Title
A Genetic Study of The Structure-Function Relationships Underlying Cryptochrome Evolution

Author
Liu, Huachun

Publication Date
2019

Peer reviewed|Thesis/dissertation

UNIVERSITY OF CALIFORNIA

Los Angeles

A Genetic Study of The Structure-Function Relationships Underlying Cryptochrome Evolution

A dissertation submitted in partial satisfaction of the

requirements for the degree Doctor of Philosophy

in Molecular Biology

by

Huachun Liu

2019

ABSTRACT OF THE DISSERTATION

A Genetic Study of The Structure-Function Relationships

Underlying Cryptochrome Evolution

by

Huachun Liu

Doctor of Philosophy in Molecular Biology

University of California, Los Angeles, 2019

Professor Chentao Lin, Chair

How cells respond to light or time is a fundamental question in biology. Cryptochromes (CRYs) are evolutionarily conserved blue light receptors or key components of the circadian oscillator found in major evolutionary lineages, from bacteria to human and have been intensively studied. However, the structure-function relationship of CRYs from evolutionary perspective is unclear. In this thesis, I interrogated the evolutionary roles of universally conserved residues (UCRs) of *Arabidopsis thaliana* cryptochrome 2 (CRY2) (Chapter 2) and developed optogenetic tools by engineering CRY2 using continuous directed evolution techniques (Chapter 3).

UCRs are invariable amino acids evolutionarily conserved among members of a protein family across diverse kingdoms of life. UCRs are considered important for stability and/or function of protein families, but it has not been experimentally examined

systematically. In Chapter 2, I experimentally analyzed 51 UCRs of *Arabidopsis* CRY2 that are universally conserved in eukaryotic cryptochromes from *Arabidopsis* to human. Surprisingly, I found that UCRs required for stable protein expression of CRY2 in plants are not similarly required for stable protein expression of human hCRY1 in human cells. Moreover, 74% of the stably expressed CRY2 proteins mutated in UCRs retained wild-type-like activities for at least one of the photoresponses I analyzed. My finding suggests that the evolutionary mechanisms underlying conservation of UCRs or that distinguish UCRs from non-UCRs determining the same functions of individual cryptochromes remain to be investigated.

CRY2 mainly regulates plant photomorphogenesis through blue-light-specific interactions with numerous protein partners. Such blue-light-specific interactions have been exploited in optogenetics to manipulated biological events in a timely and precisely manner. Chapter 3 focused on the development of a novel pair of blue-light-dependent interacting proteins: CRY2-BIC1 (Blue-light Inhibitor of Cryptochromes 1) and applied PACE (Phage Assisted Continuous Evolution) to increase the dynamic range of CRY2-BIC1 blue-light dependent interaction. I isolated variants of CRY2 with stronger interactions with BIC1 and developed soluble expression and protein-dissociating PACE to facilitate further engineering of CRY2, which could give hints on characteristics of UCRs and non-UCRs.

The dissertation of Huachun Liu is approved.

Steven Erik Jacobsen

Matteo Pellegrini

James Akira Wohlschlegel

Ke Shuai

Chentao Lin, Committee Chair

University of California, Los Angeles

2019

Acknowledgement

I'd like to first and foremost thank my mentor, Dr. Chentao Lin, for his guidance and support during my PhD training. He was always willing to meet with me about any academic or non-academic related problems I brought up to him. Apart from being a mentor, Chentao was also an inspirational idol for me, with extremely strong working ethics, highly efficient working methods, fast and rigorous reasoning, broad vision and optimistic characteristics. Chentao was critical to my success during my PhD. I also want to thank my thesis committee members, Dr. Steven Jacobsen, James Wohlschlegel, Matteo Pellegrini and Ke Shuai. I would like to thank them for their valuable advices on my research projects, taught me to discern which research questions to pursue and which not to. I'm very grateful that they spent their valuable time to help me in everything.

Infinite thanks to current and past Lin lab members for their generous share of their knowledge in both big science, trivial details of protocols and life. I would like to give my special thanks to Dr. Xu Wang and Tiantian Su, for the great enrichment to my PhD life and career by discussion about science and life we had for many times. I'm also grateful to two undergraduate students working with me, Yujia Freda Zhang and Wenyuan Andy Wei, for the enormous amount of work they did in helping my research.

The work in the Chapter 3 was not possible without the extensive knowledge, unpublished protocols and experimental materials about the PACE provided by Dr. David Liu, Ahmed Badran and Michael Packer. I would also like to express my gratitude to visiting PhD students Yifan He and Qianqian Song for their help in analysis of plant phenotypes. I also worked on another continuous directed evolution method

called OrthoRep. Although the preliminary results were not included in this thesis, my work on the system greatly deepened my understanding in directed evolution. I'm grateful for the help of Dr. Chang Liu and Garri Arzumanyan.

I thank Dr. Jeffrey Long for his continuous and sensitive support, without which my PhD life would not be the same. I was privileged to have the best home area director.

Last but not lease, I would like to offer my deepest thanks to my family and friends. My parents, Ximan Liu and Xianhua Jiang, gave me the deepest love, understanding and care. My husband, Fangtao Chi, has always been supporting me academically and emotionally, without the company of which I wound not be able to finish my PhD. My best friends, Yan Helen Yan and Yue Zhang, kept my life interesting and beautiful. Finally, I would like to thank myself, for being persistent and strong.

Table of Contents

List of Figures

Chapter 1

Chapter 2

Chapter 3

List of Tables

Chapter 2

Chapter 3

Vita

Education

- PhD student, Molecular Biology, University of California, Los Angeles, USA

  2014 – Present

- Bachelor of Agronomy**,** College of Agriculture and Biotechnology, Zhejiang University, China 2010 – 2014

Publications

- **Liu H**, Su T, He W, Wang Q, Lin C. The universally conserved residues are not universally required for stable protein expression or functions of cryptochromes. ***Molecular Biology and Evolution***, (In-press).
- Wang L, Ye X, **Liu H**, Liu X, Wei C, Huang Y, Liu Y, Tu J. (2016). Both overexpression and suppression of an *Oryza sativa* NB-LRR-like gene OsLSR result in autoactivation of immune response and thiamine accumulation. ***Scientific reports*** 6, 24079.

Presentations

- **Liu H** (Oral and poster presentation). (Jul 2019). Continuous directed evolution of plant blue-light-specific interactions for optogenetics. *GRS/GRC Synthetic Biology*.
- **Liu H** (Oral presentation). (Nov 2018). Continuous directed evolution of plant blue-light-dependent interactions. *2nd International Conference on Plant Synthetic Biology, Bioengineering, and Biotechnology*.

Awards and Honors

- UCLA Roy and Dorothy John/Fred Eiserling and Judith Lengyel Teaching Excellence Awards 2019
- UCLA Whitcome Fellowship in Molecular Biology 2018 – 2019
- China Scholarship Council (CSC) Scholarship 2014 – 2016

Positions and Employment

- Graduate Student Researcher (MCDB), UCLA 09/2014 – present
- Teaching Associate (MCDB C141 Molecular Basis of Plant Differentiation and Development), UCLA 04/2018 – 06/2018
- Teaching Associate (MCDB 165A Biology of Cells), UCLA 01/2018 – 03/2018
- Teaching Assistant (MIMG 101 Introductory Microbiology), UCLA 09/2017 – 12/2017
- Teaching Assistant (MIMG 102 Introductory Virology), UCLA 01/2017 – 03/2017
- Teaching Assistant (LS 4 Genetics), UCLA 01/2016 – 03/2016

Chapter 1: Introduction

1.1 Plant blue light receptor cryptochromes

1.1.1 Evolution of the CRY/PHL protein family

How cells respond to light or time is a fundamental question in biology. Cryptochromes (CRYs) are evolutionarily conserved blue light receptors or key components of the circadian oscillator found in major evolutionary lineages, from bacteria to human [1–5]. Photolyases (PHLs) and CRYs are related flavoproteins. PHLs are enzymes that repair DNA damage induced by ultraviolet (UV) light, while CRYs regulate photomorphogenesis and photoperiod-control of flowering in plants, and circadian rhythm in both plants and animals. UV light can cause two types of DNA damage: cyclobutane pyrimidine dimers (CPD, or Pyr<>Pyr), where C5 and C6 of a pyrimidine covalently links to the C5 and C6 of an adjacent pyrimidine, respectively, and pyrimidin-pyrimidone (6-4) photoproduct (Pyr[6-4]Pyr), where C4 of a pyrimidine covalently links with the C6 of the adjacent pyrimidine [4]. The DNA damage led by UV light can be repaired by PHLs through a process called photoreactivation, in which PHLs utilize 350-450 nm of blue light to break the pyrimidine dimers induced by UV light. Different types of PHLs can repair different types of pyrimidine dimers: CPD PHLs can repair CPD dimers, whereas (6-4) PHLs can repair (6-4) dimers [6].

CRYs were presumably evolved from PHLs ca. 1000-500 Mya [7]. CRYs lack the ability to repair UV-induced DNA damage, but instead, gain the function to regulate development in plants and circadian rhythm in both plants and animals. CRYs and PHLs share similar primary, secondary, and tertiary protein structure. They both require

non-covalent binding to flavin adenine dinucleotides (FAD) to exert normal functions. All CRYs have two protein domains: the conserved photolyase-homologous region (PHR) and the poorly conserved and intrinsically unstructured CRY-C terminal extension (CCE) domain [8]. By looking at primary protein structure, photolyases differ from CRYs in that photolyases do not contain CCE domain.

PHLs and CRYs belong to the CRY/PHL protein superfamily, and members of the protein family can be found in all 3 kingdoms of life [9]. According to phylogenetic analysis of the CRY/PHL protein family, proteins in the family can be classified into at least 6 groups: CPD PHL, (6-4) PHL, CRY-DASH (*Drosophila*, *Arabidopsis*, *Synechocystis*, *Homo*), plant PHL, plant CRY, and animal CRY [7, 10, 11]. CPD PHL, (6-4) PHL, and plant PHL are photolyases that repair various kinds of UV-induced DNA damage in different organisms. CRY-DASH has both DNA damage repair [12–15] and transcriptional regulation activities [16, 17]. Plant CRYs are blue light receptors to regulate blue-light mediated photomorphogenesis, photoperiod-controlled floral initiation, and other developmental processes [8, 18, 19]. Animal CRYs can be further divided into 2 types: type I animal CRYs are photoreceptors that gate blue light input to regulate circadian rhythm [20]; type II animal CRYs are not responsive to blue light, but instead act as transcriptional repressors to regulate circadian rhythm [21, 22]. Although both plant and animal CRYs regulate circadian rhythm, they were considered to be independently evolved from different PHLs [7]. For example, both plant and animal class I CRYs have the ability to perceive blue light [23], but evidences showed that plant CRYs were evolved from CPD PHLs, whereas animal class I CRYs were more closely related to (6-4) PHLs [7, 24].

1.1.2 CRY-mediated photoresponses in *Arabidopsis thaliana*

In territorial plants, CRYs affect many aspects of plant development at almost every stages of plant life cycle. Plant CRYs were reported to regulate seed dormancy [25], vegetative growth, reproductive growth, and some other developmental and physiological processes. For regulation of vegetative growth, CRYs can mediate blue-light dependent inhibition of hypocotyl elongation [1], cotyledon unhooking, expansion, and unfolding [26], root greening [27], and suppression of leaf senescence [28]. For regulation of reproductive growth, CRYs gate flowering time based on environmental photoperiod [29, 30], and modulate fruit development [31]. CRYs were also reported to be involved in photoresponsive genome transcription [32–34], circadian rhythm [8, 35, 36], temperature sensing [37], gravitropic responses [38], stomata opening and development [39, 40], shade avoidance [41], programmed cell death [42], phototropism [43–45], and magnetoreception [46–48], some of which are important processes taken place throughout the entire life cycle of plants, and this is not even an exhausted list of all functions of CRYs.

Plant CRYs were best studied in the model plant *Arabidopsis thaliana*. The *Arabidopsis* genome encodes 3 CRYs: CRY1 mainly regulates blue-light dependent de-etiolation processes [1][49], CRY2 mainly regulates time of flowering in long-day photoperiods [29][50], while CRY3 is a typical CRY-DASH protein targeted to mitochondria and chloroplasts and repairs CPDs in single-stranded DNA (ssDNA) and double-stranded DNA (dsDNA) [15, 51, 52]. CRY1, CRY2 as well as plant-specific red-

light receptors phytochromes (PHYs), are also involved in regulation of circadian rhythm [35, 53], but the signaling mechanisms were not comprehensively understood.

