at room temperature. The reaction mixture was then allowed to reflux at 100 °C for 48 hours. The work up was done by filtering out the solid formed in the reaction mixture and drying the organic layer over sodium sulfate followed by concentrating down the organic layer under vacuum. The crude such obtained was then purified by flash chromatography and then was characterized by $^1$H NMR. $^1$H NMR (CDCl$_3$, 500 MHz) $\delta$ (ppm) $\delta$ 8.44 (dd, $J = 16$ Hz, $J = 9.2$ Hz, 2 H) $\delta$ 7.66 (m, 1 H) $\delta$ 7.58 (m,1 H) $\delta$ 7.41 (t, $J = 6.85$ Hz, 1H) $\delta$ 7.33 (m 1 H) $\delta$ 7.15 (t, $J = 1.15$ Hz, 1H) $\delta$ 1.23 (s, 9H).
2'-Fluoro[1,1'-biphenyl]-2-amine (9)

2'-Fluoro[1,1'-biphenyl]-2-pivaloylamine was refluxed (100 °C) with 20% sulfuric acid overnight or until the reaction goes to completion. Adding 2 N sodium hydroxide till the pH became basic quenched the reaction. The aqueous layer was then extracted with dichloromethane. The combined organic layer was washed with the brine solution, dried over sodium sulfate, concentrated under the vacuum. The compound was used for the next step without any further purification. ¹H NMR (CDCl₃, 500 MHz) δ (ppm) δ 8.44 (m, 2 H), δ 7.48 (m, 2 H), δ 7.22 (m, 3H), δ 6.8 (m, 2H).

Delta-Carboline (10)

2'-Fluoro[1,1'-biphenyl]-2-amine (1.5 g, 79.6 mmol, 1 eq.) and pyridinium hydrochloride (74 g, 640 mmol, 8 eq.) was refluxed at 210-215 °C for 20-30 minutes. The hot reaction mixture was then poured in ice cold ammonia solution and stirred vigorously. The aqueous layer was then extracted with ethyl acetate. Combined organic layer was then dried over sodium sulfate and
concentrated under vacuum. The crude such obtained was purified with flash chromatography. 160 mg of pure compound was obtained, which was further characterized by $^1$H NMR. MP = 210 °C – 216 °C. $^1$H NMR (500 MHz, DMSO-d$_6$) δ 11.39 (s, 1H, NH), δ 8.4 (dd, 1H, $J$ = 1.15 Hz, $J$ = 3.45 Hz, Py-H), δ 8.16 (d, 1H, $J$ = 7.45 Hz, Py-H), δ 7.8 (dd, 1H, $J$ = 1.15 Hz, $J$ = 6.85 Hz, Py-H), δ 7.53 (d, 1H, $J$ = 8.6 Hz, Ar-H), δ 7.45 (td, 1H), δ 7.35 (dd, 1H, $J$ = 4.6 Hz, $J$ = 4.5 Hz, Ar-H), δ 7.21 (t, 1H, $J$ = 7.5, Ar-H).

\[
\text{N-(furan-2-ylmethyl)-N-(2-hydroxy-3-(5H-pyrido[3,2-b]-indol-5-yl)-propyl)-methane sulfonamide: (ZL001) (15)}
\]

δ-Carbazole (0.160 g, 0.95 mmol, 1 eq.) was dissolved in anhydrous 150 mL of N, N-dimethylformamide under argon atmosphere at room temperature. The reaction mixture was further cooled down to 0-5 °C for 15 min. To the cooled reaction mixture, added sodium hydride (60% dispersion in mineral oil, 1.4 mmol, 1.5 eq.) in batches. After addition reaction mixture was allowed to stir at room temperature for 30 min before cooling it down again to 0 - 5 °C. N-(Furan-2-ylmethyl)-N-(oxiran-2-ylmethyl)-methane sulfonamide (0.219 g, 0.95 mmol, 1 eq.) was then introduced to the reaction mixture drop wise while it was stirring over an ice bath. After complete addition, the reaction flask was equipped with a condenser and heated at 50 °C overnight under an argon atmosphere. The work up was done by quenching the reaction with water and extracting it
with ethyl acetate. The combined organic layer was washed with water and the brine solution, dried over sodium sulfate and concentrated under vacuum. The crude such obtained was crystallized by dissolving it in dichloromethane and ethyl acetate mixture and then by adding hexane drop wise. White solid such obtained was filtered and further purified by flash chromatography. 60 mg (57%) of the pure compound was obtained which was verified by the $^1$H and $^{13}$C-NMR. $^1$H NMR (DMSO-d$_6$, 500 MHz) δ (ppm) δ 8.42 (s, 1H) δ 8.16 (s, 1H) δ 7.85 (s, 1H) δ 7.46 (m, 4H) δ 7.24 (s, 1H) δ 6.91 (s, 1H) δ 5.72 (s, br, OH) δ 5.31 (s, 1H) δ 4.67 (d, $J = 15.5$ Hz, 1H) δ 4.52 (d, $J = 15.5$ Hz, 1H) δ 4.3 (m 1H) δ 4.2 (m 1H) δ 4.1 (s, br,1H) δ 2.9 (s, 3H). $^{13}$C NMR (DMSO-d$_6$, 500 MHz) δ (ppm) 141.64, 128.35, 127.36, 120.00, 117.39, 110.62, 68.19, 40.36, 40.20, 40.03.

![Chemical Structure](image)

**Isopropyl-carbamic acid phenyl ester (11)**

Isopropyl isocyanate (2.5 mL, 39.8 mmol, 1.5 eq.) was slowly added to the mixture of phenol (4.8 g, 26 mmol, 1 eq.) and triethylamine (1.4 mL, 5.3 mmol, 0.2 eq.) in dichloromethane at the room temperature under an argon atmosphere. After stirring at room temperature for 120 min, it was concentrated down under vacuum. The compound was isolated as white powder, it was then analyzed by $^1$H NMR and gas chromatography and used without further purification. 81% yield. $^1$H NMR (500 MHz, CDCl$_3$) δ 7.34 (t, $J = 7.7$ Hz, 2H, Ar-CH), δ 7.18 (t, $J = 7.15$ Hz, 1H,
Ar-CH), δ 7.1 (d, J = 7.45 Hz, 2H, Ar-CH), δ 4.84 (s, 1H, N-H), δ 3.91-3.88 (m, 1H, CH), δ 1.19 (d, J = 6.3 Hz, 6H, CH₃).

Isopropyl-carbamic acid-2-trimethylsilanyl-phenyl ester (Silyl Carbamate) (12)

A solution of Isopropyl-carbamic acid phenyl ester (2.0 g, 11.2 mmol, 1 eq.) in diethyl ether (25 mL) was charged with TMEDA (1.8 mL, 12.2 mmol, 1.1 eq.) and TBSOTf (2.8 mL, 12.2 mmol, 1.1 eq.) at 0 °C under argon atmosphere. It was gradually allowed to warm up to 25 °C and remaining TMEDA (3.35 mL, 22 mmol, 2 eq.) was added. n-BuLi (2.5 M, 11.2 mmol, 2 eq.) was slowly introduced to the reaction mixture at -78 °C over the period of 45 min. The reaction was continued to stir at -78 °C for 60 min. After that, trimethylsilyl chloride (5.0 mL, 39 mmol, 3.5 eq.) was added over the period of 35 min and the reaction was continued to stir at -78 °C for 90 min. The reaction mixture was quenched by a saturated solution of sodium bisulfate (200 mL), and the reaction was gradually allowed to warm up to room temperature. Work up was done by separating the organic layer; the organic layer was further washed with sodium bisulfate solution, dried over sodium sulfate and concentrated under vacuum. The white solid such obtained was analyzed using ¹H NMR and used without purification. Yield: 77%. ¹H NMR (500 MHz, CDCl₃) δ 7.4 (d, J = 6.3 Hz, 1H, Ar-CH), δ 7.3 (m, 1H, Ar-CH), δ 7.16 (t, J = 7.4 Hz, 1H, Ar-CH), δ 7.1 (t, J = 7.4 Hz, 1H, Ar-CH), δ 4.8 (s, 1H, NH), δ 3.9 (m, 1H, CH), δ 1.2 (d, J = 5.7 Hz, 6H, CH₃), δ 0.3 (s, 9H, CH₃).
Trimethylsilyl-2-phenyltriflate (Benzyne Precursor) (13)

1, 8-Diazabicyclo-[5.4.0]-undec-7-ene (DBU) (3.80 mL, 26 mmol, 1.5 eq.) and diethylamine (2.15 mL, 20 mmol, 1.2 eq.) was slowly introduced to the solution of silyl carbamate (4.38 g, 17 mmol, 1 eq.) in acetonitrile (30 mL) at room temperature under an argon atmosphere. The mixture was then heated at 40 °C for 45 min. It was then cooled down to room temperature; a solution of PhNTf₂ (9.3 g, 26 mmol, 1.5 eq.) in acetonitrile (25 mL) was slowly introduced dropwise over the period of 15 min. The resulting mixture was continued to stir at the room temperature under argon atmosphere for 120 min. The workup was done by washing the organic layer with saturated solution of sodium bisulfate and 10% sodium hydroxide solution twice. The organic layer was then dried over sodium sulfate, concentrated under vacuum. The resulting crude liquid was purified by flash chromatography and verified by ¹H NMR. Yield % = 65. ¹H NMR (500 MHz, CDCl₃) δ 7.54 (dd, J=5.6 Hz, 1H, Ar-CH), δ 7.45-7.42 (m, 1H, Ar-CH), δ 7.35-7.32 (m, 2H, Ar-CH), δ 0.36 (s, 9H, CH₃).
(2-Chloro-pyridin-3-yl)-phenyl-amine (14)

3-Amino-2-Chloro pyridine (0.86 g, 6.7 mmol, 2 eq.) dissolved in acetonitrile (15 mL) was added to oven dried cesium fluoride (1.02 g, 6.7 mmol, 2 eq.) under an argon atmosphere at the room temperature. Trimethylsilyl-2-phenyltriflate (1.0 g, 3.35 mmol, 1 eq.) dissolved in acetonitrile (40 mL) was slowly introduced to the reaction mixture over the period of 8 hours, by using a syringe pump. The reaction mixture was continued to stir for next 8 hours after complete addition. The reaction was monitored by thin layer (3:7 EtOAC:Hexane) and gas chromatography. The reaction was quenched by water; the aqueous layer was then extracted by ethyl acetate thrice. The combined organic layer was further washed with water and brine, dried over sodium sulfate, concentrated under vacuum. The crude solid such obtained was purified by flash chromatography. Yield = 41%. $^1$H NMR (500 MHz, CDCl$_3$) $\delta$ 7.87 (d, $J = 4.6$ Hz, 1H, Py-H), $\delta$ 7.5 (d, $J = 8$ Hz, 1H, Py-H), $\delta$ 7.34 (t, $J = 7.5$ Hz, Py-H), $\delta$ 7.2-7.1 (m, 4H, Ar-H).
5H-pyrido [3, 2-b] indole (δ-Carbazole) (10)

(2-Chloropyridin-3-yl)-phenylamine (0.153 g, 0.75 mmol, 1 eq.) was mixed with sodium acetate (0.249 g, 1.83 mmol, 2.5 eq.) and Pd (PPh₃)₂Cl₂ (0.042 g, 0.053 mmol, 8 mol%) in a 10 mL microwave vial. The vial was then flushed with argon for 5 min. N, N-Dimethyl acetamide (1 mL) was then added to the mixture, which was flush with argon for another 5 min. The vial was then heated to 180 °C in CEM Discover™ microwave oven at a power level 100 W for 10 min. The reaction mixture was then cooled down to room temperature, diluted with 60 mL of methanol and concentrated under vacuum. The crude product was purified by flash chromatography. The compound was isolated as a white solid in a 70% yield. MP = 210 °C – 216 °C. ¹H NMR (500 MHz, DMSO-d₆) δ 11.39 (s, 1H, NH), δ 8.4 (dd, 1H, J = 1.15 Hz, J = 3.45 Hz, Py-H), δ 8.16 (d, 1H, J = 7.45 Hz, Py-H), δ 7.8 (dd, 1H, J = 1.15 Hz, J = 6.85 Hz, Py-H), δ 7.53 (d, 1H, J = 8.6 Hz, Ar-H), δ 7.45 (td, 1H), δ 7.35 (dd, 1H, J = 4.6 Hz, J = 4.5 Hz, Ar-H), δ 7.21 (t, 1H, J = 7.5, Ar-H).

3-Trimethylsilyl-1H-pyridin-2-one (16)

2-Hydroxypyridine (3.0 g, 31.5 mmol, 1 eq.) was dissolved in diethyl ether (75 mL) under an argon atmosphere at room temperature. The reaction mixture was then allowed to stir over an
ice bath (0 °C to 5 °C) for about 10-15 min. After that, lithium diisopropylamide (2 M, 34.5 mL 69.3 mmol, 2.2 eq.) was drop-wise introduced to the cooled reaction mixture. After complete addition, reaction mixture was continued to stir over an ice bath for 10 min, chlorotrimethylsilane (6.2 mL, 72.5 mmol, 2.3 eq.) was slowly added to the reaction mixture while it was still stirring over an ice bath (0 °C -5 °C). The reaction mixture was then gradually warmed up to the room temperature and continued to stir overnight. Workup was done by filtering out the solid generated in the reaction, the filtrate such obtained was concentrated under vacuum. The column chromatography of the crude afforded 1.72 g white solid. $^1$H NMR (500 MHz, CDCl$_3$) $\delta$ 7.60 (d, $J$ = 4.55Hz, 1H) $\delta$ 7.44 (dt $J$ = 2.3 Hz, $J$ = 6.85 Hz, 1H) $\delta$ 6.32 (m, 1H) $\delta$ 0.27 (s, 9H).

![Structure](17)

**Trifluoro-methanesulfonic acid 3-trimethylsilanyl-pyridin-2-yl ester (pyridyne precursor) (17)**

3-Trimethylsilyl-1H-pyridin-2-one (1.72 g, 10.2 mmol, 1 eq.) was dissolved in DCM (10 mL) under argon atmosphere. 2,6-lutidine (1.4 mL, 12.3 mmol, 1.2 eq.) was then introduced to the reaction mixture drop-wise. Furthermore, reaction mixture was allowed to cool down to 0 °C, for 5 min., triflic anhydride (2 mL, 11.3 mmol, 1.1 eq.) was then gradually introduced to the cooled reaction mixture. Reaction mixture was then warmed up to room temperature and continued to stir with continuous monitoring by thin layer and gas chromatography. Allowed it to stir overnight at room temperature under argon. Reaction mixture was concentrated under vacuum and directly purified through flash chromatography 2.94 g liquid obtained. $^1$H NMR (CDCl$_3$, 500 MHz) $\delta$ 8.3
(dd $J = 4.6$ Hz, $J = 2.25$ Hz, 1H) $\delta$ 7.90 (dd, $J = 4.9$ Hz, $J = 7.4$ Hz, 1H) $\delta$ 7.30 (dt, $J = 5.15$ Hz, $J = 7.65$, 1H) $\delta$ 0.36 (s, 9H).

N-phenylpyridin-2-amine (18)

Aniline (0.30 g, 3.2 mmol, 1 eq.) dissolved in acetonitrile (10 mL) was added to cesium fluoride (1.5 g, 9.6 mmol, 3 eq.) under argon atmosphere at room temperature. The pyridyne precursor (1.04 g, 3.5 mmol, 1.1 eq.) dissolved in acetonitrile (10 mL) was slowly introduced to the reaction mixture. The reaction was continued to stir for next 8 hours after complete addition. The reaction was monitored by thin layer (3:7 EtOAc:Hexane) and gas chromatography. The work up was done by quenching the reaction with water; aqueous layer was then extracted by ethyl acetate thrice. Combined organic layer was further washed with water and brine, dried over sodium sulfate, concentrated under vacuum. The crude was purified by flash chromatography, isolated as yellow solid. Yield = 41%. $^1$H NMR (500 MHz, CDCl$_3$) $\delta$ 8.58 (d, $J = 1.15$ Hz, 1H), $\delta$ 8.42 (d, $J = 8$ Hz, 1H, Py-H), $\delta$ 7.34 (t, $J = 7.5$ Hz, Py-H), $\delta$ 7.2-7.1 (m, 4H, Ar-H).
N-(furan-2-ylmethyl)-N-(2-hydroxy-3-(phenyl(pyridin-2yl)-amino)-propyl)-methane sulfonamide (19)

N-phenylpyridin-2-amine (0.160 g, 0.95 mmol, 1 eq.) was dissolved in anhydrous 10 mL of N, N-dimethylformamide under an argon atmosphere at the room temperature. The reaction mixture was further cooled down to 0-5 °C for 15 min. To the cooled reaction mixture, added sodium hydride (1.4 mmol, 1.5 eq.) in batches over the period of 15 min. After complete addition, the reaction mixture was allowed to stir at the room temperature for 30 min before cooling it down again to 0 - 5 °C. To the cooled reaction mixture N-(Furan-2-ylmethyl)-N-(oxiran-2-ylmethyl)-methane sulfonamide (0.219 g, 0.95 mmol, 1 eq.) dissolved 5 mL of DMF was introduced drop wise. After addition, the reaction flask was equipped with a condenser and heated at 50 °C overnight under an argon atmosphere. The work up was done by quenching the reaction mixture with water and extracting it with ethyl acetate. The combined organic layer was washed with water followed by the brine solution, dried over sodium sulfate and concentrated under vacuum. The crude such obtained purified by the flash chromatography. 75 mg of pure compound was obtained. The compound was characterized by $^1$H NMR. $^1$H NMR (500 MHz DMSO-d$_6$) $\delta$ 8.07 (d, $J = 1.1$ Ar-H), $\delta$ 7.40-7.24 (m, 6-Ar-H), $\delta$ 7.54 (d, $J = 1.2$ Ar-H), $\delta$ 6.62 (d, $J = 1.5$ 1Ar-H), $\delta$ 6.3-6.2 (m, 3Ar-H), $\delta$ 5.4 (s, br, OH), $\delta$ 4.4-4.3 (dd, $J = 16$, $J = 36$, 2 H), $\delta$ 2.7 (s, CH$_3$).
2,3,4,9-Tetrahydro-1H-β-carboline (Pinoline): (21)

5-methoxytryptamine (2 g, 10.5 mmol, 1 eq.) was dissolved 10:1 mixture of acetic acid (38 mL) and methanol (3.8 mL) at the room temperature under an argon atmosphere. Paraformaldehyde (0.380 g, 12.6 mmol, 1.2 eq.) was then slowly introduced to the reaction mixture at room temperature. The solution was then heated at 80 °C for an hour, reaction was continuously monitoring by the gas chromatography. The mixture was then basified with ammonium hydroxide solution. The aqueous layer was extracted with DCM thrice. The combined organic layer was washed with water followed by the brine solution, dried over sodium sulfate and concentrated. 1.56 g of brown solid obtained which was purified by flash chromatography followed by column chromatography. The white solid obtained was verified by the $^1$H NMR. $^1$H NMR (500 MHz, DMSO- d$_6$) δ 10.3 (s, br, 1H) δ 7.11 (d, $J = 9.2$ Hz, 1H) δ 6.80 (s, 1H) δ 6.59 (dd, $J = 2.3$ Hz, $J = 8.6$ Hz, 1H) δ 4.61 (s, 2H) δ 3.84 (s, 3H), δ 2.90 (t, $J = 5.7$ Hz, 2H) δ 2.51 (t, $J = 5.7$ Hz, 2H).
**tert-butyl-6-methoxy-1,3,4,9-tetrahydro-2H-pyrido[3,4-b]-indole-2-carboxylate (22)**

Boc anhydride (2.3 g, 7.72 mmol) was dissolved in THF (70 mL) at the room temperature under an argon atmosphere. The reaction mixture was then allowed to stir at 0 °C over an ice bath for 10 min. 1.56 g of pinoline dissolved in THF (10 mL) was then slowly introduced to the cooled reaction mixture dropwise. After that, the reaction was allowed to stir at room temperature overnight. The reaction was quenched by a saturated solution of sodium bicarbonate (30 mL). The aqueous layer was then extracted with diethyl ether thrice. The combined organic later was washed with water followed by the brine solution, dried over sodium sulfate and concentrated. Flash chromatography of the crude afforded 0.930 g white solid, which was confirmed by ¹H NMR. ¹H NMR (CDCl₃, 500 MHz) δ 7.20 (d, J = 8.6 Hz, 1H) δ 6.92 (s, 1H) δ 6.82 (dd, J = 2.25 Hz, J = 8.55 Hz, 1H) δ 4.61 (s, 2H) δ 3.84 (s, 3H) δ 3.75 (t, J = 8.5 Hz, 2H) δ 2.75 (t, J = 5.7, 2H) δ 1.49 (s, 9H).
9-[3-(Furan-2-ylmethyl-methanesulfonyl-amino)-2-hydroxy-propyl]-6-methoxy-1,3,4,9-tetrahydro- β-carboline-2-carboxylic acid tert-butyl ester (23)

tert-butyl-6-methoxy-1,3,4,9-tetrahydro-2H-pyrido[3,4-b]-indole-2-carboxylate (0.710 g, 2.34 mmol, 1 eq.) was dissolved in DMF (15 mL) under an argon atmosphere at the room temperature. The reaction mixture was further cooled down to 0 °C -5 °C for 15 min. To the cooled reaction mixture, added sodium hydride (60% dispersion in mineral oil, 1.08 g, 4.60 mmol, 1.5 eq.) in batches over the period of 15 min. After complete addition, the reaction mixture was allowed to stir at room temperature for 30 min before cooling it down again to 0 °C -5 °C. N-(Furan-2-ylmethyl)-N-(oxiran-2-ylmethyl)-methane sulfonamide (0.542 g, 2.34 mmol, 1 eq.) dissolved in 10 mL of DMF was introduced into the reaction mixture drop wise while it was stirring over an ice bath. After complete addition, reaction mixture stirred at the room temperature overnight under an argon atmosphere. Work up was done by quenching the reaction with the saturated solution of ammonium chloride and extracting it with ethyl acetate thrice. The combined organic layer was washed with water followed by the brine solution. It was then dried over sodium sulfate and concentrated under vacuum. The crude obtained was then purified by flash chromatography yielding 0.590 gm of sticky white solid. The compound was used for the deprotection reaction without analysis by NMR.
N-Furan-2-ylmethyl-N-[2-hydroxy-3-(6-methoxy-1,2,3,4-tetrahydro-β-carboline-9-yl)-propyl]-methane sulfonamide (24)

9-[3-(Furan-2-ylmethyl-methanesulfonyl-amino)-2-hydroxy-propyl]-6-methoxy-1,3,4,9-tetrahydro- β-carboline-2-carboxylic acid tert-butyl ester (0.50 g, 0.93 mmol, 1 eq.) was refluxed with trifluoroacetic acid (0.722 mL, 9.30 mmol, 10 eq.) in DCM (5.00 mL) for 1 hour. The reaction was quenched by water (10.0 mL), basified by solution of 2 M sodium hydroxide. The aqueous layer was extracted with DCM thrice. The combined organic layer was washed with water followed by brine, dried over sodium sulfate and concentrated. Column chromatography afforded 0.170 g of the white solid. $^1$H-NMR (CDCl$_3$, 500 MHz) δ 7.3 (s, 1H) δ 7.14 (d, $J = 9.15$ Hz, 1H) δ 6.92 (d, $J = 2.3$ Hz, 1H) δ 6.82 (dd, $J = 2.85$ Hz, $J = 9.15$ Hz 1H) δ 6.31 (m, 1H) δ 6.17 (d, $J = 2.9$, 1H) δ 4.52-4.39 (m, 2H) δ 4.04-3.9 (m, 5H), δ 3.84 (s, 3H) δ 3.29 (q, $J = 7.45$, 1H) δ 3.22-3.11 (m, 3H) δ 2.79 (s, 3H). $^{13}$C NMR (DMSO-d$_6$, 500 MHz) δ 150.08, 143.87, 111.17, 110.65, 110.58, 100.33, 69.05, 68.96, 55.97, 55.92, 51.26, 51.11, 22.31.
6H-indolo[2,3-b]-quinoxaline (25)

A mixture of isatin (0.6 g, 4.07 mmol, 1 eq.) and ortho-phenylenediamine (0.438 g, 4.07 mmol, 1 eq.) was dissolved in acetic acid (7 mL) in a microwave tube equipped with a septum. The reaction was microwaved at 150 °C at the power level 160 W for 10 min. The reaction mixture was neutralized by pouring into an ice-cold saturated solution of sodium bicarbonate (250 mL). The aqueous layer was extracted with the ethyl acetate thrice. The combined organic layer was then washed with water followed by brine, dried over sodium sulfate and concentrated. The yellow color crude such obtained was recrystallized with ethyl acetate. To do so, the solid was dissolved in boiling ethyl acetate and stored in refrigerator overnight. The yellow needle solid such obtained was analyzed by $^1$H NMR. $^1$H-NMR (CDCl$_3$, 500 MHz) δ 8.89 (s, br, NH), δ 8.48 (d, $J = 7.45$ Hz, 1H) δ 8.33 (dd, $J = 8$ Hz, $J = 9.15$ Hz, 1H) δ 8.11 (dd, $J = 6.85$ Hz, $J = 8$ Hz, 1H) δ 7.79 (t, $J = 1.85$, 1H) δ 7.73-7.67 (m, 2H) δ 7.54 (d, $J = 8$ Hz, 1H) δ 7.41 (t, $J = 6.9$ Hz 1H).
N-(3-Benzol[b]carbazol-5-yl-2-hydroxy-propyl)-N-furan-2-ylmethyl-methanesulfonamide (26)

5H-Benzol[b]carbazole (0.360 g, 1.6 mmol, 1.5 eq) was dissolved in anhydrous 10 mL of N, N-dimethylformamide under argon atmosphere at room temperature. The reaction mixture was further cooled down to 0-5 °C for 15 min. To the cooled reaction mixture, added sodium hydride (60% dispersion in mineral oil, 0.080 g, 3.2 mmol, 1.5 eq.) in batches over the period of 10 min. Reaction mixture was allowed to stir at room temperature for 1 hour. It was again cooled down to 0-5 °C and N-(Furan-2-ylmethyl)-N-(oxiran-2-ylmethyl)-methane sulfonamide (0.231 g, 0.95 mmol, 1 eq.) dissolved in 10 mL DMF was slowly introduced. After complete addition reaction mixture was then equipped with condenser and refluxed at 50 °C overnight under argon atmosphere. Work up was done by quenching the reaction with water and extracting it with ethyl acetate thrice. The Combined organic layer was washed with brine, dried over sodium sulfate and concentrated under vacuum. Purification by flash chromatography followed by recrystallization in ethyl acetate afforded yellow solid. 1H NMR (CDCl3, 500 MHz) δ 8.44 (d, J = 8 Hz, 1H) δ 8.30 (dd., J = 8.05 Hz, J = 8.75 Hz, 1H) δ 7.76 (dt, J = 1.15 Hz, 1H) δ 7.68 (q, J = 7.45 Hz, 2H) δ 7.54 (d, J = 8 Hz, 1H) δ 7.38 (t, J = 6.9 Hz, 1H), δ 7.23 (s, 1H) δ 6.19 (d, J=1.7Hz, 1H) δ 5.13 (s, br., OH) δ 4.67 (d, J= 2.3 Hz, 1H) δ 4.55 (s, 2H) δ 4.46-4.43 (m, 2H) δ 3.49 (q, J = 1.15, 2H) δ 2.85
(s,1H). $^{13}$C NMR (CDCl$_3$, 500 MHz) 149, 146, 144, 142, 140, 138, 131, 129.5, 127.3, 121, 119, 111, 110, 70, 51.1, 47.1, 45.3, 38.1.
REFERENCES


3.1 Introduction

KL001 is a drug which is known for its period altering effect. Remarkably, KL001 is found to lengthen the circadian period of cryptochrome\textsuperscript{1-3}, which regulates gluconeogenesis\textsuperscript{4}. Incidentally, the structure-activity relationship of KL001 revealed that the carbazole moiety lacks critical hydrogen bonding with the essential polar residues in the binding pocket. After the detailed investigation it was found that the side chain of the KL001 can also be modified for the improved binding. In the previous chapter, different analogs were synthesized with either a polar moiety in the ring system or with an extra ring. In this chapter, however, the focus will be on modifying the linker and the side chain of the KL001\textsuperscript{5}. Furthermore, another focus will be on making different analogs of KL044\textsuperscript{2}, which recently have shown similar or better activity than KL001.