1.1.2.1 CRY mediated de-etiolation in *Arabidopsis thaliana*

One of the most important developmental turning points in plants is marked by the transition from darkness in soil, to the exposure of light. Plant seedlings undergo major developmental adjustments during the transition. The processes of the adjustments were collectively called de-etiolation. De-etiolation depicts the process of inhibition of hypocotyl growth, cotyledon unhooking, unfolding, expansion, and greening.

Hypocotyl inhibition is one of the most widely used phenotypic read-out of de-etiolation. In darkness, young seedlings are motivated to grow taller to reach the surface of soil to get access of light. When light became available, however, seedlings dramatically slow down the elongation of hypocotyl to repurpose the resources and energy elsewhere [54]. As a result, seedlings grown in darkness undergo skotomorphogenesis with pale and closed cotyledons and elongated hypocotyls, while those grown in light undergo photomorphogenesis, with green and open cotyledons and short hypocotyls [55]. In *Arabidopsis*, CRY1 is the major contributor in inhibition of hypocotyl elongation in blue light. When grown in the blue light, *cry1* mutant seedlings showed long hypocotyls, similar to wild-type seedlings grown in darkness [49]. CRY2 is partially and unequally redundant to CRY1. Overexpression of CRY2 leads to shorter hypocotyls in the blue light [56–59], but the hypocotyl growth inhibition activity of CRY2 is dependent on fluence rate of blue light [60]. The defects in hypocotyl inhibition in *cry2* mutants were almost invisible when seedlings were exposed to high fluence rate

(50 μmol m^{-2} s^{-1}) of blue light. When seedlings were grown in medium fluence rate (5.5 μmol m^{-2} s^{-1}) of blue light, the *cry2* mutant seedlings were obviously taller than wild-type seedlings. When grown in blue light of low fluence rate (0.6 μmol m^{-2} s^{-1}), *cry2* mutant seedlings were taller than *cry1* mutant seedlings, suggesting CRY2 mainly controls inhibition of hypocotyl elongation at low fluence rate of blue light [60]. For both CRY1 and CRY2 proteins, their activities to control hypocotyl growth are dependent on protein abundance [61, 62]. On molecular level, *Arabidopsis* CRYs interact and suppress the COP1 (constitutive photomorphogenic 1) E3 ligase [63], which degrades transcription factor HY5 (elongated hypocotyl 5) [64, 65]. HY5 is a master transcriptional regulator of hypocotyl growth [66–68].

Photoresponse of cotyledons is another major de-etiolation marker [69]. Light controls cotyledon hook opening, unfolding, expansion, and greening. As a dicot plant, each *Arabidopsis* seedling has two cotyledons. In the darkness, two cotyledons of a seedling remain attached together, and the tips of the attached/folded/closed cotyledons do not point upward, but instead point downward and stay close with hypocotyl. The cotyledon hook angle is defined by a line drawn from stem apical meristem (SAM) to the middle of a hypocotyl, and a second line drawn from SAM to bisecting the tips of cotyledons. The cotyledon unfolding angle is defined by lines drawn from the tips of cotyledons to the SAM of a seedling [70]. Cotyledon expansion describes the area of cotyledons. Cotyledon greening describes the chlorophyll content of cotyledons. *Arabidopsis* CRY1 primarily controls cotyledon dynamics at high fluence rate of blue light, while CRY2 mainly functions at low fluence rate of blue light [60]. The regulation

of cotyledon dynamics in response to blue light is not restricted to CRYs. Red-light receptors, PHY, also contribute to blue-light response of cotyledons [70].

1.1.2.2 CRY mediated photoperiodic control of flowering in *Arabidopsis thaliana*

Arabidopsis thaliana is a typical long-day plant. Photoperiod describes the period of time each day during which an organism receives light illumination. A photoperiod where day is longer than night is called long-day (LD) photoperiod, and the reverse is called short-day (SD) photoperiod. The flowering of *Arabidopsis* plants is promoted by LD photoperiod, and inhibited by short-day photoperiod.

The prevailing hypothesis to explain the control of timing of flowering by photoperiod is the external coincidence model [71]. Transcription factor florigen (FT) is one of the most downstream master regulators of flowering [72]. FT is synthesized in leaf but travel to SAM to initiate flowering gene expression. Overexpression of FT promotes flowering regardless of environmental light conditions, and mutation of FT drastically impedes floral initiation [73]. The external coincidence model describes how FT is regulated to control flowering in different photoperiod. In the coincidence model, flowering is controlled by two factors: external photoperiod and internal circadian rhythm. Flowering is only promoted when the two factors match. The two factors modulate flowering time through regulation of CONSTANS (CO) at both transcriptional and post-translational level [74]. CO is a transcriptional activator that controls the expression of FT. Expression of FT is induced when CO proteins are accumulated at high abundance [75].

The regulation of CO protein abundance is achieved by circadian controlled transcription and photoperiod-controlled protein abundance [75]. The mRNA abundance of CO oscillates rhythmically [74]. Transcription of CO remains at low level in the morning regardless of the photoperiod. In the late afternoon of LD photoperiod (corresponds to early night in the SD period), the abundance of CO transcripts is high, whereas in the afternoon of SD photoperiod (corresponds to early afternoon of LD period), the abundance of CO transcripts is low. Such rhythmic pattern is maintained by multiple contributors, including CCA1 (CIRCADIAN CLOCK OSCILLATED 1) [75, 76], LHY (LATE ELONGATED HYPOCOTYL) [77], PPR5 (PSEUDO-RESPONSE REGULATOR 5) [78], FKF1 (FLAVIN-BINDING, KELCH REPEAT, F-BOX 1) [79], GI (GIGANTEA) [80, 81] and CDF (CYCLING DOF FACTOR) [79].

Although transcription of CO is similarly highly active in the late afternoon of LD photoperiod and early night of SD photoperiod, the protein abundance of CO is completely different in the two conditions because of different post-transcriptional regulation [75]. The protein stability of CO is mainly controlled by photoperiod. In darkness, CO proteins are constantly degraded by COP1 (CONSTITUTIVE PHOTOMORPHOGENIC 1)-SPA1 (SUPPRESSOR OF *PHYA-105* 1) complex [82–84], PHYB (phytochrome B) [85], and HOS1 (HIGH EXPRESSION OF OSMOTICALLY RESPONSIVE GENES 1) [86]. On the contrary, CO proteins are stabilized by PHYA (phytochrome A) in the far-red light [85], and CRY2 [85, 87], FKF1 [88–90] in the blue light. Therefore, CO proteins accumulate in the late afternoon of LD photoperiod, but not early night of SD photoperiod. Consequently, FT expression is only activated in the late afternoon of LD period to promoter flowering.

Arabidopsis CRY2 primarily regulates time of flowering in LD photoperiod [29]. Plants of *cry2* mutants showed delayed flowering compared with wild-type (WT) plants in LD, but not SD photoperiod. *Arabidopsis* CRY2 regulates flowering time in LD photoperiod through at least two different pathways [26, 91]. In one pathway, CRY2 proteins are thought to work with PHYB coordinately to regulate flowering time under blue and red light, whilst the other pathway does not involve PHYB and regulates flowering time in monochromatic blue light. The evidences of PHYB and CRY2 working together came from genetic analysis of *phyB*, *cry2*, and *cry2phyB* mutants [26, 30, 91]. Flowering time of *cry2* and *phyB* mutant plants was comparable to that of WT in the red and blue light, respectively. PHYB inhibits floral initiation in the red light since *phyB* mutant plants showed early flowering phenotype. In contrast, the flowering time of *cry2* mutant plants was close to WT in monochromatic blue light, but the plants flowered late when subjected to both blue and red light at the same time, which suggests that CRY2 and PHYB might crosstalk in regulation of floral initiation. Double mutant *cry2phyB* plants resembled *cry2* or *phyB* mutant plants in the blue or red light, respectively. However, *cry2phyB* plants resembled *phyB* plants when applied to both blue and red light, indicating the CRY2 and PHYB proteins working in the same pathway, and that PHYB is at a more downstream position in the pathway. In sum, PHYB inhibits floral initiation in LD photoperiod in the red light. In the blue light, CRY2 inhibits the PHYB-mediated inhibition of floral initiation. As mentioned above, COP1-SPA1 complex, HOS1, and PHYB all contribute to the destabilization of CO. PHYB was found to physically interact with SPA1 [92], HOS1 [93] and CO [93], though the functional significance in degradation of CO proteins is not clear. CRY2 was reported to interact

with COP1-SPA1 to stabilize CO proteins [87]. The detailed molecular mechanism of antagonistical relationship of CRY2 and PHYB is not known.

The second pathway where CRY2 promotes flowering without involving PHYB was discovered through genetic analysis of flowering time of *cry1*, *cry2* and *cry1cry2* mutant plants in monochromatic blue light [26, 91]. Both *cry1* and *cry2* single mutant plants showed time of flowering similar to WT plants in the blue light, whereas *cry1cry2* double mutant plants showed remarkably delayed flowering in the blue light condition, which implies that CRY1 and CRY2 work redundantly to regulate floral initiation. The fact that CRY2 can promote flowering in monochromatic blue light indicates that CRY2 is able to function independent of PHYB. CIB1 (cryptochrome interaction basic helix-loop-helix)-activated FT expression probably explain the PHYB-independent CRY2-mediated floral promotion [59, 94]. CIB1 is a transcription factor that is associated with the E-box of the promoter region and promote the transcription of FT gene. CIB1 physically interacts with CRY2 in a blue-light dependent manner, and its activity in regulating transcription is partially dependent on CRY2 [59]. Evidences also showed that the CRY2-mediated inhibition of the COP1-SPA1 complex is responsible for the PHYB-independent CRY2-mediated promotion of flowering [87].

1.1.2.3 Blue-light induced proteolysis of *Arabidopsis thaliana* CRY2

When plants grown in the darkness were transferred to blue-light conditions, *Arabidopsis thaliana* CRY2 proteins undergo fast blue-light dependent degradation [56, 57, 95]. When such degradation was blocked, CRY2 proteins form photobodies inside of nuclei [96]. It was proposed that such blue-light dependent protein degradation was

a mechanism to remove active CRY2 to desensitize the photoreceptor [97]. It is plausible that the blue-light induced CRY2 proteolysis plays a role in regulation of CRY2 protein abundance based on light input. However, the detailed functions of blue-light induced proteolysis of CRY2 were not reported.

1.1.3 CRY photoactivation and inactivation

1.1.3.1 Photoreduction of CRYs and PHLs

Photoactivation describes sub-, intra-, and inter-molecular events of photoreceptors upon light excitation. Photoreduction of FAD molecules is considered to be one of the mechanisms of CRY/PHL initial photoactivation [98–100]. The isoalloxazine ring of FAD has 3 redox states: oxidized quinone (FADox), one electron reduced semiquinone ($FAD^{\bullet-}$ or FADH), and two election reduced hydroquinone ($FADH^{-}$) [101]. When illuminated with blue light, FADox acquires electrons and is reduced into $FAD^{\bullet-}$ and $FADH^{-}$. This process is called photoreduction.

Photoreduction of FAD in the CRY/PHL protein family was first identified in PHLs isolated from baker's yeast (*Saccharomyces cerevisiae*) [102, 103]. It was discovered that isolated PHLs had greatly enhanced repairing activity if illuminated with blue light prior to mixing with DNA lesion substrates. Isolated PHLs illuminated with blue light were photoreduced from FADox or FADH into $FADH^{-}$, and the $FADH^{-}$ form of PHLs was the active form to repair DNA damage. Three conserved tryptophan residues located from the surface of protein to the deeply buried FAD co-factor, called Trp-Triad residues, were considered to be involved in transferring electrons to FAD [4]. Isolated PHLs carrying mutations in Trp-Triad residues could not be photoreduced, and

therefore had little DNA lesion repairing activities [104]. However, photoreduction of PHLs was found to occur only in vitro, since the aerobic exposure of PHLs during process of protein isolation oxidized PHLs [105]. In vivo, the FAD co-factors buried inside of PHLs were in the fully reduced active form, and hence no photoreduction conversion was required for DNA repairing activities in vivo.