3.2 Structure-Activity relationship of KL001 side chain

Crystal structure of the KL001 side chain revealed that the furan ring of the side chain embed itself close to the residues enriched with aromatic and polar substituents\textsuperscript{5}. Particularly,
tryptophan (W 310) and glutamine (Q 307) are oriented very close to the furan ring. Therefore, modifying the side chain with the substituents having a propensity to participate in the pi-pi interactions or in hydrogen bonding will presumably enhance the binding efficiency of the small molecule.

While our work was ongoing, Lee et al., had designed a compound named KL044 which had a different side chain and linker. In KL044, the isopropyl linker of the KL001 was replaced by an amide linkage, while 2-cyano-6-chloro amine was used as a side chain. Structure-activity relationship of KL044 had revealed, the amide linker participates in critical hydrogen bonding interactions with the histidine and serine in the vicinity. While the carbonyl of amide interacts with the serine (S 394) (3.06 Å) and amino functionality participates in hydrogen bonding interaction.

**Figure 26: Structure-activity of the KL001 side chain:** The side chain binding site is enriched with the aromatic and the polar residues. The furan ring of the KL001, snuggles close to the tryptophan (W 310) and polar glutamine (Q 307) residues.
with 2.35 Å away histidine residue (H 357). Furthermore, the disubstituted benzyl group forms pi-pi interaction with the tryptophan in the vicinity.

**Figure 27: Structure-activity relationship of KL044**: In KL044 the isopropyl linker is replaced by the amide linkage. The amide linkage participates in important interactions with the polar residues present in the binding pocket. While carbonyl functionality of the amide linker forms a hydrogen bond with the histidine 2.3 Å away, the serine residue forms a critical hydrogen bonding with the NH functionality of amide.

Gaining insights from our detailed outlook on the structure-activity relationship and KL044 interactions, we diverted our focus on to modify the side chains and the linker in rest of the Zoltowski library. To accomplish that, we used several different side chains with the propensity of participating in CH-pi interactions and hydrogen bonding.
3.4. Design principle of ZL006

![Chemical structures]

For the sixth compound in the library nitrogen and sulfur containing bicyclic ring systems were used as a side chain in lieu of methane sulfonamide and furan ring. Additionally, the isopropyl linker was replaced by an amide linkage.

Benzothiazole was strategically chosen as a side chain for the carbazole in ZL006. A benzthiazole ring system resembles tryptophan ring, remarkably, two of the tryptophan residues are present near the side chain binding site (W 310 and W 417). We envisioned the bicyclic structure of the side chain could participate in π stacking interactions with the two-tryptophan present, also potentially participate in a stronger hydrogen bond. Most importantly, the benzothiazole nucleus itself is known to have antimicrobial, antidiabetic and antiviral and anti-inflammatory properties\(^6\text{-}^9\). Besides this, the modified amide linker is anticipated to participate in hydrogen bonding interactions with the neighboring histidine and serine.

In order to synthesize the side chain, 6-benzothiazolamine was treated with chloroacetylchloride in the presence of triethylamine in DCM at room temperature to obtain N-benzothiazolchloroacetamide 27. It was then coupled to the carbazole in presence of sodium hydride in DMF to obtain ZL006 28.
3.5 Design Principle of ZL007

Imidazo-[2,1-b]-[1,3,4]thiadiazole was used as a side chain for ZL006. Several natural products having Imidazo-[2,1-b]-thiadiazole moiety\textsuperscript{10-12} had have been shown to have biological activity\textsuperscript{29}. We anticipated Imidazo-[2,1-b]-[1,3,4]thiadiazole ring would participate in π-π interaction with trp-290, on the other hand, the extra benzene ring could snuggle itself deep inside the cavity and could potentially participate in π-π stacking with W 310 and W 417. Besides this the linker has just a carbonyl group, which could potentially participate in hydrogen bonding.
The side chain was synthesized by the published synthesis with a few modifications\textsuperscript{12}. 2-Thiazolamine was refluxed with 2-bromoacetophenone in ethanol overnight to obtain 2-phenylimidazo[2,1-b]-thiazole 28. The cyclized product obtained was then heated with chloroacetyl chloride in 1,4-dioxane to get chloroacetyl derivative of 2-phenylimidazo[2,1-b]-thiazole 29. It was later coupled with carbazole in the presence of sodium hydride and DMF to achieve ZL007 30.
3.6 Design Principle of ZL008

The design principle of another compound in the library was to keep the side chain similar to KL001 while replacing the linker to amide. The idea was to replace the side chain furan ring with another heterocyclic five-membered ring system, possibly, with 1,3 azoles. We chose to introduce 1,3-thiazole as a new side chain owing to its capability of larger pi delocalization than the corresponding oxazole.

In order to synthesize the side chain, 2-amino-thiazole was treated with chloroacetyl chloride in the presence of triethylamine in dichloromethane at room temperature to obtain N-benzothiazolchloroacetamide 32. It was then coupled with carbazole in the presence of sodium hydride in DMF to obtain ZL008 33.

Scheme 12: Synthesis of ZL008 (30%)
3.7 Design Principle of ZL009

Until now, the compounds synthesized were similar to KL001; modifications were done either on the scaffold while keeping the side chain same or vice versa. For the 9th compound, however, we had designed a novel compound in the library, by using one of the new scaffolds and a side chain. For this compound based on our detailed investigation of structure activity relationship, we used 5H-Benzob[2]carbazole 5 as a scaffold and benzo-thiazolchloroacetamide 27 as a side chain, attached by an amide linker. We anticipate the four-ring system can possibly mimic isoalloxazine ring efficiently, unlike carbazole which can only mimic xylene and pyrazine rings of the whole isoalloxazine ring system. Furthermore, we chose benzo-thiazole side chain because of its structural similarity with the tryptophan ring system present in the vicinity. The compound was synthesized by coupling 5H-Benzob[2]carbazole 5 with benzo-thiazolchloroacetamide 27 in the presence of sodium hydride and N, N-dimethylacetamide under argon atmosphere.

Scheme 13: Synthesis of ZL009 (24%)
3.8 Design Principle of ZL010

The last compound in the library was synthesized by keeping in mind that there were several polar residues present in the side chain binding pocket like glutamine (Q 307). Furthermore, as mentioned earlier, KL044 having a benzene ring with substituents facilitated binding and showed better activity than KL001. Therefore, we selected tri-methoxy benzene ring to make another variant of KL044.

![Chemical structure of ZL010](attachment:image)

The side chain 35 was synthesized by treating 3,4,5-trimethoxyaniline with chloroacetyl chloride in the presence of triethylamine as a base in dichloromethane at room temperature. The side chain was then linked with carbazole in the presence of sodium hydride in DMF to obtain ZL010 36.

![Scheme 14: Synthesis of ZL010 (36%)](attachment:image)
3.9 Materials and methods

All the reactions were performed in the dried glassware under an argon atmosphere. All reagents were purchased from Sigma Aldrich (St. Louis, MO), Fisher scientific (Waltham, MA), Alfa Aesar (Ward Hill, MA), Chem Impex (Wood Dale, IL) and used as it was received. The deuterated solvents (CDCl₃, DMSO-d⁶) were purchased from Cambridge Isotope Laboratories (Cambridge, MA). ¹H-NMR and ¹³C NMR spectra were recorded using 500 MHz Jeol multinuclear NMR spectrometer in the Department of Chemistry at Southern Methodist University. All chemical shifts are reported in the standard notation of parts per million using the peak of residual proton signals of the deuterated solvent as an internal reference. The coupling constant units are in Hertz (Hz) splitting patterns are indicated as follows: br., broad; s., singlet; d., doublet; t., triplet; q., quartet; m., multiplet; dd., doublet of doublets; dt., doublet of triplets. Microwave reactions were carried out in CEM Discover™. Column chromatography was carried out on Combi-Flash instrument by using pre-packed silica gel columns.

![Chemical Structure Image]

N-Benzothiazol-6-y1-2-chloro-acetamide (27)

6-Aminobenzthiazole (2.70 g, 17.9 mmol, 1 eq.) dissolved in dichloromethane (30.0 mL) at the room temperature under an argon atmosphere. Triethylamine (3.00 mL, 21.5 mmol, 1.2 eq.) was introduced at room temperature. The reaction mixture was then allowed to stir over an ice bath maintained at the temperature of 0 to 5 °C for 15 min. To the cooled reaction mixture, chloroacetyl chloride (1.7 mL, 21.5 mmol, 1.2 eq.) dissolved in 20 mL of DCM was then added drop-wise over
the period of 10 -15 min. After stirring it over an ice bath for 15 min, the reaction mixture was allowed to stir at room temperature overnight. The reaction mixture was concentrated, and the crude was dissolved in ethyl acetate and water. Aqueous layer was extracted with ethyl acetate thrice. The combined organic layer was washed with saturated solution of sodium bicarbonate, 5% hydrochloric acid followed by water and brine. It was then dried over sodium sulfate and concentrated. Column chromatography of crude solid afforded 2.23 g (55%) of white solid. The solid was further recrystallized with acetone. which was confirmed by $^1$H NMR. $^1$HNMR (DMSO-d$_6$) δ 10.58 (s, br., 1H) δ 9.26 (s, 1H) δ 8.49 (d, $J = 2.3$ Hz, 1H) δ 8.01 (d, $J = 12.8$ Hz, 1H) δ 7.56 (dd., $J = 8.6$ Hz, $J = 10.9$ Hz, 1H) δ 4.27 (s, 2H).

![Chemical Structure](image)

N-Benzothiazol-6-yl-2-carbazol-9-yl-acetamide (28)

Carbazole (1.00 g, 5.90 mmol, 1 eq.) was dissolved in anhydrous 50 mL of DMF under an argon atmosphere at room temperature. The reaction mixture was cooled down to 0 °C to 5 °C for 10 min. To the cooled reaction mixture, sodium hydride (60% dispersion in mineral oil 0.359 g, 8.9 mmol) was introduced over the period of 10 min. After complete addition, the reaction mixture was continued to stir over an ice bath for 15 min followed by at the room temperature for about 1 hour. After 1 hour of stirring, the reaction mixture was again cooled down to 0 °C to 5 °C, and N-
benzothiazol-6-yl-2-chloro-acetamide (1.35 g, 5.90 mmol, 1 eq.) dissolved in 5.00 mL of DMF was added to the reaction mixture drop-wise. After that, the reaction flask was equipped with reflux condenser and allowed to heat at 50 °C overnight under the argon atmosphere. The workup was done by quenching the reaction with water and extracting it with ethyl acetate (3 x 30 mL). The combined organic layer was washed with water twice followed by brine and dried over sodium sulfate, concentrated under vacuum. The white crude obtained was purified by the flash chromatography. The sticky solid such obtained was crystallized by dissolving in dichloromethane and then adding n-hexane dropwise till white crystalline solid settled down in the round bottom flask. The solid was filtered under the vacuum, dried and confirmed by 1H NMR. 1H-NMR (CDCl₃, 500 MHz) δ 10.77 (s, br., NH) δ 9.23 (s, 1H) δ 8.48 (d, J = 1.75 Hz, 1H) δ 8.14 (d, J = 7.45 Hz, 3H) δ 8.02 (d, J = 8.6 Hz, 1H) δ 7.58-7.56 (m, 3H) δ 7.41 (t, J = 6.9 Hz, 2H) δ 7.19 (t, J = 8 Hz, 2H) δ 5.3 (s, 2H). 13C-NMR (500 MHz, DMSO-d₆) δ (ppm) 167.12, 155.47, 149.81, 141.28, 136.98, 134.85, 126.27, 123.64, 122.81, 120.75, 119.62, 119.25, 112.32, 109.88, 40.37.

![Image of 6-Phenyl-imidazo[2,1-b]-thiazole (29)]

**6-Phenyl-imidazo[2,1-b]-thiazole (29)**

The mixture of 2-aminothiazole (1.0 g, 99.7 mmol, 1 eq.) and 2-bromoacetophenone (1.98 g, 99.7 mmol, 1 eq.) was dissolved in 20 mL ethanol and refluxed overnight. Ethanol was removed under vacuum; the crude such obtained was dissolved in water and DCM. The aqueous layer was extracted with DCM thrice. The combined organic layer was washed with saturated solution of sodium bicarbonate followed by water and brine. It was then dried over sodium sulfate and
concentrated. The orange colored solid (1.25 g) obtained was taken for next step without further purification. $^1$H-NMR (DMSO-d$_6$, 500 MHz) δ 8.19 (s, 1H) δ 7.90 (d, $J$ = 4.6 Hz, 1H) δ 7.80 (dd, $J$ = 1.15 Hz, $J$ = 8.6 Hz, 2H) δ 7.34 (dt, $J$ = 7.45 Hz, $J$ =13.7 Hz, 2H) δ 7.23- 7.21 (m, 2H).

![Chemical Structure](image)

2-Chloro-1-(6-Phenyl-imidazo[2,1-b]-thiazol-5yl)-ethanone (30)

6-Phenyl-imidazo-[2,1-b]-thiazole (1.25 g, 6.25 mmol, 1 eq.) was dissolved in 1,4-dioxane in a two necked 50 mL flask, equipped with condenser and an argon balloon. The mixture was heated at 70 °C and chloroacetyl chloride (1.5 mL, 18.7 mmol, 3 eq.) was introduced drop-wise. After complete addition reaction mixture was continued to stir at 70 °C for 15 min. thereafter, refluxed for 3 hours, reaction was monitored through gas chromatography. The reaction mixture was cooled down to room temperature and filtered. The residue solid was washed with cold ethanol. 0.390 g of pale white solid obtained. $^1$H-NMR (500 MHz, DMSO-d$_6$) δ 8.48 (d, $J$ = 8.6 Hz, 1H) δ 7.61-7.50 (m, 6H) δ 4.30 (s, 2H).
2-Carbazol-9-yl-(6-phenyl-imidazo[2,1-b]-thiazol-5-yl)-ethanone (31)

Carbazole (0.210 g, 1.26 mmol, 1 eq.) was dissolved in anhydrous 10 mL of DMF under an argon atmosphere at room temperature. The reaction mixture was cooled down to 0 °C to 5 °C for 10 min. To the cooled reaction mixture, sodium hydride (60% dispersion in mineral oil 0.075 g, 1.89 mmol, 1.5 eq.) was introduced over the period of 15 min, and then allowed to stir room temperature for 1 hr. After 1 hour of stirring, it was again cooled down to 0 °C -5 °C, and 2-Chloro-1-(6-Phenyl-imidazo[2,1-b]-thiazol-5yl)-ethanone (0.35 g, 1.26 mmol, 1 eq.) dissolved in 5 mL of DMF was added to the reaction mixture drop-wise. After that, the flask was equipped with reflux condenser and allowed to heat at 50 °C overnight under argon atmosphere. The reaction was quenched with water and extracted with ethyl acetate. The combined organic layer was washed with water followed by brine, dried over sodium sulfate and concentrated under vacuum. The flash chromatography purification yielded sticky solid which was crystalized by dissolving in dichloromethane and then adding n-hexane drop wise till white crystalline solid settled down in the round bottom flask. The solid was filtered, dried and confirmed by 1H-NMR. 1H-NMR (CDCl3, 500 MHz) δ 7.3 (s, 1H) δ 7.14 (d, J = 9.15Hz, 1H) δ 6.92 (d, J = 2.3 Hz, 1H) δ 6.82 (dd, J = 2.85 Hz, J = 9.15 Hz 1H) δ 6.31 (m, 1H) δ 6.17 (d, J = 2.9, 1H) δ 4.52- 4.39 (m, 2H) δ 4.04-3.9 (m, 5H), δ 3.84 (s, 3H) δ 3.29 (q, J = 7.45, 1H) δ 3.11 (s, 2H).
2-Chloro-N-(thiazol-2-yl)-acetamide (32)

2-Amino-thiazole (2.00 gm, 19.9 mmol, 1 eq.) was dissolved in dichloromethane (DCM 30.0 mL) at room temperature under the argon atmosphere. Triethylamine (3.30 ml, 24 mmol, 1.2 eq.) was then introduced at room temperature. The reaction mixture was then allowed to cool down to 0 to 5 °C for 15 min. Chloroacetyl chloride (1.9 ml, 24 mmol, 1.2 eq.) dissolved in 20 mL DCM was then dropwise introduced at 0 °C, over the period of 10 min. After stirring it over an ice bath for 15 min, the reaction mixture was allowed to stir at room temperature overnight. Concentrating it under vacuum afforded crude, which was dissolved in ethyl acetate and water. Aqueous layer extracted with ethyl acetate thrice. The combined organic layer was washed with saturated solution of sodium bicarbonate, 5% hydrochloric acid followed by water and brine. It was then dried over sodium sulfate and concentrated. A column chromatography of crude solid afforded 1.63 g of white solid, which was confirmed by $^1$H NMR. $^1$H-NMR (500 MHz, CDCl$_3$) $\delta$ 7.51 (d, $J = 3.45$ Hz, 1H), $\delta$ 7.05 (d, $J = 3.45$ Hz, 1H), $\delta$ 4.27 (s, 2H).
2-(9H-carbazol-9-yl)-N-(thiazol-2-yl)-acetamide (33)

Carbazole (0.7 g, 4.1 mmol, 1 eq.) was dissolved in anhydrous 10 mL of DMF under the argon atmosphere at room temperature. The reaction mixture was cooled down to 0 °C to 5 °C for 10 min. To the cooled reaction mixture, sodium hydride (60% dispersion in mineral oil 0.251 g, 6.3 mmol, 1.5 eq.) was introduced over the period of 10 min, and then allowed to stir room temperature for 1 hr. After 1 hour of stirring, it was again cooled down to 0 °C to 5 °C, and 2-Chloro-N-(thiazol-2-yl)-acetamide (0.740 g, 4.1 mmol, 1 eq.) dissolved in 5 mL of DMF was added to the reaction mixture drop-wise. Thereafter, the flask was equipped with reflux condenser and allowed to heat at 50 °C overnight under argon atmosphere. Work up was done by quenching the reaction with water and extracting it with ethyl acetate. Combined organic layer was washed with brine, dried over sodium sulfate and concentrated under vacuum. The crude obtained was then purified by flash chromatography and sticky solid was obtained. $^1$H-NMR (500, MHz, DMSO-d$_6$) δ 8.14 (d, $J =$ 7.35 Hz, 2 Ar H), δ 7.54 (d, $J =$ 8 Hz, 3ArH), δ 7.48 (d, $J =$ 4.05 Hz, 2 Ar H), δ 7.48 (t, $J =$ 8.6, 3 Ar H), δ 7.21-7.18 (m, 3 Ar H), δ 5.3 (s, 2H). $^{13}$C-NMR (500 MHz, DMSO-d$_6$) δ 167.11, 158.14, 141.13, 138.29, 126.36, 122.85, 120.78, 119.77, 114.40, 109.80, 40.36.
N-(benzo[d]thiazol-5-yl)-2-(6H-indolo[2,3-b]-quinoxalin-6-yl)-acetamide (34)

5a,11a-dihydro-5H-benzo[b]carbazole (0.410 g, 1.87 mmol, 1 eq.) was dissolved in anhydrous 15 mL of DMF under the argon atmosphere at room temperature. The reaction mixture was cooled down to 0 °C to 5 °C for 10 min. To the cooled reaction mixture, sodium hydride (60% dispersion in mineral oil 0.112 g, 2.80 mmol, 1.5 eq.) was introduced over the period of 15 min and then allowed to stir room temperature for one hr. After 1 hour of stirring, it was again cooled down to 0 °C to 5 °C, and N-Benzothiazol-6-yl-2-chloro-acetamide (0.422 g, 1.87 mmol, 1 eq.) dissolved in 10 mL of DMF added to the reaction mixture drop-wise. After that, the flask was equipped with reflux condenser and allowed to heat at 50 °C overnight under argon atmosphere. Work up was done by quenching the reaction with water and extracting it with ethyl acetate. The combined organic layer was washed with brine, dried over sodium sulfate and concentrated under vacuum. The crude obtained was then purified by flash chromatography. The yellow solid such obtained was recrystallized with ethyl acetate. $^1$H-NMR (500 MHz, DMSO-d$_6$) δ 10.81 (s, NH), δ 9.25 (s, 1 Ar H), δ 8.46 (d, $J = 1.7$ Hz, 2 Ar H), δ 8.40 (d, $J = 7.4$ Hz, 1 Ar H), δ 8.01 (d, $J = 7.4$ Hz, 1 Ar H), δ 7.77- δ 7.73 (m, 4 Ar H), δ 7.61 (d, $J = 8.1$ Hz, 1 H), δ 5.43 (s, 2 H). $^{13}$C-NMR (400 MHz, DMSO-d$_6$) δ (ppm) 166.39, 155.45, 149.78, 145.92, 145.51, 140.26, 139.40, 136.79, 134.78, 131.91, 129.60, 127.92, 126.80, 123.57, 122.62, 121.82, 119.22, 117.69, 112.36, 111.24. 44.85.
2-Chloro-N-(3,4,5-trimethoxyphenyl)-acetamide (35)

3,4,5-Trimethoxyaniline (2.00 gm, 11 mmol) was dissolved in dichloromethane (DCM 20.0 mL) at the room temperature under argon atmosphere. Triethylamine (1.83 mL, 13 mmol) was then introduced at room temperature. The reaction mixture was then allowed to cool down to 0 to 5 °C for 15 min. Chloroacetyl chloride (1.0 mL, 13 mmol) dissolved in 10 mL DCM was then added to the reaction mixture drop-wise at 0 °C, over the period of 10 min. After stirring it over ice bath for 15 min, reaction mixture was allowed to stir at room temperature overnight. Concentrating it under vacuum afforded crude, which was dissolved in ethyl acetate and water. Aqueous layer extracted with ethyl acetate thrice. Combined organic layer was washed with saturated solution of sodium bicarbonate, 5% hydrochloric acid followed by water and brine. It was then dried over sodium sulfate and concentrated. Flash chromatography of crude solid afforded 2.6 g of white solid, which was confirmed by ¹H NMR. ¹H-NMR (500 MHz, CDCl₃) δ 8.12 (s, NH), δ 6.82 (s, 2 Ar H), δ 4.18 (s, 2H), δ 3.86 (s, 9 H).
2-(9H-carbazol-9-yl)-N-(3,4,5-trimethoxyphenyl)-acetamide (36)

Carbazole (0.4 g, 2.39 mmol, 1 eq.) was dissolved in anhydrous 10 mL of DMF under argon atmosphere at the room temperature. The reaction mixture was cooled down to 0 °C to 5 °C for 10 min. To the cooled reaction mixture, sodium hydride (60% dispersion in mineral oil 0.144 g, 3.59 mmol, 1.5 eq.) was introduced in portions over the period of 10-15 min. Following complete addition, the reaction mixture was stirred at room temperature for 1 hour. The reaction mixture was cooled down to 0 °C to 5 °C and 2-chloro-1-(6-phenyl-imidazo[2,1-b]-thiazol-5yl)-ethanone (0.621 g, 2.39 mmol, 1 eq.) dissolved in 5 mL of DMF was introduced dropwise to the ice cooled reaction mixture. Thereafter, the flask was equipped with reflux condenser and allowed to heat at 50 °C overnight under argon atmosphere. Work up was done by quenching the reaction with water and extracting it with ethyl acetate. The combined organic layer was washed with brine, dried over sodium sulfate and concentrated under vacuum. The crude obtained was then purified by flash chromatography and sticky solid was obtained. The sticky solid was crystalized by dissolving in dichloromethane and then adding n-hexane drop wise till white crystalline solid settled down in the round bottom flask. The solid obtained was filtered and $^1$H-NMR (500 MHz, DMSO-d$_6$) $\delta$ 10.46 (s, NH), $\delta$ 8.14 (d, $J$ = 7.5 Hz, 2 ArH), $\delta$ 7.54 (d, $J$ = 8 Hz, 2 ArH), $\delta$ 7.41 (t, $J$ = 1.15 Hz, 2ArH), $\delta$ 7.19 (t, $J$ = 1.15 Hz, 2ArH), $\delta$ 6.95 (s, 2ArH), $\delta$ 5.21 (s, 2ArH), $\delta$ 3.67 (s, $J$ = 8 Hz, 6H), $\delta$ 3.56 (s, 2H). $^{13}$C-NMR (500 MHz, DMSO-d$_6$) $\delta$ (ppm) 166.69, 153.30, 141.26,
135.46, 134.01, 133.97, 126.25, 122.79, 120.73, 119.59, 119.30, 109.85, 109.67, 97.16, 60.71, 60.63, 56.15, 55.94, 46.54, 46.43.
REFERENCES


CHAPTER 4

CELL BASED CIRCADIAN ASSAY

4.1 Introduction

The design principle of all the compounds in the Zoltowski library relies on the careful investigation of the binding pocket of FAD and the structure-activity relationship of the KL001, already known for efficient binding and activity. The activity of compounds in the Zoltowski library as a cryptochrome modulators is tested by a cell-based luciferase activity using a human osteosarcoma U2OS cell lines stably expressing clock reporter Bmal1-dLuc. To our delight, most of the compounds of the library showed robust rhythm at all the concentrations tested, and some of them delay the rhythm at higher concentrations. This chapter discusses the results obtained from the circadian activity assay of the Zoltowski library. Additionally, a comparative study of the period shift with that of the KL001 is explored.

4.2 Cell Based Circadian Assay

For the circadian rhythm assay, it is crucial to select a cell line which shows a robust circadian rhythm and can survive longer under confluent conditions, and hence are suitable for long-term circadian recordings. Notably, several research groups have used human osteosarcoma U2OS cells for the circadian rhythm assay owing to its robust rhythm. Additionally, the rhythmic
expression of the clock gene can be monitored noninvasively by using firefly luciferase reporter. In most of the cells, the clock genes Bmal1 and per2 show robust circadian rhythm. Therefore, for our experiments, we used U2OS cells harboring Bmal-dLuc reporter for monitoring circadian rhythm following drug administration.