PHLs repair DNA using light as substrate through a cyclic electron shuttle mechanism [106]. Other than the photoreduction process described above, light can elicit two kinds of electron transportation on FAD in PHLs, both cyclic, one is intrinsic electron transfer within FAD and the other is electron transfer from FAD to DNA lesion and then back to FAD [106]. The intrinsic cyclic electron transfer occurred within the folded FAD, between the isoalloxazine ring and the adenine moiety [107]. This intrinsic electron transfer is not directly used to repair DNA damage. Instead, the cyclic electron transfer between FAD molecules and DNA lesion is the photochemistry involved in repairment of DNA [108, 109]. When isoalloxazine ring of FAD is excited, electron can be shuttled to thymine dimers to break the covalent linkage induced by UV light. Depending on the intermediates during the electron transferring process, two different electron transfer schemes were identified to repair DNA: direct electron tunneling where electrons were delivered to pyrimidine dimers though superexchange mechanism, and two-step electron hopping where adenine moiety of FAD served as an intermediate of receiving electrons from FAD before transferring the electrons to pyrimidine dimer [106]. Different photolyases usually used both of the mechanisms, but only predominantly used one of them to repair DNA [106].

Photoreduction can also occur in cryptochromes. FAD molecules of CRY proteins extracted from *Arabidopsis* and *Drosophila melanogaster* (dCRY) are in FADox form [101]. FAD molecules of *Arabidopsis* CRYs can be fully photoreduced into $FADH^-$ form both ex vivo [99] and in vitro [61, 62], and FAD of dCRY proteins can be reduced into $FAD^{\bullet-}$ [110–112]. The relationship between photoreduction activities and physiological functions remained to be debated. Although in some occasions, abolishing photoreduction via mutation of Trp-Triad residues is correlated with reduced or abolished physiological activities [113, 114], there are also examples where abolishing photoreduction did not obviously change physiological functions [61, 62]. For example, both in vivo and in vitro photoreduction was abolished in the *Arabidopsis* $CRY2^{W397F}$ mutant proteins. However, $CRY2^{W397F}$ mutant plants retained normal functions in inhibition of hypocotyl elongation and floral initiation in the LD period [61]. Similarly, replacing each of the Trp-Triad residues with phenylalanine residues in dCRY abolished photoreduction, but not light-induced proteolysis activities [111, 115]. Meanwhile, chemical reduction of dCRY was not sufficient to elicit light responses [116]. Most evidences do not support the hypothesis that photoreduction is required for functions of CRYs or PHLs.

1.1.3.2 Photoactivation of *Arabidopsis thaliana* CRY2

For all PHLs and light sensing CRYs, photoactivation begins with the absorption of light by FAD co-factors. For PHLs, light provides the energy to excite electrons from FAD molecules to break covalently linked thymidine dimers. For cryptochromes, it is hypothesized that light triggers conformational changes by dissociating the CCE domain

from the PHR domain, leading to downstream reactions [117], such as CRY phosphorylation [95, 118], ubiquitylation [57], degradation [56], homooligomerization [119] and interaction with protein partners, including PPK (photoregulatory protein kinases) [120], CIB1 [59], SPA1 [61], HBI1 (HOMOLOG OF BEE2 INTERACTING WITH IBH 1) [121], IAA7 [122], IAA17 [122]. The precise mechanism of light-induced conformational change is not clear. When excited by light energy, PHLs also undergo intrinsic cyclic electron shuttle without changing redox states of FAD, in addition to the cyclic electron shuttle that repairs DNA damage [101]. It is reasonable to speculate that the same intrinsic cyclic electron shuttling process is able to occur in CRYs, and possibly contribute to the change of conformation of CRYs.

As one of the possible results of domain disengagement, homooligomerization of the *Arabidopsis* CRY2 was reported to be an indispensable step in photoactivation of the photoreceptor [119]. The first evidence indicating the importance of homooligomerization of CRY2 was from the study of transgenic plants expression GUS (β-glucuronidase)-CRY2 or GFP (green fluorescent protein)-CRY2 fusion proteins. Plants expression GUS-CRY1/2 CCE (GUS fused with the CCE domains of CRY1/2) fusion proteins underwent constitutive photomorphogenesis [123], while the GFP-CRY2 plants did not show similar effect [96]. GUS is tetrameric enzyme [124] and hence enforces the tetramerization of CRY2 proteins fused to GUS, while GFP only have very weak dimeric activity at very high protein concentration [125]. It was later reported that chemically induced dimerization of the CRY2 effector CCE domain in darkness was sufficient to convey light-induced CRY2 functions [126], which not only demonstrated the importance of homooligomerization of CRY2, but also indicated that it is a blue-light

dependent process. Recent research using CRY2 proteins fused with different epitope tags expressed at similar stoichiometric amounts showed blue-light dependent homooligomerization in both *Arabidopsis* plants and heterologous human HEK (human embryonic kidney) 293T cells [119].

1.1.3.3 Photoinactivation of *Arabidopsis thaliana* CRY2

At least 3 mechanisms were used by the *Arabidopsis* plants to inactivate CRY2. First, since homooligomerization of CRY2 is conceived to be vital for function, it is reasonable to infer that the disruption of CRY2 homooligomerization leads to inactivation of CRY2 signaling. The activation of *Arabidopsis* CRY2 (represented by CRY2-CIB1 interaction) is instant in mammalian cells, with the blue-light dependent interaction of CRY2-CIB1 occurred within 300 ms after blue light illumination [127]. In the heterogeneous system, there was no endogenous *Arabidopsis* proteins or factors to facilitate the reversion of CRY2-CIB1 into monomers, but such reversion could be accomplished spontaneously in darkness within 12 minutes [127]. Second, in *Arabidopsis* plants, inactivation of CRY2 proteins could be assisted by BIC1/2 (blue-light inhibitor of cryptochromes 1/2) [119]. BIC1/2 proteins were inhibitors of CRY2, the overexpression of which resulted in inhibition of CRY2 functions in inhibition of hypocotyl growth in the blue light and promotion of floral initiation in the LD photoperiod. The *bic1bic2* mutant plants led to increased sensitivity to blue light [119]. BIC1/2 proteins physically interact with CRY2 to maintain CRY2 proteins in monomer form. However, it was not clear if it was accomplished through inhibition of homooligomerization, or promotion of dissociation of homooligomers. Third,

homooligomerized CRY2 can be removed by blue-light dependent proteolysis to desensitize the excited photoreceptors [56].

1.2 Directed evolution of proteins

Darwinian evolution describes how spontaneous mutation of genetic materials and natural selection of fitness are combined to shape the diverse phenotypes of biological organisms living in diverse environment. In the past thousands of years, people have also been artificially selecting features that are desirable for human rather than those favor the survival in nature. This artificial selection produces common crops and domesticated animals. Spontaneous mutations serve as the genetic pool of selection for both Darwinian natural evolution and artificial selection. However, the rate of spontaneous mutation in relevant biological organisms is generally low, and therefore the genetic pool for selection of desired phenotype is small. To enlarge the pool of genetically different materials, scientists introduced various mutagenesis methods to achieve higher mutagenesis rate. The chance to select a desired phenotype of an organism is greatly enhance by the diversified starting genetic materials. Laboratory adaptive evolution (LAE) used human induced mutagenesis to select desired individual genomes [128]. For example, the method allows for metabolic engineering of microorganisms, and is frequently used to improve industrial strains for optimized production [129]. In contrast to LAE, directed evolution (DE) uses individual genes, gene products, or pathways instead of using individual genomes as selection unit [130]. Directed evolution focuses on improving or re-purposing functions of individual gene products, and usually requires more focus mutagenesis on the gene(s) of interest,

rather than applying genome-wide random mutagenesis. Both LAE and DE consist of genetic diversification and screening/selection. Rational protein design is also focused on improvement of single gene product(s). It harnesses structural information of the targeted proteins and employs predictions or knowledge of structure-function relationship to design mutations that could possibly improve the protein of interest in the desired function. On the contrary, engineering proteins using directed evolution methods does not require prior knowledge of protein structure or structure-function relationship. Recently, with the rapid progress of the understanding of protein folding, rational protein design becomes increasingly accurate and is more frequently used to complement the brute-force directed evolution methods.

1.2.1 Diversification methods in directed evolution

The success of a directed evolution experiment highly relies on the ability to generate diversified genetic variations, because the number of screened variations is correlated with the probability of occurrence of desired activities. Therefore, diversification schemes are generally designed to cover as much sequence space as possible. However, comprehensive mutagenesis is impossible. For example, complete diversification of a 10-residue peptide includes 10^{13} (20^{10}) mutants [131]. As a result, a more realistic aim for diversification is to more universally and randomly cover the entire sequence space, with unbiased mutagenic spectra.

A wide range of mutagenesis methods with different features were developed to be used in directed evolution. These include in vitro chemical or physical mutagenesis [132], error-prone polymerase chain reaction (epPCR) [133], de novo synthesis of

diversified DNA oligonucleotides [134], DNA shuffling [135], and in vivo mutator strains [136]. Chemical mutagens include base analogs, deamination agents, alkylating agents and others [137]. Physical mutagens introducing base variations is mainly achieved by UV light, although electromagnetic radiation at other wavelengths is also used to generate more dramatic deletion of DNA. Chemical and physical mutagens are more frequently used in adaptive evolution, but there are also examples in directed evolution, such as the use of a common chemical agent, EMS (ethyl methanesulfonate) to mutagenize a regulatory gene xylS [138].

Error-prone PCR is one of the most widely used methods to diversify specific genes [133]. In epPCR, low fidelity DNA polymerases, or reaction buffer that decreases the fidelity of DNA polymerases are used during PCR amplification of gene(s) of interest. Error-prone PCR is highly commercialized, easy to use, inexpensive, and capable of achieving high and customized mutagenesis rate.

De novo synthesis of diversified DNA oligonucleotides is frequently used for focused scanning saturation mutagenesis of a small specific region of a protein [139]. Structure-function relationship knowledge is often used to guide the selection of target domains/motifs or residues for mutagenesis.

DNA shuffling and other DNA combination techniques can combine different mutations within a diverse population or from different population of variants. Most DNA combination techniques consist of fragmentation and ligation steps, and both steps can be sequence-dependent or independent to accomplish different purposes [135]. DNA combination is desirable to concentrate beneficial mutations, to eliminate passenger mutations, and to swap domains from different sources to acquire better-performing

mutants. The combination of beneficial mutations distributed among different copies of DNA may yield better-performing variants compared with variants bearing only single beneficial mutation. Occasionally, passenger mutations are selected because they are linked to beneficial mutations *in cis*. These passenger mutations convey neutral or even detrimental effects to the desired activities. DNA shuffling and other DNA combination techniques can "combine out" those passenger mutations [131].

In vitro mutagenesis is effective, reliable, and easy to be customized based on different goals. However, screening or selection of desired activity is often conducted or includes in vivo steps [140]. In that case, molecular cloning is required to transfer mutant libraries created in vitro into in vivo screening or selection conditions. Molecular cloning is tedious and expensive and always loses part of the mutational diversity during the unavoidable transformation step. It is therefore desirable to create mutator strains, where construction of mutant library can be performed in vivo to circumvent the drawbacks of in vitro mutagenesis. Several *Escherichia coli* (*E. coli*) mutator strains, such as XL1-Red [141], are created by compromising endogenous DNA repair machinery. These mutator strains can increase rate of spontaneous mutation by ca. 5000 fold [141]. However, the elevated mutation rate is not focused on a single evolving gene. Concurrent genomic mutations may interfere screening/selection of the evolving genes, either through decoupling the activity of the evolving gene product from screening/selection scheme, or through breakdown of the supporting cellular machinery.

Therefore, it would be helpful to have orthogonal mutagenesis systems capable of targeted mutagenesis in vivo. The following are two examples of in vivo orthogonal

mutagenesis systems developed in *E. coli* [142] and yeast (*Saccharomyces cerevisiae*) [143] via using bacteriophage and orthogonal DNA replication system, respectively.

The *E. coli*-bacteriophage based system is called PACE (phage-assisted continuous evolution) [142]. To generate genetic diversity, PACE used a mutagenesis plasmid (MP) expressing mutator proteins that compromise DNA repair and fidelity-maintaining machinery when induced by arabinose. To select for desired activities, activities are coupled with propagation of M13 filamentous bacteriophage, by linking the activity of interest to the production of a M13 phage protein III (pIII; encoded by gene III, gIII), which is critical for M13 phage infection. For example, in a PACE system selecting for stronger protein-protein interaction (PPI) affinity (Fig. 1-1A) [144], strong PPI leads to transcription of gIII via a bacterial-two-hybrid system and hence support M13 phage propagation, while weak PPI cannot support phage propagation. To achieve the above scheme, in addition to the MP, the accessory plasmid (AP) and the selection phage (SP) are employed. SP is built based on M13 phage genome, by replacing gIII with the evolving gene-of-interest. The AP contains the reporter system that links activity-of-interest to the production of pIII. Basically, the system exploited the continuous culture and selection of phages (Fig. 1-1 B and C). *E. coli* host cells are made to continuously flow through a vessel (called lagoon) containing a replicating population of M13 phage, the genome of which is SP. Fresh *E. coli* cells flowing into the lagoon only contains MP and AP. SP is introduced into the *E. coli* cells through phage infection. Since the replication of SP is accomplished in *E. coli* cells, the evolving gene on SP is subjected to mutagenesis by MP. If the mutated evolving gene product is not able to stimulate production of pIII, the resulting phage progeny is non-infectious, and will be washed-out

from the lagoon by the continuous flow of fresh *E. coli* cells. Otherwise, the phages can persist in the lagoon. The MP can unbiasedly introduce mutations on any DNA replicating in *E. coli* cells. The orthogonality of mutagenesis is achieved by the constant flow of *E. coli* cells, since arabinose that induces mutagenesis is only added in the lagoon, but not in the container culturing host *E. coli* cells. In the constant flow, the time for *E. coli* cells staying in the lagoon is longer than the time for M13 phage replication, but shorter than the doubling time of *E. coli* cells, so only mutations accumulated on SP will affect selection results.

Chang Liu and colleagues developed a orthogonal mutagenesis system in yeast cells [139, 143], called OrthoRep (orthogonal replication) (Fig. 1-2). The system employs cytoplasmic linear plasmids derived from *Kluveromyces lactis*. When introduced into *Saccharomyces cerevisiae*, the linear plasmids exploit their own DNA replication machinery rather than using the endogenous machinery that is used to replicate genomic DNA. Therefore, different DNA polymerases (DNAP) are used to replicate the linear plasmids and cellular genome, and the DNAPs are orthogonal to each other. Orthogonal mutagenesis is achieved by using low-fidelity DNAP to replicate the linear plasmids and high fidelity DNAP to replicate genome.

1.2.2 Screening and selection methods in directed evolution

Screening and selection are integral parts of directed evolution. Screening is to gather the information of each members of a library, and choose some members meeting certain criteria for future rounds of diversification and screening/selection. Selection, however, does not collect quantitative information for members of a library.

Instead, the method classifies individual variants into two group: those that meet the selection criteria, and those do not. Only variants meeting the selection criteria will be propagated and further diversified for future rounds of directed evolution. Compared with selection, more information about the library can be acquired by screening to better inform the stringency of future screening. The cost of the detailed information is relatively low throughput compared with selection.

Some activities can be screened or selected by using the activity of interest, or surrogate substrates as direct readouts. For example, in directed evolution experiments to produce better-performing GFP, the desired activity can be screened directly by using fluorescent signals as readout [145]. The evolving fucosidase activities can also be screened by directly observing the color change of chromogenic substrates X-Fuc (5-bromo-4-chloro-3-indolyl b-D-fucopyranoside) [146]. However, a significant portion of the desired activities cannot be easily screened or selected directly. These activities are frequently coupled with transcriptional outputs to enable efficient screening or selection. For example, when evolving glycosynthase activity using a yeast three-hybrid system, glycosynthase activity was coupled with the transcription of the auxotrophic reporter *LEU2* by reconstituting a transcriptional activator through covalent bond formation of heterodimeric small molecules linked to AD and DBD of a transcription factor, respectively [147].

1.2.3 Continuous directed evolution

Traditionally, the majority of directed evolution efforts are noncontinuous (stepwise), frequently requiring human operations for rounds of diversification and

screening/selection. Therefore, stepwise directed evolution is laborious, time-consuming, and can introduce inconsistent factors by human manipulation. Continuous directed evolution describes methods that automate diversification and selection. Reduced human intervention can greatly accelerate directed evolution process.

Continuous directed evolution requires particular mutagenesis and selection schemes [148]. It is very difficult, if not impossible, to use screening in continuous directed evolution. Three strategies are used to achieve the automation of mutagenesis and selection: 1) in vitro mutagenesis coupled with in vitro selection; 2) in vivo mutagenesis coupled with in vivo selection, enabled by automated culture dilution; 3) automation of the processes of conversion of the DNA mutant library between in vivo and in vitro status, such as transformation and library DNA extraction processes.

In vitro continuous directed evolution was reported to evolve RNA ribozymes that catalyze template-directed RNA ligation [149]. This is accomplished in 4 steps: 1) RNA substrates containing T7 promoter sequences and sequences complemented with the evolving ribozymes are provided. 2) The substrates are ligated to the 5' end of the ribozymes if the ribozymes had the desired functions. 3) DNA primers are annealed to the 3' end of the ribozymes, and the ribozymes are reverse transcribed from the DNA primers. Meanwhile, mutations are introduced into ribozymes, because reverse transcriptase (RT) cannot proofread. 4) If the ligation is successful in step 2, the reverse transcription DNA products would possess T7 promoter. T7 RNA polymerases (RNAP) then produce the ribozymes from T7 promoter, for next round of mutagenesis and selection.

In vivo continuous directed evolution systems have been constructed in prokaryotes [142, 150], lower eukaryotes (yeast) [143, 151], and higher eukaryotes (mammalian cells) [152–154]. In vivo orthogonal mutagenesis is the key to such systems, because it avoids accumulation of mutations in the background genomes to enable continuous culture. The PACE system is the first developed in vivo continuous directed evolution system [142], the working mechanisms of which was detailed above. EvolvR [150] is another in vivo continuous evolution system constructed in *E. coli* cells. To achieve orthogonal/targeted mutagenesis, the system employs CRISPR (clustered regularly interspaced short palindromic repeats)-guided nickase fused with nick-translating DNAP. The targeted sequence is first nicked, and then the nick-translating error-prone DNAP synthesizes mutated DNA strands, with a mutational window with length of up to 350 nucleotides. EvolvR has been used to evolve resistances to antibiotics spectinomycin and streptomycin, conferred by 30S ribosomal protein S5 (encoded by *rpsE* gene) and S12 (encoded by *rpsL* gene), respectively. Theoretically, by changing gRNA (guide RNA), EvolvR can easily evolve any genomic gene(s) *in situ*. Engineering of industrial *E. coli* strains usually involves optimization of genes on plasmids, followed by incorporation of the genes into the genome of industrial strains. However, it is very likely that the genome integration of the genes may change their behavior. This is where the EvolvR system comes in handy, because it engineers a gene at the gene's endogenous genome context. Additionally, the design of EvolvR has the potential to be transferred into other organisms.

In vivo continuous directed evolution systems in lower eukaryotes are exemplified by the OrthoRep [139, 143] and ICE (in vivo continuous evolution) [151]

systems, both of which were constructed in yeast (*Saccharomyces cerevisiae*). The ICE system places the evolving genes on a Ty1-based transposon, and the transcription of the transposon is under the control of an exogenous promoter. After the transcription of the transposon into mRNA, reverse transcriptase (RT) is produced from the mRNA, and reverse transcribes the transposon RNA, introducing mutations into the evolving gene(s) at the same time because of the error-prone nature of RT. The DNA resulted from reverse transcription is incorporated into genome and the evolved genes are expressed to enable selection.

In vivo continuous directed evolution in unicellular organisms heavily depends on continuous dilution of the cultures of biological organisms. The EvolvR, ICE and OrthoRep systems all required unautomated dilution of cultures per 16 hours [150] or 3-4 days [139, 151]. In the PACE systems, automatic continuous dilution of cultures is achieved using a sophisticated and expensive pumping system, which is a hurdle for the system to be adopted by other laboratories. A noncontinuous version of the PACE technique, PANCE (phage-assisted noncontinuous evolution), was developed to reduce the dependency on the pumping system, but it requires dilution by human per 8-12 hours[155]. Ahmad S Khalil and colleagues developed an automated and generalizable cell culture growth and dilution systems called eVOLVER [156]. The system employs standardized and open-source wetware, hardware, and network-based tools, to enable easy adoption by a broad community of users. The eVOLVER is composed of 3 parts: fluidic modules, electronic modules, and web-based software. The fluidic modules include multiple parallel smart sleeves, which are small containers to culture cells. Each smart sleeve can independently monitor and control the optical density, temperature,

stir rate and light conditions of the cells. The fluidic modules also include automated culture dilution modules that prevent the usage of expensive pumping systems. The electronics modules consist of the control modules of smart sleeves and Raspberry Pi that communicates with computer or cloud server and sends instructions to command modules. Web-based software can be used to remotely control parameters of individual smart sleeves. The eVOLVER technique significantly improves the number of parameters and conditions that can be tested at the same time and reduces human effort in cell culture growth and dilution. It has the potential to dramatically increase the efficiency of in vivo continuous directed evolution.

There are two types of mammalian in vivo continuous directed evolution systems: dCas9 (CRISPR associated protein 9 endonuclease dead)-based [152, 153] and mammalian virus-based [154] systems. The dCas9-based systems, targeted mutations are introduced by cooperation of dCas9 and cytidine deaminase that deaminates cytosine (C) to uracil (U). The association between dCas9 and cytidine deaminase is achieved either via direct protein fusion [153], or via fusion of gRNA with a RNA hairpin capable of recruiting MS2 protein-fused cytidine deaminase [152]. Target proteins were evolved to acquire resistance to cancer therapeutics. However, the editing window is somewhat limited, because 7-10 gRNA is needed to cover 80% of C/G bases within 100 base pairs [153].

The virus-based mammalian continuous directed evolution system, VEGAS (viral evolution of genetically actuating sequences), uses mammalian RNA alphavirus Sindbis for orthogonal diversification and selection [154]. RNA alphavirus Sindbis is a plus single-stranded RNA (+ ssRNA) virus, the RNA genome of which directly serves as the

template for translation of viral proteins [157]. Virus particles can be introduced into cells by transfecting DNA plasmids encoding the RNA genome of the virus. The genome of the virus is replicated by an intrinsic RNA-dependent RNA polymerase (RdRP). The RdRP is deficient in proofreading, leading to high error rate during RNA replication. To enable selection, the non-structural and structural parts of the genome of the virus was separated into different plasmid vectors. The non-structural part and the evolving genes were delivered together on the same DNA vector, while the structural part is delivered on a different DNA vector, and the production of the viral structural proteins were controlled by the activity of interest. The VEGAS system is unique in that it enables evolution of full-length mammalian genes that are not accessible to other systems. For example, the authors successfully evolved GPCR (G protein-coupled receptor) to be constitutively active.

The automation of in vivo-in vitro conversion of libraries is exemplified by the MAGE (multiplex automated genome engineering) system developed by George M. Church and colleagues [158]. *E. coli* genomes are engineered by ssDNA recombination mediated by bacteriophage λ-Red ssDNA-binding protein β. First, cells are grown to mid-log phase, and β protein production is induced. Second, media is replaced with water, and ssDNA oligomers are electroporated into *E. coli* cells. Finally, the cells are recovered and fed into the first growth step. As the name “MAGE” imply, the entire genome engineering processes is automated.

1.3 References

1. Ahmad, M., and Cashmore, A.R. (1993). HY4 gene of A. thaliana encodes a protein with characteristics of a blue-light photoreceptor. Nature 366, 162–166.
2. Lin, C., Robertson, D.E., Ahmad, M., Raibekas, A.A., Jorns, M.S., Dutton, P.L., and Cashmore, A.R. (1995). Association of flavin adenine dinucleotide with the Arabidopsis blue light receptor CRY1. Science 269, 968–970.
3. Cashmore, A.R. (2003). Cryptochromes: Enabling plants and animals to determine circadian time. Cell 114, 537–543.
4. Sancar, A. (2003). Structure and Function of DNA Photolyase and Cryptochrome Blue-Light Photoreceptors. Chem. Rev. 103, 2203–2238.
5. Lin, C., and Shalitin, D. (2003). Cryptochrome structure and signal transduction. Annu. Rev. Plant Biol. 54, 469–496.
6. Yamada, D., Dokainish, H.M., Iwata, T., Yamamoto, J., Ishikawa, T., Todo, T., Iwai, S., Getzoff, E.D., Kitao, A., and Kandori, H. (2016). Functional Conversion of CPD and (6–4) Photolyases by Mutation. Biochemistry 55, 4173–4183.
7. Mei, Q., and Dvornyk, V. (2015). Evolutionary history of the photolyase/cryptochrome superfamily in eukaryotes. PLoS One 10, e0135940.
8. Yu, X., Liu, H., Klejnot, J., and Lin, C. (2010). The Cryptochrome Blue-Light Receptors. Arab. B. 8, e0135.
9. Lin, C., and Todo, T. (2005). The cryptochromes. Genome Biol. 6, 220.
10. Kanai, S., Kikuno, R., Toh, H., Ryo, H., and Todo, T. (1997). Molecular evolution of the photolyase-blue-light photoreceptor family. J. Mol. Evol. 45, 535–548.
11. Lucas-Lledó, J.I., and Lynch, M. (2009). Evolution of mutation rates:

Phylogenomic analysis of the photolyase/cryptochrome family. Mol. Biol. Evol. 26, 1143–1153.