4.2.1: Compounds screening

The protocol of compound screening was adapted from Hirota et al. The compound screening was done in an opaque bottom 96 well plate. After trypsinization, the cells were suspended in a batch of culture media and plated 100 µL containing 2000 cells per well. The plate was then incubated at 37 °C for 2 days (48 hours) to facilitate the cells to reaching full confluency. The cell growth was monitored by plating a clear bottom plate with the same amount of culture media in parallel.

To measure luminescence, the media was replaced with 100 µl explant media containing HEPES buffer to maintain the optimal pH, a B-27 supplement which synchronizes the cellular clock and luciferin (1 mM). All the drugs were dissolved in 100% DMSO. The drugs were introduced into each well by keeping the DMSO concentration constant. KL001 was used a reference drugs, while some of the wells were only supplied with DMSO. The plate was then covered with a thin transparent film. The rhythm was then monitored using BioTek cytation 5 plate reader. The data was collected for the five consecutive days (120 hours) at the temperature 35 °C, integration time of 14sec, interval time 1 hour 40 min.
4.2.2 Luminescence Analysis

The luminescence data was analyzed by using R-project computing environment. We used the R-code analysis algorithm for curve fitting and data display written by Hirota et al. (2008).\(^1\)\(^5\). The raw luminescence data were fit using following equation:

\[
\text{Luminescence} = \text{Baseline} + mt + \text{Amplitude} \times (e^{-kt}) \cos \left(\frac{2\pi(t - \text{Phase})}{\text{Period}}\right)
\]

The output .txt file was used to plot Luminescence vs Concentration graph by using the Origin graphing software (Origin Lab 2018).

In addition to observing the rhythm, period lengths were calculated using another R script non-parametric algorithm JTK_Cycle developed by Professor Karl Kornacker (Ohio state Univeristy) and Professor John Hogenesch (University of Pennsylvania School of Medicine).\(^2\). The output file was then used to plot a bar graph of Period vs. concentration. The period length data was then used to accurately compare the period altering affects with that of KL001.
Figure 28: Schematic representation of the cell-based circadian study: The potency of the small compounds to bind with cryptochrome and mimic FAD was tested using BMAL1-dLuc reporter. Cryptochrome is a repressor of BMAL1 thereby expression of the BMAL1 can be used to monitor the cryptochrome activity. 96 well plate plated with U2OS cells harboring BMAL1-dLuc reporter and the luminescence monitoring was done in the biotek plate reader. The output was analyzed by the R computing environment to observe the period altering effects. The curve fitting program was used to plot the graph between luminescence vs. time. The JTK cycle was used to tabulate the period shift behavior of the compounds.

4.3 Results and discussions

The R-computing script was used to analyze the output data obtained from the plate reader. The curve-fitting was done using the R-script written by Hirota et al. and the output generated with the script was used to plot the graph Luminescence vs. time. The bar graph period vs. concertation
was plotted using the JTK_cycle R-script. Notably, the JTK cycle script was written for the initial 48-time points. Therefore, the bar graph only gives the snapshot of the initial 48 time points while the experiment was done for 120 hours with 73 data points.

4.3.1 Activity of KL001 as a cryptochrome modulator

The activity of KL001 was tested and used as a reference for all small-molecule activity. KL001 showed an efficient period-altering ability, consistent with the published reports. At low concentrations, the period was delayed by 1.7 hours. However, the rhythms are completely abolished at the higher concentrations (8 µM, not shown in the bar graph) indicating that even though KL001 is a potent drug, it could lead to many other clock-related dysfunctions at the higher concentrations.

![Graphs showing the activity of KL001](image)

**Figure 29: Activity of KL001:** A) KL001 shifts period in the dose-dependent manner, however rhythm gets abolished at higher concentration. B) The bar-graph clearly shows a dose-dependent increase in period shift.
4.3.2 Activity of ZL001 as a cryptochrome modulator

While we were still developing the cell lines to analyze the potency of the drugs in altering the cellular rhythm, our collaborator Dr. Carrie Partch UC Santa Cruz compared the efficacy of ZL001 with that of KL001 using live cell luciferase imaging.

The U2OS cell line expressing luciferase reporter gene administered with the drugs at the different concentrations was tested for period altering effects. ZL001 delayed the period at the higher concentration without killing the rhythm, at the same time a robust rhythm was observed at 0.1 µM concentration. The reference drug, KL001 (carbazole-scaffold), however, had abolished the rhythm entirely at the higher concentration (24 µM) (blue). The results depicted ZL001 has a mild action in delaying the rhythm.
Figure 30: A comparison of CRY modulators: KL001 (A) and ZL001 (B): X axis: Time (days), Y axis: Amplitude of luciferase activity. Both the compounds have shown a period shift. At the higher concentration, KL001 abolishes the rhythm while still maintained with ZL001. ZL001 showed robust rhythm at 0.1 µM (red) concentration (Collaborator: Dr. Carrie Partch UC Santa Cruz)

ZL001 screening was also conducted in our laboratory by the procedure mentioned earlier. Our data corroborated the results obtained from the Partch lab. ZL001 was able to maintain the rhythm at the lower concentration, while delays the period by 1.7 hours at the higher concentration. On the contrary, KL001 is toxic to circadian function at a higher concentration. The result dictates that KL001 could potentially have many rhythm related side effects such as sleep deprivation or other metabolic diseases, while ZL001 might not as rhythm was maintained at a higher
concentration as well. Also, since ZL001 was able to maintain the rhythm at the lower concentration, therefore, it could be used to mimic FAD for the biophysical studies.

Figure 31: ZL001 delays rhythm at the higher concentrations: (A) Structure of ZL001 (B) A graph of luminescence vs. time shows delayed rhythm at the highest concentration (8 µM, teal) in the experimental condition, a slight shift is also seen at 2.7 µM (light brown). (C) A bar graph plotted between period vs. concentration shows the period shift from 25.05 (zero drug) to 26.72. (D) Comparison of ZL001 with KL001: KL001 shows dose dependent increase in the period shift activity, however abolishes the rhythm at the higher concentration, while ZL001 is active at the higher concentration and maintains rhythm at all the concentrations used in the experimental condition.
4.3.3 Activity of ZL002 and ZL003 as a cryptochrome modulator

A fascinating result was obtained from the ZL002, and ZL003 compounds; contrary to our initial hypothesis, that only planar compounds will be able to mimic FAD and thereby maintain or alter the rhythm, both ZL002 and ZL003 altered the period at higher concentrations. Presumably, the free rotation of the scaffold might allow it to participate in crucial interactions. The extra nitrogen in both compounds might facilitate hydrogen bonding interactions with the polar residues in the vicinity. In terms of clock function, ZL002 delayed the rhythm at 0.9 and 2.7 µM concentrations by 1.7 hours, while ZL003 demonstrated the same behavior at the higher concentration (8 µM). Incidentally, the extra nitrogen along with a chloro group present in the ring system of ZL003 are likely facilitating the binding in the initial concentrations.
Figure 32: ZL002 delays rhythm at the higher concentrations: (A) Structure of ZL002 (B) A graph of luminescence vs. time clearly shows delayed rhythm at the higher concentrations (0.9 µM pink and 2.7 µM, brown and) in the experimental condition (C) A bar graph plotted between period vs. concentration shows the period shift of 2.7 hours at 0.9 and 2.7 µM. (D) Comparison of ZL002 with KL001: KL001 has a better activity at the lower concentration, while ZL002 is active at the higher concentration.
Figure 33: ZL003 delays rhythm at the higher concentrations: (A) Structure of ZL003 (B) A graph of luminescence vs. time clearly shows delayed rhythm at the highest concentration 8 µM (teal) in the experimental condition. (C) A bar graph plotted between period vs. concentration shows the period shift from 25.05 (zero drug) to 26.72. (D) Comparison of ZL003 with KL001: KL001 has a better activity at the lower concentration, while ZL003 is active at the higher concentration.
4.3.4 Activity of ZL004 as a cryptochrome modulator

Pinoline was used as a scaffold as it is a natural biological metabolite of melatonin. Thus, if it demonstrates physiological effects on the circadian period, it could be an endogenous ligand of CRY that modulates activity. Remarkably, Pinoline without any side chain attached had a period altering effects at the higher concentration. Notably the period delay of 1.7 hours was seen at 2.7 and 8 µM concentrations. More importantly, the result indicated that Pinoline is able to bind at the flavin binding pocket of the CRY and can potentially be tested for its ability as a natural therapeutic agent.

![Diagram of Pinoline and its period altering effects](image)

**Figure 34: Pinoline, a natural metabolite shows period altering effects:** (A) Pinoline showed period lengthening effects at higher concentrations, robust rhythm was observed at all the concentrations indicating the pinoline binding at the FAD binding site. (B) The period lengthening of about 1.7 hours were observed at 2.7 and 8 µM concentrations.

Intrigued by the results we tested the pinoline scaffold attached with the KL001 like side chain. The compound showed robust rhythm at all the concentrations used in the experimental
settings. The period lengthening effect of 1.7 hours were observed at 0.9 and 2.7 µM concentrations. Particularly, the drug showed robust rhythm indicated by the higher amplitudes in the fig 35B implying binding ability of the compound. These latter observations are particularly intriguing. The period lengthening effects are consistent with docking to the FAD binding pocket to delay CRY degradation, but in contrast to KL001, which suppresses clock function, here we observe higher amplitude rhythms consistent with enhanced clock activity. Motivated with our initial results, we plan to test the activity of the compound at the higher concentrations.

Figure 35: ZL004 maintains robust rhythm while delays it at the higher concentrations: (A) Structure of ZL004 (B) The curve fitting graph depicts robust rhythm was maintained at all the concentrations with rhythm delaying affects seen at the higher concentrations. (C) The period delaying effect can be clearly visible at 0.9 µM and 8 µM. (D) Comparison between KL001 and ZL004: At the 0.9 µM concentration period altering effect of ZL004 is almost similar to that of KL001.
4.3.5 Activity of ZL005 as a cryptochrome modulator

In ZL005, benz-quinoxaline was used a scaffold with the belief that it will mimic isoalloxazine ring efficiently, owing to its four-ring structure. Interestingly, similar to the Pinoline scaffold, benz-quinoxaline ring system maintained robust rhythm at all concentrations, with higher amplitudes, as shown in fig 36B. More specifically, the compound demonstrated period lengthening effects at the higher concentration which clearly indicates binding at the flavin binding pocket in CRY. The period was significantly delayed by 1.7 hours at the 0.9 µM to 8 µM concentrations.
Figure 36: ZL005 sustained robust rhythm and delayed it consistently at the higher concentrations: (A) Structure of ZL005 (B) A curve fitting graph shows robust rhythm which gets delayed at the higher concentrations (yellow, teal). (C) A bar graph shows 1.7 hours period shift from 0.9 µM to 8 µM concentrations (D) Comparison of ZL005 and KL001: The period delaying effect of ZL005 is almost similar to that of KL001 at the 0.9 µM concentration.

4.3.6 Activity of ZL006 as a cryptochrome modulator

In the sixth compound the side chain and the linker were modified, while carbazole was used a side chain. The result was particularly intriguing, despite ZL006 retaining KL001 like carbazole scaffold, ZL006 did not alter the rhythm indicating the compound did not bind the flavin
binding pocket. The result indicated that the side chain and the linker also play a critical role in the efficient binding of a drug. In future, we plan to test this compound for its activity in the higher concentrations to gauge if the concentration plays any role in the binding ability of the compound.

Figure 37: ZL006 exhibits robust rhythm with no period alteration affects: (A) Structure of ZL006 (B) A graph of luminescence vs. time shows robust rhythm maintained at all the concentrations of the drug. (C) A bar graph plotted between period vs. concentration shows no period shifts. (D) Comparison of ZL006 with KL001: KL001 has a better activity at the lower concentration, while ZL006 has no period altering effects.
4.3.7 Activity of ZL007 as a cryptochrome modulator

For ZL007, the carbazole moiety was retained, however, the side chain and the linker had been changed in ZL007. The compound shows no period altering activity, indicating side chain might not able be to snuggle close to the tryptophan residues owing to shorter linker used. We plan not to use this compound for further studies.

Figure 38: ZL007 shows robust rhythm with no period alteration: (A) Structure of ZL007 (B) A graph of luminescence vs. time shows the rhythm maintained at all the concentrations of the drug. (C) A bar graph plotted between period vs. concentration shows no period shifts. (D) Comparison of ZL007 with KL001: Kl001 has a better activity at the lower concentration, while ZL006 has no period altering effects.
4.3.8 Activity of ZL008 as a cryptochrome modulator

In ZL008, an amide linker and a thiazole side chain are used as a side chain while keeping the carbazole scaffold similar to KL044 and KL001. The compound lacks any period altering effect implying lack of binding at the flavin binding pocket, which is intriguing as carbazole had been shown to be essential for the efficient binding in the flavin binding pocket. The result of ZL008 along with that of ZL006 indicates the side chain of the compound complements the binding ability of the carbazole moiety. For instance, in KL044 the electron withdrawing substituents in the side chain aromatic ring showed to play an essential role in binding. In future, we plan to use substituents in the side chain moiety (indeed, promising results obtained with the tenth compound in the library ZL010).
4.3.9 Activity of ZL009 as a cryptochrome modulator

ZL009 is the novel compound in the library, here the benzquinaxaline ring is coupled with the side chain previously used in ZL006. It was intriguing as the compound substantially
lengthened the period at the 2.7 μM concentration. Moreover, the result was captivating as the rhythm is delayed by 3.3 hours similar to that of KL001. Most importantly, the compound not only shifts the circadian period substantially, but also leads to the amplified rhythms (fig 40B), particularly, at the 2.7 μM concentration indicating possibly different mode of action.