12. Tagua, V.G., Pausch, M., Eckel, M., Gutiérrez, G., Miralles-Durán, A., Sanz, C., Eslava, A.P., Pokorny, R., Corrochano, L.M., and Batschauer, A. (2015). Fungal cryptochrome with DNA repair activity reveals an early stage in cryptochrome evolution. Proc. Natl. Acad. Sci. U. S. A. 112, 15130–15135.
13. Huang, Y., Baxter, R., Smith, B.S., Partch, C.L., Colbert, C.L., and Deisenhofer, J. (2006). Crystal structure of cryptochrome 3 from Arabidopsis thaliana and its implications for photolyase activity. Proc. Natl. Acad. Sci. U. S. A. 103, 17701–17706.
14. Klar, T., Pokorny, R., Moldt, J., Batschauer, A., and Essen, L.O. (2007). Cryptochrome 3 from Arabidopsis thaliana: Structural and Functional Analysis of its Complex with a Folate Light Antenna. J. Mol. Biol. 366, 954–964.
15. Pokorny, R., Klar, T., Hennecke, U., Carell, T., Batschauer, A., and Essen, L.O. (2008). Recognition and repair of UV lesions in loop structures of duplex DNA by DASH-type cryptochrome. Proc. Natl. Acad. Sci. U. S. A. 105, 21023–21027.
16. Hitomi, K. (2000). Bacterial cryptochrome and photolyase: characterization of two photolyase-like genes of Synechocystis sp. PCC6803. Nucleic Acids Res. 28, 2353–2362.
17. Brudler, R., Hitomi, K., Daiyasu, H., Toh, H., Kucho, K.I., Ishiura, M., Kanehisa, M., Roberts, V.A., Todo, T., Tainer, J.A., and Getzoff, E.D. (2003). Identification of a new cryptochrome class: Structure, function, and evolution. Mol. Cell 11, 59–67.
18. Wang, X., Wang, Q., Nguyen, P., and Lin, C. (2014). Chapter Seven -

Cryptochrome-Mediated Light Responses in Plants. In Signaling Pathways in Plants, Y. Machida, C. Lin, and F. B. T.-T. E. Tamanoi, eds. (Academic Press), pp. 167–189.

19. Liu, B., Yang, Z., Gomez, A., Liu, B., Lin, C., and Oka, Y. (2016). Signaling mechanisms of plant cryptochromes in Arabidopsis thaliana. J. Plant Res. 129, 137–148.

20. Collins, B., Mazzoni, E.O., Stanewsky, R., and Blau, J. (2006). Drosophila CRYPTOCHROME is a circadian transcriptional repressor. Curr. Biol. 16, 441–449.

21. Michael, A.K., Fribourgh, J.L., Van Gelder, R.N., and Partch, C.L. (2017). Animal Cryptochromes: Divergent Roles in Light Perception, Circadian Timekeeping and Beyond. Photochem. Photobiol. 93, 128–140.

22. Rosensweig, C., Reynolds, K.A., Gao, P., Laothamatas, I., Shan, Y., Ranganathan, R., Takahashi, J.S., and Green, C.B. (2018). An evolutionary hotspot defines functional differences between CRYPTOCHROMES. Nat. Commun. 9, 1138.

23. Yuan, Q., Metterville, D., Briscoe, A.D., and Reppert, S.M. (2007). Insect cryptochromes: Gene duplication and loss define diverse ways to construct insect circadian clocks. Mol. Biol. Evol. 24, 948–955.

24. Scheerer, P., Zhang, F., Kalms, J., Von Stetten, D., Krauß, N., Oberpichler, I., and Lamparter, T. (2015). The class III cyclobutane pyrimidine dimer photolyase structure reveals a new antenna chromophore binding site and alternative photoreduction pathways. J. Biol. Chem. 290, 11504–11514.

25. Barrero, J.M., Downie, A.B., Xu, Q., and Gubler, F. (2014). A role for barley CRYPTOCHROME1 in light regulation of grain dormancy and germination. Plant Cell 26, 1094–1104.

26. Mockler, T.C., Guo, H., Yang, H., Duong, H., and Lin, C. (1999). Antagonistic actions of Arabidopsis cryptochromes and phytochrome B in the regulation of floral induction. Development 126, 2073–2082.

27. Usami, T., Mochizuki, N., Kondo, M., Nishimura, M., and Nagatani, A. (2004). Cryptochromes and phytochromes synergistically regulate Arabidopsis root greening under blue light. Plant Cell Physiol. 45, 1798–1808.

28. Meng, Y., Li, H., Wang, Q., Liu, B., and Lin, C. (2013). Blue light-dependent interaction between cryptochrome2 and CIB1 regulates transcription and leaf senescence in soybean. Plant Cell 25, 4405–4420.

29. Guo, H., Yang, H., Mockler, T.C., and Lin, C. (1998). Regulation of flowering time by Arabidopsis photoreceptors. Science 279, 1360–1363.

30. El-Din El-Assal, S., Alonso-Blanco, C., Peeters, A.J.M., Wagemaker, C., Weller, J.L., and Koornneef, M. (2003). The Role of Cryptochrome 2 in Flowering in Arabidopsis. Plant Physiol. 133, 1504–1516.

31. El-Assal, S.E.D., Alonso-Blanco, C., Hanhart, C.J., and Koornneef, M. (2004). Pleiotropic effects of the Arabidopsis cryptochrome 2 allelic variation underlie fruit trait-related QTL. Plant Biol. 6, 370–374.

32. Facella, P., Lopez, L., Carbone, F., Galbraith, D.W., Giuliano, G., and Perrotta, G. (2008). Diurnal and circadian rhythms in the tomato transcriptome and their modulation by cryptochrome photoreceptors. PLoS One 3, e2798.

33. Yang, Z., Liu, B., Su, J., Liao, J., Lin, C., and Oka, Y. (2017). Cryptochromes Orchestrate Transcription Regulation of Diverse Blue Light Responses in Plants. Photochem. Photobiol. 93, 112–127.

34. Yang, Y.J., Zuo, Z.C., Zhao, X.Y., Li, X., Klejnot, J., Li, Y., Chen, P., Liang, S.P., Yu, X.H., Liu, X.M., and Lin, C.T. (2008). Blue-light-independent activity of Arabidopsis cryptochromes in the regulation of steady-state levels of protein and mRNA expression. Mol. Plant 1, 167–177.

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

36. Devlin, P.F., and Kay, S.A. (2000). Cryptochromes are required for phytochrome signaling to the circadian clock but not for rhythmicity. Plant Cell 12, 2499–2509.

37. Blázquez, M.A., Ahn, J.H., and Weigel, D. (2003). A thermosensory pathway controlling flowering time in Arabidopsis thaliana. Nat. Genet. 33, 168–171.

38. Ohgishi, M., Saji, K., Okada, K., and Sakai, T. (2004). Functional analysis of each blule light receptor, cry1, cry2, phot1, and phot2, by using combinatorial multiple mutants in arabidopsis. Proc. Natl. Acad. Sci. U. S. A. 101, 2223–2228.

39. Mao, J., Zhang, Y.C., Sang, Y., Li, Q.H., and Yang, H.Q. (2005). A role for Arabidopsis cryptochromes and COP1 in the regulation of stomatal opening. Proc. Natl. Acad. Sci. U. S. A. 102, 12270–12275.

40. Kang, C.Y., Lian, H.L., Wang, F.F., Huang, J.R., and Yang, H.Q. (2009). Cryptochromes, phytochromes, and COP1 regulate light-controlled stomatal development in arabidopsis. Plant Cell 21, 2624–2641.

41. Pedmale, U. V., Huang, S.S.C., Zander, M., Cole, B.J., Hetzel, J., Ljung, K., Reis, P.A.B., Sridevi, P., Nito, K., Nery, J.R., Ecker, J.R., and Chory, J. (2016). Cryptochromes interact directly with PIFs to control plant growth in limiting blue light. Cell 164, 233–245.

42. Danon, A., Coll, N.S., and Apel, K. (2006). Cryptochrome-1-dependent execution of programmed cell death induced by singlet oxygen in Arabidopsis thaliana. Proc. Natl. Acad. Sci. U. S. A. 103, 17036–17041.

43. Ahmad, M., Jarillo, J.A., Smirnova, O., and Cashmore, A.R. (1998). Cryptochrome blue-light photoreceptors of Arabidopsis implicated in phototropism. Nature 392, 720–723.

44. Lascève, G., Leymarie, J., Olney, M.A., Liscum, E., Christie, J.M., Vavasseur, A., and Briggs, W.R. (1999). Arabidopsis contains at least four independent blue-light-activated signal transduction pathways. Plant Physiol. 120, 605–614.

45. Tsuchida-Mayama, T., Sakai, T., Hanada, A., Uehara, Y., Asami, T., and Yamaguchi, S. (2010). Role of the phytochrome and cryptochrome signaling pathways in hypocotyl phototropism. Plant J. 62, 653–662.

46. Ahmad, M., Galland, P., Ritz, T., Wiltschko, R., and Wiltschko, W. (2007). Magnetic intensity affects cryptochrome-dependent responses in Arabidopsis thaliana. Planta 225, 615–624.

47. Harris, S.R., Henbest, K.B., Maeda, K., Pannell, J.R., Timmel, C.R., Hore, P.J., and Okamoto, H. (2009). Effect of magnetic fields on cryptochrome-dependent responses in Arabidopsis thaliana. J. R. Soc. Interface 6, 1193–1205.

48. Pooam, M., Arthaut, L.D., Burdick, D., Link, J., Martino, C.F., and Ahmad, M.

(2019). Magnetic sensitivity mediated by the Arabidopsis blue-light receptor cryptochrome occurs during flavin reoxidation in the dark. Planta 249, 319–332.

49. Lin, C., Ahmad, M., Gordon, D., and Cashmore, a R. (1995). Expression of an Arabidopsis cryptochrome gene in transgenic tobacco results in hypersensitivity to blue, UV-A, and green light. Proc. Natl. Acad. Sci. U. S. A. 92, 8423–8427.

50. El-Din El-Assal, S., Alonso-Blanco, C., Peeters, A.J.M., Raz, V., and Koornneef, M. (2001). A QTL for flowering time in Arabidopsis reveals a novel allele of CRY2. Nat. Genet. 29, 435–440.

51. Kleine, T., Lockhart, P., and Batschauer, A. (2003). An Arabidopsis protein closely related to Synechocystis cryptochrome is targeted to organelles. Plant J. 35, 93–103.

52. Selby, C.P., and Sancar, A. (2006). A cryptochrome/photolyase class of enzymes with single-stranded DNA-specific photolyase activity. Proc. Natl. Acad. Sci. U. S. A. 103, 17696–17700.

53. Yanovsky, M.J., Mazzella, M.A., Whitelam, G.C., and Casal, J.J. (2001). Resetting of the circadian clock by phytochromes and cryptochromes in Arabidopsis. J. Biol. Rhythms 16, 523–530.

54. Zhong, S., Shi, H., Xue, C., Wei, N., Guo, H., and Deng, X.W. (2014). Ethylene-orchestrated circuitry coordinates a seedling's response to soil cover and etiolated growth. Proc. Natl. Acad. Sci. U. S. A. 111, 3913–3920.

55. Yu, Y., and Huang, R. (2017). Integration of ethylene and light signaling affects hypocotyl growth in Arabidopsis. Front. Plant Sci. 8, 57.

56. Guo, H., Duong, H., Ma, N., and Lin, C. (1999). The Arabidopsis blue-light

receptor cryptochrome 2 is a nuclear protein regulated by a blue-light dependent post-transcriptional mechanism. Plant J. 19, 279–289.

57. Yu, X., Klejnot, J., Zhao, X., Shalitin, D., Maymon, M., Yang, H., Lee, J., Liu, X., Lopez, J., and Lin, C. (2007). Arabidopsis cryptochrome 2 completes its posttranslational life cycle in the nucleus. Plant Cell 19, 3146–56.

58. Yu, X., Shalitin, D., Liu, X., Maymon, M., Klejnot, J., Yang, H., Lopez, J., Zhao, X., Bendehakkalu, K.T., and Lin, C. (2007). Derepression of the NC80 motif is critical for the photoactivation of Arabidopsis CRY2. Proc. Natl. Acad. Sci. U. S. A. 104, 7289–94.

59. Liu, H., Yu, X., Li, K., Klejnot, J., Yang, H., Lisiero, D., and Lin, C. (2008). Photoexcited CRY2 interacts with CIB1 to regulate transcription and floral initiation in *Arabidopsis*. Science 322, 1535–1539.

60. Lin, C., Yang, H., Guo, H., Mockler, T., Chen, J., and Cashmore, A.R. (1998). Enhancement of blue-light sensitivity of Arabidopsis seedlings by a blue light receptor cryptochrome 2. Proc. Natl. Acad. Sci. U. S. A. 95, 2686–90.

61. Li, X., Wang, Q., Yu, X., Liu, H., Yang, H., Zhao, C., Liu, X., Tan, C., Klejnot, J., Zhong, D., and Lin, C. (2011). Arabidopsis cryptochrome 2 (CRY2) functions by the photoactivation mechanism distinct from the tryptophan (trp) triad-dependent photoreduction. Proc. Natl. Acad. Sci. U. S. A. 108, 20844–9.

62. Gao, J., Wang, X., Zhang, M., Bian, M., Deng, W., Zuo, Z., Yang, Z., Zhong, D., and Lin, C. (2015). Trp triad-dependent rapid photoreduction is not required for the function of Arabidopsis CRY1. Proc. Natl. Acad. Sci. U. S. A. 112, 9135–9140.

63. Wang, H., Ma, L.G., Li, J.M., Zhao, H.Y., and Xing Wang Deng (2001). Direct

interaction of Arabidopsis cryptochromes with COP1 in light control development. Science 294, 154–158.

64. Ang, L.H., Chattopadhyay, S., Wei, N., Oyama, T., Okada, K., Batschauer, A., and Deng, X.W. (1998). Molecular interaction between COP1 and HY5 defines a regulatory switch for light control of Arabidopsis development. Mol. Cell 1, 213–222.

65. Osterlund, M.T., Hardtke, C.S., Ning, W., and Deng, X.W. (2000). Targeted destabilization of HY5 during light-regulated development of Arabidopsis. Nature 405, 462–466.

66. Oyama, T., Shimura, Y., and Okada, K. (1997). The Arabidopsis HY5 gene encodes a bZIP protein that regulates stimulus- induced development of root and hypocotyl. Genes Dev. 11, 2983–2995.

67. Jing, Y., Zhang, D., Wang, X., Tang, W., Wang, W., Huai, J., Xu, G., Chen, D., Li, Y., and Lin, R. (2013). Arabidopsis chromatin remodeling factor PICKLE interacts with transcription factor HY5 to regulate hypocotyl cell elongation. Plant Cell 25, 242–256.

68. Gangappa, S.N., and Botto, J.F. (2016). The Multifaceted Roles of HY5 in Plant Growth and Development. Mol. Plant 9, 1353–1365.

69. Kami, C., Lorrain, S., Hornitschek, P., and Fankhauser, C. (2010). Light-regulated plant growth and development. In Current Topics in Developmental Biology, pp. 29–66.

70. Neff, M.M., and Chory, J. (1998). Genetic interactions between phytochrome A, phytochrome B, and cryptochrome 1 during Arabidopsis development. Plant

Physiol. 118, 27–35.

71. Song, Y.H., Shim, J.S., Kinmonth-Schultz, H.A., and Imaizumi, T. (2015). Photoperiodic Flowering: Time Measurement Mechanisms in Leaves. Annu. Rev. Plant Biol. 66, 441–464.

72. Turnbull, C. (2011). Long-distance regulation of flowering time. J. Exp. Bot. 62, 4399–4413.

73. Kobayashi, Y., Kaya, H., Goto, K., Iwabuchi, M., and Araki, T. (1999). A pair of related genes with antagonistic roles in mediating flowering signals. Science 286, 1960–1962.

74. Suárez-López, P., Wheatley, K., Robson, F., Onouchi, H., Valverde, F., and Coupland, G. (2001). CONSTANS mediates between the circadian clock and the control of flowering in Arabidopsis. Nature 410, 1116–1120.

75. Shim, J.S., Kubota, A., and Imaizumi, T. (2017). Circadian clock and photoperiodic flowering in arabidopsis: CONSTANS is a Hub for Signal integration. Plant Physiol. 173, 5–15.

76. Kamioka, M., Takao, S., Suzuki, T., Taki, K., Higashiyam, T., Kinoshita, T., and Nakamichi, N. (2016). Direct repression of evening genes by CIRCADIAN CLOCK-ASSOCIATED1 in the Arabidopsis circadian clock. Plant Cell 28, 696–711.

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

78. Nakamichi, N., Kita, M., Niinuma, K., Ito, S., Yamashino, T., Mizoguchi, T., and Mizuno, T. (2007). Arabidopsis clock-associated pseudo-response regulators PRR9, PRR7 and PRR5 coordinately and positively regulate flowering time through the canonical CONSTANS-dependent photoperiodic pathway. Plant Cell Physiol. 48, 822–832.

79. Imaizumi, T., Schultz, T.F., Harmon, F.G., Ho, L.A., and Kay, S.A. (2005). FKF1 F-box protein mediates cyclic degradation of a repressor of CONSTANS in Arabidopsis. Science 309, 293–297.

80. Mizoguchi, T., Wright, L., Fujiwara, S., Cremer, F., Lee, K., Onouchi, H., Mouradov, A., Fowler, S., Kamada, H., Putterill, J., and Coupland, G. (2005). Distinct roles of GIGANTEA in promoting flowering and regulating circadian rhythms in Arabidopsis. Plant Cell 17, 2255–2270.

81. Sawa, M., Nusinow, D.A., Kay, S.A., and Imaizumi, T. (2007). FKF1 and GIGANTEA complex formation is required for day-length measurement in Arabidopsis. Science 318, 261–265.

82. Jang, S., Marchal, V., Panigrahi, K.C.S., Wenkel, S., Soppe, W., Deng, X.W., Valverde, F., and Coupland, G. (2008). Arabidopsis COP1 shapes the temporal pattern of CO accumulation conferring a photoperiodic flowering response. EMBO J. 27, 1277–1288.

83. Laubinger, S., Marchal, V., Gentillhomme, J., Wenkel, S., Adrian, J., Jang, S., Kulajta, C., Braun, H., Coupland, G., and Hoecker, U. (2006). Arabidopsis SPA proteins regulate photoperiodic flowering and interact with the floral inducer CONSTANS to regulate its stability. Development 133, 3213–3222.

84. Liu, L.J., Zhang, Y.C., Li, Q.H., Sang, Y., Mao, J., Lian, H.L., Wang, L., and Yang, H.Q. (2008). COP1-mediated ubiquitination of CONSTANS is implicated in cryptochrome regulation of flowering in Arabidopsis. Plant Cell 20, 292–306.

85. Valverde, F., Mouradov, A., Soppe, W., Ravenscroft, D., Samach, A., and Coupland, G. (2004). Photoreceptor regulation of CONSTANS protein in photoperiodic flowering. Science 303, 1003–6.

86. Lazaro, A., Valverde, F., Piñeiro, M., and Jarillo, J.A. (2012). The Arabidopsis E3 ubiquitin ligase HOS1 negatively regulates CONSTANS abundance in the photoperiodic control of flowering. Plant Cell 24, 982–999.

87. Zuo, Z., Liu, H., Liu, B., Liu, X., and Lin, C. (2011). Blue Light-Dependent Interaction of CRY2 with SPA1 Regulates COP1 activity and Floral Initiation in Arabidopsis. Curr. Biol. 21, 841–847.

88. Imaizumi, T., Tran, H.G., Swartz, T.E., Briggs, W.R., and Kay, S.A. (2003). FKF1 is essential for photoperiodic-specific light signalling in Arabidopsis. Nature 426, 302–306.

89. Song, Y.H., Smith, R.W., To, B.J., Millar, A.J., and Imaizumi, T. (2012). FKF1 conveys timing information for CONSTANS stabilization in photoperiodic flowering. Science 336, 1045–1049.

90. Song, Y.H., Estrada, D.A., Johnson, R.S., Kim, S.K., Lee, S.Y., MacCoss, M.J., and Imaizumi, T. (2014). Distinct roles of FKF1, GIGANTEA, and ZEITLUPE proteins in the regulation of constans stability in Arabidopsis photoperiodic flowering. Proc. Natl. Acad. Sci. U. S. A. 11, 17672–17677.

91. Mockler, T., Yang, H., Yu, X.H., Parikh, D., Cheng, Y. chia, Dolan, S., and Lin, C.

(2003). Regulation of photoperiodic flowering by Arabidopsis photoreceptors. Proc. Natl. Acad. Sci. U. S. A. 100, 2140–2145.

92. Lu, X.-D., Zhou, C.-M., Xu, P.-B., Luo, Q., Lian, H.-L., and Yang, H.-Q. (2015). Red-Light-Dependent Interaction of phyB with SPA1 Promotes COP1–SPA1 Dissociation and Photomorphogenic Development in Arabidopsis. Mol. Plant 8, 467–478.

93. Lazaro, A., Mouriz, A., Piñeiro, M., and Jarillo, J.A. (2015). Red Light-Mediated Degradation of CONSTANS by the E3 Ubiquitin Ligase HOS1 Regulates Photoperiodic Flowering in Arabidopsis. Plant Cell 27, 2437–54.

94. Liu, Y., Li, X., Ma, D., Chen, Z., Wang, J., and Liu, H. (2018). CIB 1 and CO interact to mediate CRY 2-dependent regulation of flowering. EMBO Rep. 19.

95. Shalitin D; Yang H; Mockler TC; Maymon M; Guo H; Whitelam GC; Lin C (2002). Regulation of Arabidopsis cryptochrome 2 by blue-light- dependent phosphorylation. Nature 417, 763–767.

96. Yu, X., Sayegh, R., Maymon, M., Warpeha, K., Klejnot, J., Yang, H., Huang, J., Lee, J., Kaufman, L., and Lin, C. (2009). Formation of nuclear bodies of Arabidopsis CRY2 in response to blue light is associated with its blue light-dependent degradation. Plant Cell 21, 118–130.

97. Zuo, Z.C., Meng, Y.Y., Yu, X.H., Zhang, Z.L., Feng, D.S., Sun, S.F., Liu, B., and Lin, C.T. (2012). A study of the blue-light-dependent phosphorylation, degradation, and photobody formation of Arabidopsis CRY2. Mol. Plant 5, 726–733.

98. Weber, S. (2005). Light-driven enzymatic catalysis of DNA repair: A review of

recent biophysical studies on photolyase. Biochim. Biophys. Acta - Bioenerg. 1707, 1–23.

99. Engelhard, C., Wang, X., Robles, D., Moldt, J., Essen, L.-O., Batschauer, A., Bittl, R., and Ahmad, M. (2014). Cellular metabolites enhance the light sensitivity of Arabidopsis cryptochrome through alternate electron transfer pathways. Plant Cell 26, 4519–31.

100. Hoang, N., Schleicher, E., Kacprzak, S., Bouly, J.-P., Picot, M., Wu, W., Berndt, A., Wolf, E., Bittl, R., and Ahmad, M. (2008). Human and Drosophila Cryptochromes Are Light Activated by Flavin Photoreduction in Living Cells. PLoS Biol. 6, e160.

101. Liu, B., Liu, H., Zhong, D., and Lin, C. (2010). Searching for a photocycle of the cryptochrome photoreceptors. Curr. Opin. Plant Biol. 13, 578–586.

102. RUPERT, C.S. (1962). Photoenzymatic repair of ultraviolet damage in DNA. I. Kinetics of the reaction. J. Gen. Physiol. 45, 703–724.

103. RUPERT, C.S. (1962). Photoenzymatic repair of ultraviolet damage in DNA. II. Formation of an enzyme-substrate complex. J. Gen. Physiol. 45, 725–741.