Notably, the four-membered scaffold ring system was selected such that it can overlap or mimic the isoalloxazine ring of flavin, while side chain ring system was similar to that of the tryptophan present in the side-chain binding pocket. The similar scaffold used in ZL005 (discussed above) with the KL001 side chain was also able to delay rhythm by 1.7 hours, on the contrary the same side chain used in ZL006 did not alter the rhythm. Further, oscillations were maintained with significantly enhanced circadian amplitudes. Such affects are particularly intriguing as decline in circadian period have been associated with aging, and age-related physiological impairment. This observation enabled us to conclude that possibly, the four-membered scaffold plays a critical role in a period altering abilities.
Figure 40: ZL009, a novel compound shows both robust rhythm and the period lengthening effects: (A) Structure of ZL009. (B) The novel compound maintains robust rhythm at all the concentrations, period lengthening observed at 2.7 µM. (C) The bar graph shows a significant delay in the period at 2.7 µM concentration. The period delay is about 3.3 hours similar to that of KL001. (D) Comparison of ZL009 with KL001: The period lengthening effect at 2.7 µM concentration is similar to that of KL001.
4.3.10 Activity of ZL010 as a cryptochrome modulator

The tenth compound in the library was synthesized with the insights taken by KL044. The structure-activity analysis of KL044 indicated that the substitutes present in the side chain affects the binding efficiency of the drug. In ZL010 polar substituents are attached in an aromatic side chain with an amide linker and a carbazole scaffold. It was intriguing that the compound demonstrates considerable period shifts of 1.7 hours at the higher concentrations but results in ablation of circadian amplitude. Mechanistically, these results are consistent with those observed for KL001. Combining observations from molecules like KL001, ZL010 (amplitude lowering) and ZL009/ZL004 (amplitude enhancing) may facilitate greater understanding of the complexities of CRY regulation as it pertains to clock maintenance. The results are captivating, and we plan to use the compound for the further analysis.
Figure 41: ZL010 showed robust rhythm at the lower concentration, while period lengthening effects at the higher concentration: (A) Structure of ZL010 (B) A graph of luminescence vs. time shows a delayed rhythm at the higher concentrations 2.7 and 8 µM of the drug. (C) A bar graph plotted between period vs. concentration shows period shifts by 1.2 hours (D) Comparison of ZL010 with KL001: KL001 has a better activity at the lower concentration, while ZL010 showed period lengthening effects at the higher concentration.

4.4 Conclusion

Although further studies are required, based on our results, we can safely conclude that scaffold plays a critical role in the period altering effects. To our delight, the hypothesis of introducing polar interactions in the isoalloxazine binding site proved to be essential for the
efficient binding. Notably, the introduction of polar moieties in the scaffold was able to maintain robust rhythm at the lower concentrations, while some of them like ZL001, ZL004, ZL005 showed period altering effects at the higher concentrations. To rationalize the finding, the four-membered ring system altered the period similar to that of KL001.

Furthermore, substituents present in the side chain aromatic ring system facilitate the period lengthening effects. For instance, in the Zoltowski library none of the heterocyclic side chain used were able to alter the period. However, the trimethoxybenzene side chain delayed the rhythm by 2.7 hours. It was intriguing that the tryptophan like side chain used in the ZL006 with the KL001-like carbazole scaffold depicted no period altering effects, whereas the same side chain in ZL009 with the benz-quinoxaline ring system showed activity similar to that of KL001.

Even though none of the compounds in the Zoltowski library was able to delay the rhythm like KL001, but interestingly none of them had utterly killed the rhythm. Some of the compounds in the library can be used for biophysical studies of the signaling mechanism of the cryptochrome.

With these results in hand, we plan to synthesize different analogs of the four-membered ring scaffold with the polar side chain linked with an amide linker. Moreover, the compounds which have shown period altering activity, we plan to test their activity at the higher concentrations. Additionally, we plan to perform degradation assays to determine the toxicity of these drugs.

4.5 Long-term goals

The circadian clock is the prime determinant of all the physiological behavior, given that most of the clock related dysfunctions are chronic hence specificity of the drugs is required to avoid the potential side effects. The long-term goal of this research is to design the site-specific
drugs for the critical circadian proteins and individual interactions. Phenotype-based screening of FDA approved chemical library (LOPAC) had led to explore several small molecules having a therapeutic effect such as KL001\textsuperscript{4,8-11}. Since the circadian clock is so intricate a target-based drug design is needed to avoid fatal side effects. For instance, mammals have two cryptochromes CRY1 and CRY2. Since only CRY 2 has been found to be associated with breast cancer,\textsuperscript{11} therefore, site-specific CRY 2 inhibitors could eliminate the off-target effects and thereby the side effects. Furthermore, the research can be extended to target individual interactions such as cryptochrome: androgen receptors which are associated with prostate cancer. Other significant interactions exist such as CRY:myc, CRY:HIF1\textsuperscript{12}.

In recent years, X-ray crystallography has provided a vast platform for the site-specific drug designing. A high-resolution crystal structure of a protein or a protein complex can provide better insight into the binding pocket; thereby a site-specific drug can be designed. As we saw earlier, KL001 was found to be a potent drug out of several LOPAC screened, but on analyzing the crystal structure of CRY2 in complex with KL001 or with FBXL3, they were able to determine the mechanistic basis of drug action. Crystal structure revealed the information about the binding site which led the same group to synthesize yet another potent drug KL044.

The long-term goal of our laboratory will be getting the crystal structure of the proteins, and individual complexes and from that follow the site-specific drug designing strategy.

4.6 Materials and methods
All the cell culture flasks and 96-well plate were ordered from Fischer Scientific. The U2OS cells harboring \textit{Bmal-dLuc} reporter was obtained from ATCC (The American type culture collection). All the reagents Dulbecco’s Modified Eagle Media (DMEM) composed of L-glutamine, HEPES, sodium bicarbonate. (Gibco 21063029), fetal Bovine Serum (FBS) and antibiotics
(penicillin/streptomycin, 100U/ml), luciferin, 2% B27 were ordered from Fischer scientific until otherwise reported.

4.6.1 Cell Culture

Composition of Medium: (Adapted from Hirota and Kay method).

A sincere thanks to Samuel G. Weber for the media preparation and cell plating before every experiment.

**Culture media:**

<table>
<thead>
<tr>
<th>Medium Description</th>
<th>Volume</th>
</tr>
</thead>
<tbody>
<tr>
<td>DMEM (high glucose, no phenol red)</td>
<td>450 mL</td>
</tr>
<tr>
<td>10% FBS</td>
<td>50 mL</td>
</tr>
<tr>
<td>100unit/ml Penicillin and 100ug/ml streptomycin</td>
<td>5 mL</td>
</tr>
<tr>
<td>1mM Sodium pyruvate</td>
<td>5 mL</td>
</tr>
</tbody>
</table>

**Explant media:**

<table>
<thead>
<tr>
<th>Medium Description</th>
<th>Volume</th>
</tr>
</thead>
<tbody>
<tr>
<td>DMEM (high glucose, no phenol red), containing 100unit/ml Penicillin and 100ug/ml streptomycin, HEPES, 1mM Sodium pyruvate</td>
<td>12 mL</td>
</tr>
<tr>
<td>2% B27</td>
<td>2 µL (added just before the experiment)</td>
</tr>
<tr>
<td>Gentamicin</td>
<td>0.1 mg/mL (added just before the experiment)</td>
</tr>
<tr>
<td>Luciferin</td>
<td>1 mM (added just before the experiment)</td>
</tr>
</tbody>
</table>

U2OS cells harboring *Bmal-dLuc* reporter were maintained in the culture media at 37 °C. After trypsinization of the cells, they were seeded in the culture media at the density of 2000 cells/well in an opaque 96-well plate and incubated at 37 °C for two consecutive days for allowing U2OS to reach to the highest confluency. At the same time, a transparent 96-well plate was seeded
to monitor the cell growth under the microscope. After two days of incubation, the media was replaced by an explant media containing B-27 supplement to reset the clock of the cells. The explant media was also supplemented with luciferin.

4.6.2 Drug Screening

All the drugs were dissolved in DMSO. The drugs were introduced in each well by keeping the DMSO concentration same. KL001 was used as a reference drug, while some of the wells were only supplied with DMSO to observe the effect of the neat DMSO. The plate was then covered with a thin transparent film. The rhythm was monitored using BioTek Cytation 5 plate reader. The data was collected for the five consecutive days (120 hours) at the temperature 35 °C with the integration time of 14 seconds and interval time of 1 hour 40 min. The data obtained was used for the curve fitting data analysis using R programming script.
REFERENCES


3. Oshima,T; Yamakana, I; Kumar, A; Yamaguchi, J; Ohkawa,T. N; Muto, K; Kawamura, R; Hirota, T; Yagita, K; Irle, S; Kay, S. A; Yoshimura, T; Itami, K. C-H activation generated period-shortening molecules that target cryptochrome in the mammalian circadian clock. Angew. Chemie 127, 7299-7303 (2015).


CHAPTER 5
LOV DOMAIN PROTEIN LKP2 IN A. THALIANA AND B. RAPA

5.1 Introduction

Plants being sessile have developed robust sensory mechanisms to gauge the surrounding environment and adapt to facilitate their growth and development. Out of all environmental cues, light is the most important source which plants sense to dictate the developmental process\(^1\). Plants not only use light for the photosynthetic process but also to regulate several developmental and metabolic processes such as photoperiodic flowering, hypocotyl elongation\(^2\), stomatal opening\(^3\), and phototropism\(^4,5\), and chloroplast movement\(^6\). In order to do so, plants employ several photoreceptor proteins which sense broad spectrum of light. The major classes of photoreceptors in plants are the blue light photoreceptors, phototropins and cryptochromes, and the red/far-red photoreceptors phytochromes\(^7,8\). All photoreceptors can sense various wavelength and intensities of light and assist the plant in adjusting to the surrounding environment.
Figure 42: Different photoreceptors in the plant sense light and regulate several physiological processes

All these photoreceptors are well studied at the molecular level for their role in regulating plant physiological processes\(^9\). Incidentally, in the year 2000 a new kind of the blue light photoreceptor family named the ZTL family was identified in *Arabidopsis thaliana*. The ZTL family in *A. thaliana* includes three members: Zeitlupe (ZTL), flavin binding kelch repeat F-box1 (FKF1) and LOV kelch protein2 (LKP2)\(^{10,11}\). All these proteins belong to the LOV domain superfamily of blue-light photoreceptors. The members of ZTL family were found to work in concert to measure the day length and hence to determine the flowering time and to regulate the other circadian processes\(^{12}\). To date, the function of ZTL\(^{10,15,18}\) and FKF1\(^{10,16,17}\) is well studied in *A. thaliana*. However, not much is known about the function of the third member of the family LKP2. This chapter will mainly focus on characterization and understanding the photocycle of LKP2 in *A. thaliana*. In addition, we extend our research to study the important circadian rhythm proteins in the Brassicaceae family specifically *Brassica rapa*, which is an essential agricultural
crop that includes turnip, Napa cabbage, Chinese cabbage etc. Since circadian rhythm controls several developmental processes, studying crucial circadian protein in important agricultural crop can potentially impact the agricultural industry.\(^\text{19}\)

### 5.2 LOV Domain

LOV (Light Oxygen Voltage) domain are blue light photoreceptors, originally discovered in the plant phototropins. LOV proteins function as a subset of the Period-Aryl hydrocarbon receptor nuclear translocator-Singleminded (PAS) domain superfamily, where LOV domains non-covalently binds the flavin cofactors (FAD or FMN) as a photo-responsive element.\(^\text{10,13,14}\). The PAS fold consists of 5 antiparallel beta strands 2-1-5-4-3 flanked by a series of four \(\alpha\)-helices on one side of the face termed as “A face”. The A-face acts as the photoreceptor core as it binds with the flavin cofactor within a cavity formed by typical \(\alpha\)-helices and \(\beta\)-strands arrangement. The B face consists of the other side of the beta strands and are typically involved in protein:protein interactions in PAS and LOV domain proteins. The primary characteristic of LOV proteins are dictated by their flavin chemistry, which plays an important role in determining the activity and signal transduction landscape of LOV domain proteins.\(^\text{20,21}\). In *Arabidopsis thaliana*, LOV domain proteins like Zeitlupe (*ZTL*), flavin binding kelch repeat F-box1 (*FKF1*) and LOV kelch protein2 (*LKP2*) employ small molecule flavin chemistry to measure the day length and hence determine the flowering time and regulate the other circadian processes.\(^\text{18}\).
Figure 43: Schematic diagram of PAS domain fold: PAS domain is composed of 5 antiparallel β sheets (green arrows) flanked by one side series of four alpha helices (pink cylinders).

Figure 44: Crystal structure of the Zeitlupe exhibiting a typical PAS fold.
5.2.1 Flavin chemistry

The photocycle of the LOV domain is chiefly directed by the small molecule cofactor flavin adenine dinucleotide (FAD) or flavin mononucleotide (FMN). Flavin chemistry is further modulated by the neighboring group interactions or steric constrains in the spatial environment, and ultimately the kinetics parameters of the protein\textsuperscript{21,22}. Upon light activation, a covalent adduct is formed between the flavin cofactor and a conserved cysteine residue in the LOV core. The cysteine residue is present on the E\textalpha{} region within a conserved GXNCRFLQ motif. With the formation of flavin-cysteinyl adduct (C4a) a series of electron and proton transfer reactions occur leading to the protonation at the N5 position of the flavin ring. Protonation relays changes in the hydrogen bonding network of the protein starting from the flavin binding pocket to the other regions of the protein affecting inter-or intra-protein interactions. Upon return to the dark, the adduct state undergoes thermal decay back to its ground state that corresponds to dark state with lifetimes varying from seconds to days\textsuperscript{20-24}.

Recently, these variations in the photo-adduct lifetimes was shown to be essential for proper function. For instance, ZTL functions at the day-night transition which is supported by its relatively short photocycle kinetics with a half-life of 2 hours\textsuperscript{18}. On the other-hand LKP2 has a relatively long photo-adduct lifetime with a half-life of approximately six days. Therefore, the photo-adduct lifetime may provide some insight in predicting the biological function of the protein.