104. Kim, S.T., Heelis, P.F., Okamura, T., Hirata, Y., Mataga, N., and Sancar, A. (1991). Determination of rates and yields of interchromophore (folate .fwdarw. flavin) energy transfer and intermolecular (flavin .fwdarw. DNA) electron transfer in Escherichia coli photolyase by time-resolved fluorescence and absorption spectroscopy. Biochemistry 30, 11262–11270.

105. Payne, G., Heelis, P., Rohrs, B., and Sancar, A. (1987). The Active Form of Escherichia coli DNA Photolyase Contains a Fully Reduced Flavin and Not a

Flavin Radical, both in Vivo and in Vitro. Biochemistry 26, 7121–7127.

106. Zhang, M., Wang, L., and Zhong, D. (2017). Photolyase: Dynamics and electron-transfer mechanisms of DNA repair. Arch. Biochem. Biophys. 632, 158–174.

107. Liu, Z., Zhang, M., Guo, X., Tan, C., Li, J., Wang, L., Sancar, A., and Zhong, D. (2013). Dynamic determination of the functional state in photolyase and the implication for cryptochrome. Proc. Natl. Acad. Sci. U. S. A. 110, 12972–12977.

108. Kao, Y.T., Saxena, C., Wang, L., Sancar, A., and Zhong, D. (2005). Direct observation of thymine dimer repair in DNA by photolyase. Proc. Natl. Acad. Sci. U. S. A. 102, 16128–16132.

109. Zhang, M., Wang, L., Shu, S., Sancar, A., and Zhong, D. (2016). Bifurcating electron-transfer pathways in DNA photolyases determine the repair quantum yield. Science 354, 209–213.

110. Berndt, A., Kottke, T., Breitkreuz, H., Dvorsky, R., Hennig, S., Alexander, M., and Wolf, E. (2007). A novel photoreaction mechanism for the circadian blue light photoreceptor Drosophila cryptochrome. J. Biol. Chem. 282, 13011–13021.

111. Öztürk, N., Song, S.H., Selby, C.P., and Sancar, A. (2008). Animal Type 1 cryptochromes: Analysis of the redox state of the flavin cofactor by site-directed mutagenesis. J. Biol. Chem. 283, 3256–3263.

112. Paulus, B., Bajzath, C., Melin, F., Heidinger, L., Kromm, V., Herkersdorf, C., Benz, U., Mann, L., Stehle, P., Hellwig, P., Weber, S., and Schleicher, E. (2015). Spectroscopic characterization of radicals and radical pairs in fruit fly cryptochrome - protonated and nonprotonated flavin radical-states. FEBS J. 282, 3175–3189.

113. Zeugner, A., Byrdin, M., Bouly, J.-P., Bakrim, N., Giovani, B., Brettel, K., and Ahmad, M. (2005). Light-induced electron transfer in Arabidopsis cryptochrome-1 correlates with in vivo function. J. Biol. Chem. 280, 19437–40.

114. Lin, C., Top, D., Manahan, C.C., Young, M.W., and Crane, B.R. (2018). Circadian clock activity of cryptochrome relies on tryptophan-mediated photoreduction. Proc. Natl. Acad. Sci. U. S. A. 115, 3822–3827.

115. Song, S.H., Öztürk, N., Denaro, T.R., Arat, N.Ö., Kao, Y.T., Zhu, H., Zhong, D., Reppert, S.M., and Sancar, A. (2007). Formation and function of flavin anion radical in cryptochrome 1 blue-light photoreceptor of monarch butterfly. J. Biol. Chem. 282, 17608–17612.

116. Ozturk, N., Selby, C.P., Annayev, Y., Zhong, D., and Sancar, A. (2011). Reaction mechanism of Drosophila cryptochrome. Proc. Natl. Acad. Sci. U. S. A. 108, 516–521.

117. Ozturk, N., Selby, C.P., Zhong, D., and Sancar, A. (2014). Mechanism of photosignaling by drosophila cryptochrome ROLE of the REDOX status of the flavin chromophore. J. Biol. Chem. 289, 4634–4642.

118. Shalitin, D., Yu, X., Maymon, M., Mockler, T., and Lin, C. (2003). Blue light-dependent in vivo and in vitro phosphorylation of Arabidopsis cryptochrome 1. Plant Cell 15, 2421–9.

119. Wang, Q., Zuo, Z., Wang, X., Gu, L., Yoshizumi, T., Yang, Z., Yang, L., Liu, Q., Liu, W., Han, Y.-J., Kim, J.-I., Liu, B., Wohlschlegel, J.A., Matsui, M., Oka, Y., and Lin, C. (2016). Photoactivation and inactivation of Arabidopsis cryptochrome 2. Science 354, 343–347.

120. Liu, Q., Wang, Q., Deng, W., Wang, X., Piao, M., Cai, D., Li, Y., Barshop, W.D., Yu, X., Zhou, T., Liu, B., Oka, Y., Wohlschlegel, J., Zuo, Z., and Lin, C. (2017). Molecular basis for blue light-dependent phosphorylation of Arabidopsis cryptochrome 2. Nat. Commun. 8, 15234.

121. Wang, S., Li, L., Xu, P., Lian, H., Wang, W., Xu, F., Mao, Z., Zhang, T., and Yang, H. (2018). CRY1 interacts directly with HBI1 to regulate its transcriptional activity and photomorphogenesis in Arabidopsis. J. Exp. Bot. 69, 3867–3881.

122. Xu, F., He, S., Zhang, J., Mao, Z., Wang, W., Li, T., Hua, J., Du, S., Xu, P., Li, L., Lian, H., and Yang, H.-Q. (2018). Photoactivated CRY1 and phyB Interact Directly with AUX/IAA Proteins to Inhibit Auxin Signaling in Arabidopsis. Mol. Plant 11, 523–541.

123. Yang, H.-Q., Wu, Y.-J., Tang, R.-H., Liu, D., Liu, Y., and Cashmore, A.R. (2000). The C Termini of Arabidopsis Cryptochromes Mediate a Constitutive Light Response. Cell 103, 815–827.

124. Matsuura, T., Hosoda, K., Ichihashi, N., Kazuta, Y., and Yomo, T. (2011). Kinetic analysis of β-galactosidase and β-glucuronidase tetramerization coupled with protein translation. J. Biol. Chem. 286, 22028–22034.

125. Phillips, G.N. (1997). Structure and dynamics of green fluorescent protein. Curr. Opin. Struct. Biol. 7, 821–827.

126. Rosenfeldt, G., Viana, R.M., Mootz, H.D., von Arnim, A.G., and Batschauer, A. (2008). Chemically Induced and Light-Independent Cryptochrome Photoreceptor Activation. Mol. Plant 1, 4–14.

127. Kennedy, M.J., Hughes, R.M., Peteya, L. a, Schwartz, J.W., Ehlers, M.D., and

Tucker, C.L. (2010). Rapid blue-light-mediated induction of protein interactions in living cells. Nat. Methods 7, 973–975.

128. Conrad, T.M., Lewis, N.E., and Palsson, B.Ø. (2011). Microbial laboratory evolution in the era of genome-scale science. Mol. Syst. Biol. 7, 509.

129. Portnoy, V.A., Bezdan, D., and Zengler, K. (2011). Adaptive laboratory evolution-harnessing the power of biology for metabolic engineering. Curr. Opin. Biotechnol. 22, 590–594.

130. Cobb, R.E., Si, T., and Zhao, H. (2012). Directed evolution: An evolving and enabling synthetic biology tool. Curr. Opin. Chem. Biol. 16, 285–291.

131. Packer, M.S., and Liu, D.R. (2015). Methods for the directed evolution of proteins. Nat. Rev. Genet. 16, 379–394.

132. Kodym, A., and Afza, R. (2003). Physical and chemical mutagenesis. Methods Mol. Biol. 236, 189–204.

133. McCullum, E.O., Williams, B.A.R., Zhang, J., and Chaput, J.C. (2010). Random mutagenesis by error-prone PCR. Methods Mol. Biol. 634, 103–109.

134. Hughes, R.A., and Ellington, A.D. (2017). Synthetic DNA synthesis and assembly: Putting the synthetic in synthetic biology. Cold Spring Harb. Perspect. Biol. 9, a023812.

135. Crameri, A., Raillard, S.A., Bermudez, E., and Stemmer, W.P.C. (1998). DNA shuffling of a family of genes from diverse species accelerates directed evolution. Nature 391, 288–291.

136. Selifonova, O., Valle, F., and Schellenberger, V. (2001). Rapid Evolution of Novel Traits in Microorganisms. Appl. Environ. Microbiol. 67, 3645–3649.

137. Singer, B., and Kusmierek, J.T. (1982). Chemical Mutagenesis. Annu. Rev. Biochem. 52, 655–691.

138. Ramos, J.L., Stolz, A., Reineke, W., and Timmis, K.N. (1986). Altered effector specificities in regulators of gene expression: TOL plasmid xylS mutants and their use to engineer expansion of the range of aromatics degraded by bacteria. Proc. Natl. Acad. Sci. U. S. A. 83, 8467–8471.

139. Ravikumar, A., Arzumanyan, G.A., Obadi, M.K.A., Javanpour, A.A., and Liu, C.C. (2018). Scalable, Continuous Evolution of Genes at Mutation Rates above Genomic Error Thresholds. Cell 175, 1946–1957.

140. Nannemann, D.P., Birmingham, W.R., Scism, R.A., and Bachmann, B.O. (2011). Assessing directed evolution methods for the generation of biosynthetic enzymes with potential in drug biosynthesis. Future Med. Chem. 3, 803–819.

141. Greener, A., Callahan, M., and Jerpseth, B. (1997). An Efficient Random Mutagenesis Technique Using an E. coli Mutator Strain. Appl. Biochem. Biotechnol. - Part B Mol. Biotechnol. 7, 189–195.

142. Esvelt, K.M., Carlson, J.C., and Liu, D.R. (2011). A system for the continuous directed evolution of biomolecules. Nature 472, 499–503.

143. Ravikumar, A., Arrieta, A., and Liu, C.C. (2014). An orthogonal DNA replication system in yeast. Nat. Chem. Biol. 10, 175–177.

144. Badran, A.H., Guzov, V.M., Huai, Q., Kemp, M.M., Vishwanath, P., Kain, W., Nance, A.M., Evdokimov, A., Moshiri, F., Turner, K.H., Wang, P., Malvar, T., and Liu, D.R. (2016). Continuous evolution of Bacillus thuringiensis toxins overcomes insect resistance. Nature 533, 58–63.

145. Sawano, A. (2000). Directed evolution of green fluorescent protein by a new versatile PCR strategy for site-directed and semi-random mutagenesis. Nucleic Acids Res. 28, 78e – 78.

146. Zhang, J.H., Dawe, G., and Stemmer, W.P.C. (1997). Directed evolution of a fucosidase from a galactosidase by DNA shuffling and screening. Proc. Natl. Acad. Sci. U. S. A. 94, 4504–4509.

147. Lin, H., Tao, H., and Cornish, V.W. (2004). Directed evolution of a glycosynthase via chemical complementation. J. Am. Chem. Soc. 126, 15051–15059.

148. d'Oelsnitz, S., and Ellington, A. (2018). Continuous directed evolution for strain and protein engineering. Curr. Opin. Biotechnol. 53, 158–163.

149. Wright, M.C., and Joyce, G.F. (1997). Continuous in vitro evolution of catalytic function. Science 276, 614–617.

150. Halperin, S.O., Tou, C.J., Wong, E.B., Modavi, C., Schaffer, D. V., and Dueber, J.E. (2018). CRISPR-guided DNA polymerases enable diversification of all nucleotides in a tunable window. Nature 560, 248–252.

151. Crook, N., Abatemarco, J., Sun, J., Wagner, J.M., Schmitz, A., and Alper, H.S. (2016). In vivo continuous evolution of genes and pathways in yeast. Nat. Commun. 7, 13051.

152. Hess, G.T., Frésard, L., Han, K., Lee, C.H., Li, A., Cimprich, K.A., Montgomery, S.B., and Bassik, M.C. (2016). Directed evolution using dCas9-targeted somatic hypermutation in mammalian cells. Nat. Methods 13, 1036–1042.

153. Ma, Y., Zhang, J., Yin, W., Zhang, Z., Song, Y., and Chang, X. (2016). Targeted AID-mediated mutagenesis (TAM) enables efficient genomic diversification in

mammalian cells. Nat. Methods 13, 1029–1035.