5.2.3 LOV photocycle and kinetics

The photo adduct formation reaction can easily be studied in the laboratory setting by observing the shift in the UV-vis absorption spectra. Notably, in the dark state the LOV domain
protein shows the characteristic absorption spectra of oxidized flavin with peaks at 450nm and vibrational bands at 425 nm and 475 nm (Fig. 44C). Upon blue light activation, the isoalloxazine ring of flavin forms a covalent linkage with the thiol moiety of a conserved cysteine. Specifically, blue light activation promotes oxidized flavin at the dark state into the singlet excited state, that undergoes intersystem crossing to form a triplet state. The triplet state, then abstracts an electron from the cysteine and generates a radical pair. Subsequent radical recombination leads to the formation of C4a adduct, which can clearly be observed by a single absorption band peak at 390 nm.\textsuperscript{20,25-28}
Figure 45: Mechanism of the C4a Adduct formation: (A) Upon blue light activation, a conserved cysteine residue in GXNCRFLQ motif of LOV domain participates in the C4a adduct formation with the isoalloxazine ring of flavin (B), (C) The conversion of light and dark state can be monitored by observing the flavin spectra; in the dark state classic flavin peak can be observed at 450 nm with the two shoulder peaks at 425 and 475 nm. In light state only one peak is observed at 380 nm. (D) Blue light activation excites flavin to the singlet excited state followed by the triplet state. Single electron transfer event takes place from the conserved cysteine, which results in the formation of a radical pair. The reduced flavin generates upon radical recombination.
5.2.4 Factors affecting LOV photocycle

There are several factors which dictate the cofactor flavin chemistry and thereby the LOV domain photocycle and the overall structure of the protein. Primarily, there are three factors which modulate the flavin photochemistry. First and foremost, is the position of cysteine in the dark state protein where two different conformers have been observed that differ according to the proximity of the thiol moiety to the C4a position of the isoalloxazine ring. Computational and kinetic studies indicate that the light state of the protein is stabilized if the position close to the C4a adduct is favored, whereas the conformer favoring the position further away leads to a less-stable adduct. Several studies have indicated that introducing residues that sterically bias the cysteine conformer to one state or the other can tune adduct state lifetime²⁰,²³,²⁹-³¹.

The second factor that controls the C4a adduct stability involves conformational changes occurring in the flavin isoalloxazine ring system. C4a adduct formation changes the hybridization of the C4a position from the planer sp² to the tetrahedral sp³ hybridization states, which alters steric contacts to the neighboring residues, mostly on the residues below or above the isoalloxazine ring system. Thereby, electronic and steric effects of those residues govern the stability of the adduct. For instance, Zoltowski et al. showed that the stability of the light state adduct was enhanced when isoleucine was introduced at the flavin re-face in a fungal protein Vivid of Neurospora crassa.²³

The third crucial factor involves hydrogen bonding and the N5th position of the isoalloxazine ring system of flavin. Deprotonation of the N5 proton has been shown to be rate-limiting in adduct decay. Thus, factors which either stabilize or destabilize the N5 proton directly modulate the kinetics of adduct decay. For instance, a conserved glutamine (Q) adopts a position
as such that its amino functionality participates in the hydrogen bonding with the O₄ of the isoalloxazine ring in the dark state. On the contrary, upon blue light illumination; the C4a adduct forms which changes the environment at the active site, forces glutamine to flip 180°. The “flipped conformation” of the glutamine adopts a conformation such that now its carbonyl functionality forms hydrogen bond with the N5th proton, thereby facilitating in proton abstraction. Later in the chapter, we will observe the effect of glutamine flip in the LOV domain photocycle of the LKP2 protein⁴¹,³².

Figure 46: Glutamine Flip: The amino functionality of the glutamine forms a hydrogen bond with the O₄. Upon blue light activation, the C4a adduct formation leads to the change in the hydrogen bonding pattern. As a result, a conserved glutamine flips and orients in such a way that the carbonyl functionality of the glutamine forms a hydrogen bonding with the N5H proton.
5.3 ZTL Family

The Zeitlupe (ZTL) family forms an important subclass of the LOV domain family. They are the blue light photoreceptors and play a significant role in both the plant circadian clock and in mediating photoperiodic responses. The ZTL family includes three blue light photoreceptor proteins: Zeitlupe (ZTL), Flavin bind Kelch repeat 1 (FKF1) and LOV kelch Protein 2 (LKP2). In contrast to the other LOV domain proteins, these proteins consist of two unique domains F-box and Kelch repeats. Structurally, the F-box domain is flanked by the photo reactive LOV domain at the N-terminus and Kelch domains at the C-terminus. The presence of the extra two domains provides the ZTL family a unique feature from the other LOV domain proteins; while the LOV domain acts as a photo sensory domain and perceives blue light, the other two domains are involved in protein degradation and protein-protein interactions. The main function of the F-box containing protein is to maintain the pace of the circadian rhythm by regulating the time of day for specific ubiquitination of substrate and consequent degradation. In this manner, the ZTL family forms a clade of proteins capable of light-dependent targeted degradation\textsuperscript{1,10,11,33}. 
Figure 47: The Schematic representation of domains in ZTL family: ZTL family consists of F-box domain flanked by the LOV domain at the N-terminal and six Kelch repeats at the C-terminal. LOV domain participates in photoactivation, while both F-box and kelch repeats take part in ubiquitination process.

All three members of the ZTL family are known to work in concert to maintain the circadian rhythm and photoperiodic flowering. The current model of the ZTL family is as follows: During dawn blue light activated LOV domain favors Zeitlupe (ZTL) to form a complex with Gigantea (GI), the complex dissociates in dusk and the F-box domain of the ZTL complexes with TOC1 (Timing of Cab 1) and PRR5 and to them for proteasomal degradation. FKF1 on the other hand, forms a complex with Gigantea and regulates photoperiodic flowering by repressing CDF (Cyclin Dof factor)\textsuperscript{11,18}. All these protein target, directly alter gene transcription, thereby modulating clock-controlled gene transcription in a light-dependent manner.
Figure 48: A schematic representation of circadian and photoperiodic clock in *A. thaliana*. ZTL forms a complex with Gigantia (GI) during the day time, at the same time FKF1 forms a complex with GI and degrades Cyclin DoF Factor (CDF), thereby regulate photoperiodic timing. At night, ZTL gets dissociated with GI and degrades TOC1 and thereby represses the transcription of circadian genes. The symbol ─– indicates degradation pathway. (Adapted from: Pudasaini et al. *Biochemistry* 2013, 52, 7150-7158)

Although specific functions of ZTL and FKF1 are known, the function of LKP2 is still ambiguous. Notably, research has shown LKP2 is an important member of the ZTL family and plays a crucial role in maintaining circadian rhythm. Shultz et al. (2001) first identified LKP2 gene and have shown LKP2 functions either within or very close to the circadian oscillator in *A. thaliana*. In the study they found overexpression of LKP2 led to several arrhythmic phenotypes for the multiple circadian outputs both in constant dark and light. Incidentally, the overexpressed mutant showed loss of photoperiodic flowering time and long hypocotyls. Moreover, another
independent study by Baudry et al. found that ztl fkf1 lkp2 triple mutant showed weaker circadian rhythm than the double mutant ztl fkf1 suggesting LKP2 participates in the circadian rhythm process. Mechanistically, what differentiates LKP2 from the other members of the ZTL family is unknown.

Herein, we investigated the thermal photocycle of *Arabidopsis thaliana*. We anticipate the kinetics study can provide important clues about the active site configuration that could possibly decipher the actual role of LKP2 in the plant circadian clock and what differentiates its function from ZTL and FKF1. Further, we will extend our research to study LKP2 protein of *Brassica rapa*, where due to gene duplication, roles of ZTL and FKF1 is undertaken by set of three LKP2.

5.3.1 *Arabidopsis thaliana* LKP2

The PAS domain of the LKP2 shares 77% amino acid sequence identity with Zeitlupe and 67% identity with FKF1. The sequence alignment of proteins in ZTL family revealed, LKP2 differs with other LOV proteins at the position 153, where both other members ZTL and FKF1 as well as all other characterized LOV proteins contain a conserved glutamine residue at this site. As discussed earlier in fig 41, a conserved glutamine flip plays an important role in LOV photocycle and LOV signal transduction and was believed to be indispensable for function.
In LKP2 the polar glutamine (Q) residue is replaced by a hydrophobic leucine (L) residue. Importantly, same substitution abrogates function in all other LOV domains. We therefore, concluded to introduce a point mutation in LKP2 (L153Q) to mutate the leucine residue to the glutamine observed in all other LOV domains. Presumably, the comparison of the kinetic photocycle of both wild type and mutant could help us delineating the signaling pathway of LKP2.

5.3.2 *Brassica rapa* LKP2

Sequence alignment of ZTL family proteins with the F-box harboring LOV domain proteins of *Brassica rapa* revealed only proteins with the sequence similarity to LKP2 and no copies of ZTL and FKF1. Incidentally, evolutionarily, *B. rapa* has undergone genome duplication resulting in complete loss of some of the gene and polyploidization of others. Interestingly, in
Brassica rapa, triplication of LKP2 is observed and it is believed that these additional copies have undertaken the role of ZTL and FKF1. Importantly, all three harbor the leucine at position 153, which should lead to inactive proteins.

Hitherto plant circadian rhythm studies have been done in A. thalaina which had been a model organism for the circadian rhythm research. We extended our research and studied thermal kinetics study of all three LKP2 present in B. rapa termed as BrLKP2-A, BrLKP2-B and BrLKP2-C. Sincere thanks to Shital Kale and Shannon Swisher for providing the BrLKP2-C data. The data reported in Shital Kale’s thesis used wrong temperature values.

5.4 Results and discussions

Reversion of the light state (C4a adduct state) to the dark state (oxidized flavin) was monitored at time interval of 20 min for several days. The data was plotted between absorbance vs. time at different temperature ranging from 29-43 °C. Each data set was collected in triplicates.
The half-life of the protein was calculated by the average of rate constant calculated after data fitting with an exponential decay equation.

\[ y = A_1 \cdot \exp \left( -\frac{x}{t_1} \right) + y_0 \]

**Figure 51:** (A) The spectra of LKP2 protein exemplifies the C4a adduct formation. The dark state (black) at 450 nm with the shoulder peaks at 425 and 475 nm represents oxidized flavin. Blue light bleaching results in a light state peak (red) representing a C4a adduct formation. (B) Reversion of the light to the dark state demonstrate first order kinetics as shown by the absorbance band at 450 nm (red) and 478 nm (black).

To better understand the photocycle of the LKP2 and BrLKP2, Arrhenius and Eyring analysis of dark state recovery was performed and different kinetic parameters such as activation energy and entropy were calculated.

5.4.1 Thermal kinetics of *A. thaliana* LKP2

Thermal kinetics studies indicate that compared to other LOV domain proteins, LKP2 has slower adduct decay rate with the longer half-life \((t)\) of 137 hr., while that of mutant is reduced to
7 hours. The observation clearly shows the importance of glutamine (Q) in the active site of the LOV domain proteins. Furthermore, the Eyring and Arrhenius analysis of LKP2 revealed strong temperature dependence of LKP2 owing to the higher enthalpy of activation (115.9 KJ/mol). LKP2 mutant (L153Q), on the contrary showed less temperature dependence due to the low enthalpy of activation (99.4 KJ/mol.). Notably, the mutation has caused the enthalpy of activation similar to that of the fast-cycling LOV domain proteins (70-100 KJ/mol.)24.

Figure 52: Thermal kinetics comparison of *A. thaliana* LKP2 and LKP2 (L153Q) mutant: A and C: Arrhenius and Eyring plot of LKP2 (16-165). B and D Arrhenius and Eyring plot of the mutant.
Table 4: Kinetic parameters of *A. thaliana* LKP2 and LKP2 (L153Q) mutant

<table>
<thead>
<tr>
<th>Kinetic Parameters</th>
<th>Wild Type</th>
<th>Mutant</th>
</tr>
</thead>
<tbody>
<tr>
<td>Activation Energy (Ea) (KJ/mol)</td>
<td>115.9</td>
<td>99.4</td>
</tr>
<tr>
<td>Enthalpy (∆H*) (KJ/mol)</td>
<td>113.3</td>
<td>96.9</td>
</tr>
<tr>
<td>Entropy (∆S*) (J/mol.K)</td>
<td>31.1</td>
<td>-3.8</td>
</tr>
<tr>
<td>t (25 °C)</td>
<td>136 hrs.</td>
<td>7 hrs.</td>
</tr>
<tr>
<td>Rate, K (s⁻¹)</td>
<td>2.0 x 10⁻⁶</td>
<td>3.6 x 10⁻⁵</td>
</tr>
</tbody>
</table>

5.4.2 Thermal kinetics of BrLKP2- B and C

5.4.2a Thermal kinetics of BrLKP2 B

The slow cycling BrLKP2-B has a half-life of 27 hours, while the half-life of the mutant is reduced approximately by 10-fold to 2.6 hours. BrLKP2 WT shows weak temperature dependence compared to *At*LKP2, due to relatively low activation energy (87.14 KJ/mol. vs. 115.9 KJ/mol.) that is compensated by an unfavorable entropy (-58.7 J/K vs 31.1 J/K).

The mutant with the half-life of 2.6 hours demonstrated activation energies within the range of fast cycling LOV proteins (70-100 KJ/mole). Moreover, Eyring analysis indicated the low enthalpy of activation is compensated by a large unfavorable entropy (-67.49 J/K)
Figure 53: Thermal kinetics of BrLKP2-B and BrLKP2 (L149Q) mutant: A and C: Arrhenius and Eyring plot of BrLKP2-B (29-165). B and D Arrhenius and Eyring plot of the mutant L149Q mutant

Table 5: Kinetic parameters of *B.rapa* LKP2-B and *B.rapa* LKP2-B (L149Q) mutant

<table>
<thead>
<tr>
<th>Kinetic Parameters</th>
<th>Wild Type</th>
<th>Mutant</th>
</tr>
</thead>
<tbody>
<tr>
<td>Activation Energy (Ea) (KJ/mol)</td>
<td>87.14</td>
<td>78.14</td>
</tr>
<tr>
<td>Enthalpy (ΔH°) KJ/mole</td>
<td>84.57</td>
<td>75.56</td>
</tr>
<tr>
<td>$\tau$ (at 25 °C)</td>
<td>27 hours</td>
<td>2.6 hours</td>
</tr>
<tr>
<td></td>
<td>(measured)</td>
<td>(calculated)</td>
</tr>
</tbody>
</table>
5.4.2b Thermal kinetics of BrLKP2 C

The kinetics parameters of BrLKP2-C were similar to BrLKP2-B where the half-life of the wild type was 29 hours (vs. 27 hours of B-variant).

The BrLKP2-C L→Q behaved similar to the mutant of the L→Q mutant of the B variant with a half-life of 2.5 hours (2.6 hours of B variant). However, the low enthalpy of activation of the BrLKP2-C mutant was compensated by relatively large unfavorable entropy of activation (144.22 J/K vs. -67.49 J/K). The data clearly revealed a far weaker temperature dependence, characteristic of a very low enthalpy of activation compensated by a large negative entropy of activation.

**Figure 54: Thermal kinetics of BrLKP2-C:** A and C: Arrhenius and Eyring plot of BrLKP2-C (29-165). B and D Arrhenius and Eyring plot of the mutant (L149Q).
Table 6: Kinetic parameters of *B. rapa* LKP2 and LKP2 (L153Q) mutant

<table>
<thead>
<tr>
<th>Kinetics parameters</th>
<th>BrLKP2C_WT</th>
<th>BrLKP2C_L2Q Mutant</th>
</tr>
</thead>
<tbody>
<tr>
<td>Activation Energy (Ea) KJ/mole</td>
<td>85.76</td>
<td>55.10</td>
</tr>
<tr>
<td>Enthalpy ΔH ± KJ/mole</td>
<td>88.30</td>
<td>52.58</td>
</tr>
<tr>
<td>Entropy ΔS ± J/mole.K</td>
<td>50.03</td>
<td>-144.22</td>
</tr>
<tr>
<td>τ at 25 °C (Hours)</td>
<td>29</td>
<td>2.5</td>
</tr>
</tbody>
</table>

Based on our studies, we propose LKP2 shows different temperature dependencies which may allow for temperature compensation, temperature sensing pathways as has been indicated for other photoreceptor families. Our results indicated *At*-LKP2 protein demonstrated higher temperature dependencies compared to both the variants of *Br*-LKP2 (B and C). Moreover, both the variants of BrLKP2 (B and C) showed nearly similar temperature dependencies. Notably, the life time of the adduct decay was longer in the *At*-LKP2 than *Br*-LKP2. Furthermore, our studies corroborate that the invariant glutamine in the LOV domain family is critical for functioning as significant difference was seen in the lifetime of the LKP2 photocycle when leucine (L) was mutated to glutamine (Q).

**5.5 Materials and methods**

Cloning and Protein Purification

The nucleotide encoding LKP2 was obtained from The *Arabidopsis* Information Resource (Tair). N-terminal and C-terminal truncated LKP2 construct (16-165) was cloned by PCR amplification. The PCR products were examined using 1% agarose gel followed by the gel purification. The constructs were cloned into 6-His parallel vector with ampicillin resistance using
BamH1 and Hind III restriction sites. LKP2 (16-165) construct was further used to introduce a L153 Q point mutation using Quick change protocol (Stratagene). Both, LKP2 construct (16-165) and the mutant (L153Q) sequence were verified by DNA sequencing (Genewiz). Proteins were expressed in *Escherichia coli* BL21(DE3) cells. Cells were grown at 37 °C until optical density (OD$_{600}$) reached 0.5, at which the temperature was reduced to 18 °C for 30 min. The protein expression was induced by 0.2 mM isopropyl thiogalactoside (RPI). After 22 h, the cell pellets were harvested and stored in 100 mM NaCl, 50 mM Hepes (pH 8.0) 10 % glycerol at 4 °C. LKP2 construct and mutant pellets were purified using Ni-NTA affinity resin (Qiagen) at 22 °C; Washing and elution of protein was done by 20 mM imidazole and 250 mM imidazole, respectively. His$_6$ affinity tag was cleaved by incubating the eluted protein with 2mg/ml TEV protease enzyme (Tobacco Etch Virus Protease) overnight. The residual imidazole and the His$_6$-TEV was removed by buffer exchange followed by another round of affinity purification. The protein was further purified by size exclusion column (superdex S-200) already equilibrated with the 100 mM NaCl, 50 mM Hepes (pH 8.0) 10 % glycerol buffer.

Cloning and Protein Purification of BrLKP2

The nucleotide encoding BrLKP2A, B and C were obtained from Jin A. Kim from the Rural Development Agency, Republic of South Korea. Both N and C-terminal were truncated BrLKP2 construct (29-165) was cloned by PCR amplification. The PCR products were examined using 1% agarose gel followed by the gel purification. The constructs were cloned into GST parallel vector with ampicillin resistance using EcoR1 and Not1 restriction sites. A point mutation of L143Q was introduced using Quick change protocol (Stratagene). Both, the constructs BrLKP2-B, C and their mutants (L153Q) sequence were verified by DNA sequencing (Genewiz). Proteins were expressed in *Escherichia coli* JM109 cell lines. Cells were grown at 37°C until optical density
(OD_{600}) reached 0.5, at which the temperature was reduced to 18 °C for 30 min. The protein expression was induced by 0.2 mM isopropyl thiogalactoside (RPI). After 22 h, the cell pellets were harvested and stored in 100 mM NaCl, 50 mM Hepes (pH 8.0) 10 % glycerol at 4 °C.

BrLKP2 construct and mutant pellets were purified using glutathione affinity resin (Qiagen) at 22 °C. Washing was done by 100 mM NaCl, 50 mM Hepes (pH 8.0) 10 % glycerol. The GST tag was cleaved by incubating the eluted protein with 2 mg/mL TEV protease enzyme (Tobacco Etch Virus Protease) overnight. The cleaved protein was eluted in the buffer, another round of nickel-nitriloacetic acid chromatography was done to remove His6-TEV followed by a final purification with a superdex S-200 size exclusion chromatography.

UV-Vis Absorbance Spectroscopy and Kinetics:

UV–visible absorbance spectroscopy of LKP2 16-165 and LKP2 (L153Q) 16-165 were conducted either on Agilent 8453 spectrophotometer or BioTek plate reader. Spectra for thermal reversion was recorded at the concentrations of 50 µM. All kinetics experiments were performed by illuminating the protein samples with a flood light source (150 W), while incubating the protein on ice to populate the light state. Each data set were done in triplicates to check the consistency in the data collection. After illumination of the dark state protein, spectra were recorded at different times until the absorbance reached to the saturation state. Thermal reversion kinetics of LKP2 and BrLKP2 constructs were obtained at the absorbance of 450 nm and 478 nm with respect to time.

Data fitting was done using mono- and bi-exponential equations as required to extract kinetic parameters. All time constants were reported as 1/ k adduct scission that was averaged between the values obtained at 450 and 478 nm. All other kinetics parameters were calculated based on Arrhenius and Eyring equation.
REFERENCES


APPENDIX I

CRYSTAL STRUCTURE OF CRYPTOCHROME LIKE ALGAL PROTEIN: OTCPF1

A1. Introduction

Light is the prime source of energy in all forms of life ranging from the simplest organisms like photosynthetic algae to the mammals. Algae are the supreme producer of the biomass on the Earth and contribute significantly to the global production and biogeochemical cycling. Just like plants, photosynthetic algae utilize light for the growth and development processes. Light dictates essential physiological process in algae such as photosynthesis, gene expression, and circadian clock entrainment. Recently, several algal genome sequences including diatoms and green algae (such as Ostrococcus tauri and Chlamydomonas reinhardtii) were made available. Interestingly, photoreceptors such as red and blue light photoreceptors including cryptochromes and photolyase type of photoreceptors were found. Phylogenetic analysis of Ostrococcus tauri revealed 5 genes belonging to cryptochrome photolyase family (CPF) family.

The cryptochrome photolyase family1 (CPF1) family consist of the proteins of varied functions. The family consists of three major groups of the proteins: the CPD (cyclobutane pyrimidine dimer) photolyase, (6-4) photolyases and the cryptochrome. While both CPD
photolyase and (6-4) photolyases participate in ultraviolet-induced DNA repair activity, the cryptochrome does not participate in the DNA repair activity. The cryptochrome acts as a photoreceptor protein in plants and insects, whereas in animals it works in a light-independent manner and is a part of the core circadian clock. Interestingly an algal protein OtCPf1 has been found to possess dual functionality; not only it showed the (6-4)-photolyase activity but also interacted with CLOCK and BMAL1 characteristic of the mammalian circadian clock.

We particularly became interested in studying the otcpf1 as it possesses the characteristic of the mammalian cryptochromes. Since photosynthetic alga show robust circadian rhythm and alga being a unicellular or simplest cell to study the function and signaling mechanism of the critical circadian proteins can easily be understood.

A1.2 Protein Crystallization

To obtain better insights of the cryptochrome and its signaling mechanism, we crystalized deletion mutant of OtCPf1 (Δ258-293). We deleted the disordered loop from the otcpf1 full length, to facilitate the protein to crystalize. Hanging drop method was used to crystalize the protein after purification in the reduced environment in the presence of DTT (1,4-Dithiothreitol).

Initially, protein screening was done at the room temperature in the dark room using Hampton crystal screens. However, multitude of nucleation centers were formed and as a result the crystals appeared were not big enough to give a conclusive data after X-ray crystallography. All our efforts to avoid getting multiple nucleation centers such as changing pH of the buffer, reducing the precipitant or the protein concentration did not yield better results. Delightfully, purification and crystallization procedures done at the lower temperature (4 °C) was able to reduce the number of nucleation centers in each well yielding two or three crystals per well. X-ray
crystallography study of the crystals provided us 2.9 Å of data set, which we are in the process of solving.

A2 Results and discussion

Currently, the unfinished data has provided us some insights of the protein structure. The comparison of the OtCPf1 protein with the cryptochrome-2 revealed the presence of similar residues in the active site. Moreover, the residues which are essential for the KL001 binding at the active site (as discussed in chapter 2) are also present (Fig 46 C). We anticipate our success in crystallizing otcpf1 will open new avenues to study the drug binding in the cryptochrome binding site. Possibly, will assist in site-specific drug designing and in the biophysical studies.
Figure 55: Structure of OtCPf1 (A) The cartoon structure of OtCPF1 (B) FAD bound in the FAD binding pocket of OtCPF1. (C) Residues which were found to be essential for the KL001 binding are present and labeled. (D) Side chain of the FAD snuggle close to the tryptophan and glutamine residues.
A3 Future direction.

As OtCPf1 has a structural similarity with the mammalian cryptochrome, therefore, we plan to crystallize the protein with one of the drugs from the ZL library and get the crystal structure to study the structure-activity relationship. In addition, we plan to use the Otcpf1 protein for determining the binding efficiency of the small molecules.

A4 Materials and methods

A4.1 Cloning and Protein Purification

The codon optimized full length Otcpf1 were cloned into pGST parallel vector. The sequence was verified by DNA sequencing (Genewiz). Cells were grown in Luria-Bertani (LB) medium at 37 °C, until the OD₆₀₀ reached to 0.5. The temperature of cells was decreased to 18°C for 45 min. Protein expression was induced by 0.2 mM of isopropyl thiogalactoside (IPTG). Cell harvesting was done after 22 hours; cells were pelleted and were stored in 50 mM Tris (pH = 7.4), 100 mM NaCl and 10% glycerol.

OtCPf1 was purified with glutathione affinity resin (Qiagen) at 4 °C. Bound protein OtCPf1-GST in column was treated with 2 mg of TEV protease per milliliter of resin overnight at 4 °C. 20 mM of Imidazole was added to the eluted protein; an additional round of Ni-NTA chromatography then cleaved His6-TEV. Protein was then incubated with 5 mM Dithiothreitol for 30 min at 25 °C to avoid oligomerization. The size Exclusion Chromatography was then used to further purify protein using 50 mM Tris (pH = 7.4), 100 mM NaCl and 10% glycerol. Crystal tray were set using protein concentration of 6 mg/mL, 9 mg/mL at room temperature and at 4 °C.
Structural Analysis:

Initially, OtCPf1 deletion mutant crystals were obtained with the reservoir solution containing 0.1 M HEPES sodium (pH=7.5) 1.4 M sodium citrate tribasic dihydrate in dark at room temperature at the 6 mg/mL and 9 mg/mL concentration. Crystals were also obtained at 4 °C with the reservoir solution of 1.6 M Sodium citrate tribasic dihydrate pH 6.5 (Hampton screen in HR2-112)

Diffraction data were collected at the F1 beamline at the Cornell energy synchrotron source (CHESS). Data was collected with 20% ethylene glycol as a cryoprotectant. The diffraction data obtained were scaled and reduced in HKL2000. The phase information was obtained by molecular replacement with phaser and phenix and a search model of photolyase (Protein Data Bank code (3FY4). The iterative rebuilding cycle was completed in coot and refinement was achieved by Phenix.
REFERENCES


APPENDIX II.

SCANNED $^1$H AND $^{13}$C NMR SPECTRA