154. English, J.G., Olsen, R.H.J., Lansu, K., Patel, M., White, K., Cockrell, A.S., Singh, D., Strachan, R.T., Wacker, D., and Roth, B.L. (2019). VEGAS as a Platform for Facile Directed Evolution in Mammalian Cells. Cell 178, 748–761.

155. Suzuki, T., Miller, C., Guo, L.T., Ho, J.M.L., Bryson, D.I., Wang, Y.S., Liu, D.R., and Söll, D. (2017). Crystal structures reveal an elusive functional domain of pyrrolysyl-tRNA synthetase. Nat. Chem. Biol. 13, 1261–1266.

156. Wong, B.G., Mancuso, C.P., Kiriakov, S., Bashor, C.J., and Khalil, A.S. (2018). Precise, automated control of conditions for high-throughput growth of yeast and bacteria with eVOLVER. Nat. Biotechnol. 36, 614–623.

157. Nasir, A., and Caetano-Anollés, G. (2015). A phylogenomic data-driven exploration of viral origins and evolution. Sci. Adv. 1, e1500527.

158. Wang, H.H., Isaacs, F.J., Carr, P.A., Sun, Z.Z., Xu, G., Forest, C.R., and Church, G.M. (2009). Programming cells by multiplex genome engineering and accelerated evolution. Nature 460, 894–898.

1.4 Figures

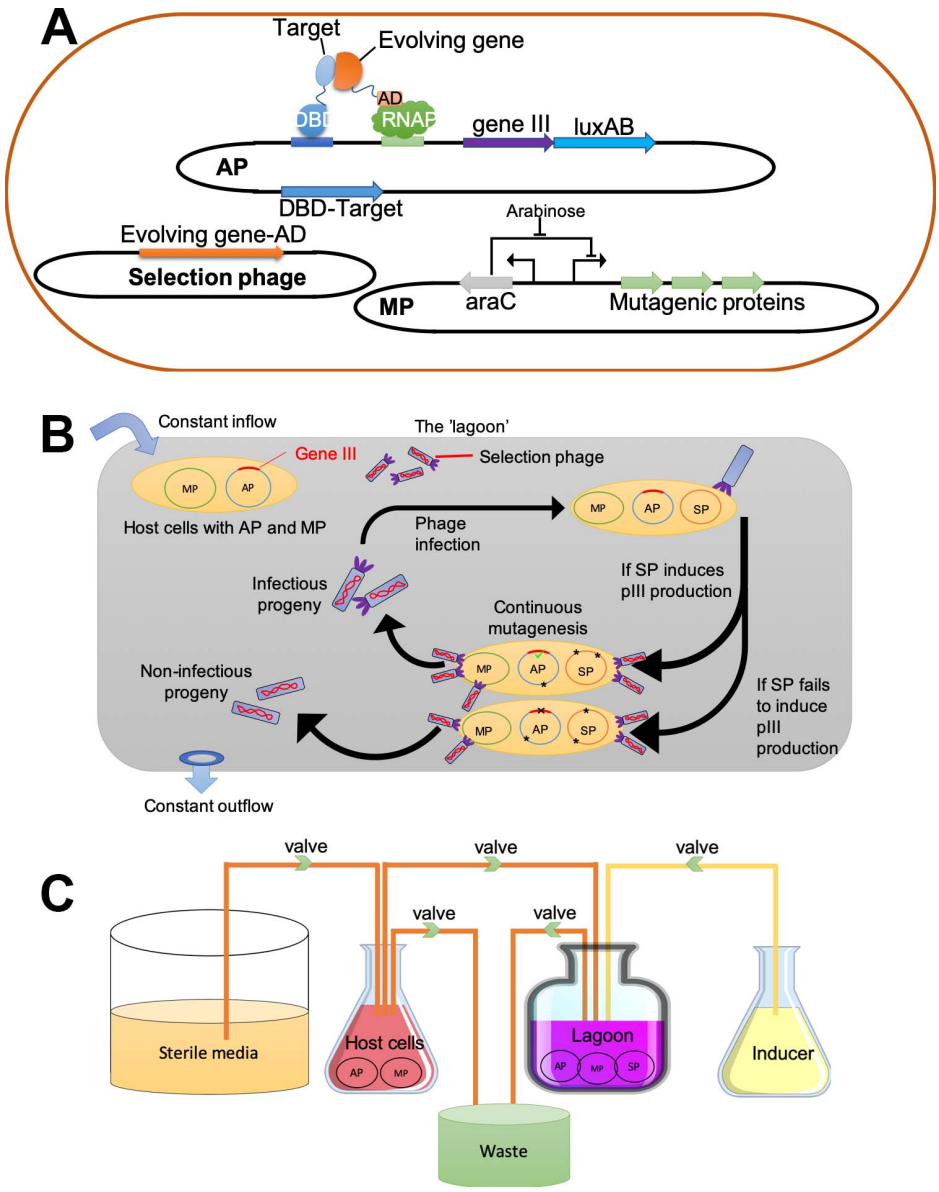

Fig. 1-1: The PACE system.

A) PACE system to evolve stronger protein-protein interactions. Evolving protein is encoded by selection phage (SP). On AP, expression of gIII is under the control of bacterial two-hybrid. The evolving gene is fused with a transcriptional activation domain (AD) and the target protein is fused with a DNA-binding domain (DBD). Transcription of

gIII can only be initiated if the evolving protein can bind to the target protein. Adapted from [144].

B) After phage infection of *E. coli* host cells, each SP encodes an evolving protein is subjected to random mutagenesis by MP. Protein III is required for the infectivity of M13 phages. Therefore, infectious phage progeny will be produced if the mutated evolving protein on SP can initiate production of pIII from AP. Non-infectious progeny will be produced if the mutated evolving protein cannot initiate production of pIII. This cycle can be repeated automatically to enrich infectious progeny.

C) The PACE devices. Media, *E. coli* cells, lagoons and mutagenesis inducers are in separate containers. Media is continuously pumped into the *E. coli* cell containers; *E. coli* cells and mutagenesis inducers are continuously pumped into the lagoons; both *E. coli* cells and the selection culture in the lagoons are kept at a pre-set volume, and overflow culture is pumped into waste containers.

B and C are adapted from [142].

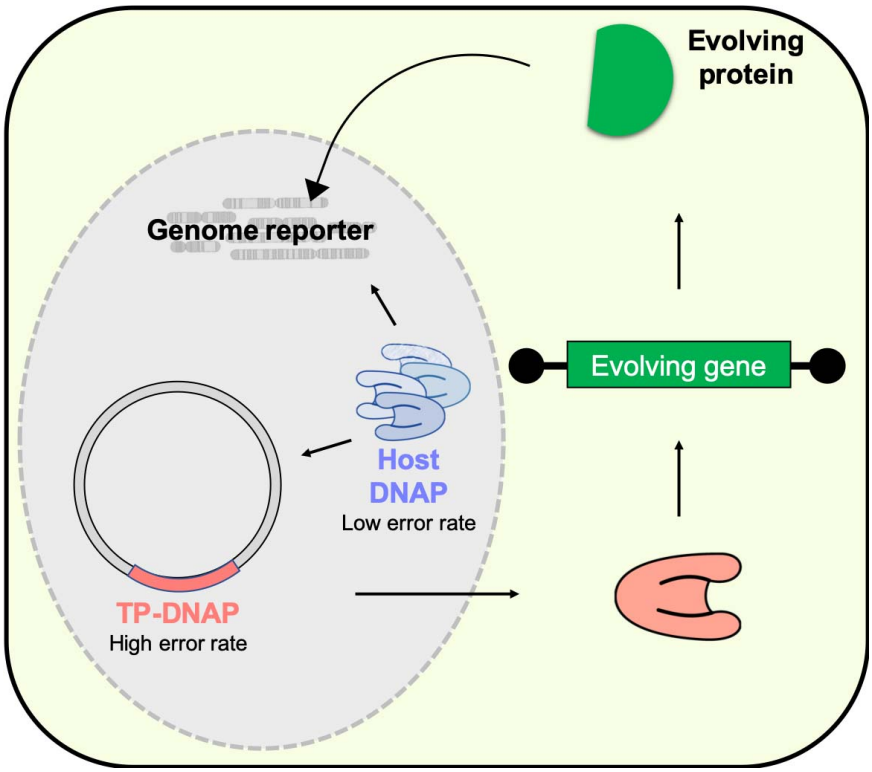

Fig. 1-2: The OrthoRep system.

Cytoplasmic linear plasmids carrying the evolving gene(s) are replicated by the low-fidelity TP-DNAP (terminal protein DNAP), while the high-fidelity host DNAP replicates other DNA.

Chapter 2: The universally conserved residues are not universally required for stable protein expression or functions of cryptochromes

2.1 Abstract

Universally conserved residues (UCRs) are invariable amino acids evolutionarily conserved among members of a protein family across diverse kingdoms of life. UCRs are considered important for stability and/or function of protein families, but it has not been experimentally examined systematically. Cryptochromes are photoreceptors in plants or light-independent components of the circadian clocks in mammals. We experimentally analyzed 51 UCRs of *Arabidopsis* cryptochrome 2 (CRY2) that are universally conserved in eukaryotic cryptochromes from *Arabidopsis* to human. Surprisingly, we found that UCRs required for stable protein expression of CRY2 in plants are not similarly required for stable protein expression of human hCRY1 in human cells. Moreover, 74% of the stably expressed CRY2 proteins mutated in UCRs retained wild-type-like activities for at least one of the photoresponses analyzed. Our finding suggests that the evolutionary mechanisms underlying conservation of UCRs or that distinguish UCRs from non-UCRs determining the same functions of individual cryptochromes remain to be investigated.

2.2 Introduction

Cryptochrome is one of the most ancient and common photoreceptors found in nature [1–5]. Cryptochromes are homologous to DNA photolyases that repair DNA lesions resulting from ultraviolet light [1, 2]. Cryptochromes do not repair DNA, instead they act as blue-light receptors to regulate photomorphogenic development in plants or transcriptional regulators to control circadian clock in plants and animals [4, 6]. Cryptochromes are composed of two domains, the universally conserved N-terminal PHR (Photolyase-Homologous Region) domain and the unstructured and poorly conserved CCE (Cryptochrome C-terminal Extension) domain [4]. The PHR domain of cryptochromes contains universally conserved residues (UCRs), which are invariable amino acids of members of a protein family from distantly related phylogenetic lineages [7]. It is intuitive that UCRs must be essential to the overall structure integrity of the proteins, such that mutations are prevented from accumulating during evolution of the protein family [8, 9]. And it is commonly hypothesized that UCRs determine the common structure elements that are universally important to the stability and functions of individual members of the protein family. Mutations altered in functionally important UCRs of cryptochromes from diverse lineages, such as *Arabidopsis* [1, 10–14], *Drosophila* [15], and mammals [16–18], have been reported. However, the structural and functional importance of UCRs has not been systematically investigated for signaling proteins. One technical difficulty to experimentally test the above hypothesis appears to lie in how to measure the specific activities or the protein abundance-adjusted physiological activities of signaling proteins in vivo. We developed a method to systematically analyze the functional importance of most (~90%) UCRs of *Arabidopsis*

CRY2 in vivo. Our results are consistent with the notion that the UCRs are important for the function of CRY2, because every mutation altered the UCRs of CRY2 impaired at least one activity of CRY2 tested. However, we found that most (~94%) stably expressed CRY2 proteins mutated in UCRs remained photophysiologically or photobiochemically active in one or more photoresponses examined. We also found that the UCRs required for stable protein expression of *Arabidopsis* CRY2 are not similarly required for stable protein expression of a human cryptochrome 1 (hCRY1). We further demonstrated that the specific functions of *Arabidopsis* and mouse cryptochromes are determined by both UCRs and non-UCRs of the respective cryptochromes, arguing it is important to elucidate the evolutionary mechanisms to distinguish UCRs from non-UCRs of cryptochromes or other protein families.

2.3 Results

2.3.1 The double and triple mutants of the universally conserved Trp-triad residues of CRY2 remained photobiologically active in vivo

All members of the photolyase/cryptochrome proteins contain three universally conserved tryptophan residues, referred to as Trp-triad [19]. Trp-triad are known to be critical to the photoreduction of cryptochromes in vitro, whereby the photon-absorbing chromophore FAD (Flavin Adenine Dinucleotide) is reduced [2, 19–21]. The Trp-triad-dependent photoreduction has been hypothesized to be the photobiochemical mechanism underlying the function of the photolyase/cryptochrome proteins [19, 21–29]. According to this hypothesis, photolyases and cryptochromes become biochemically and physiologically active upon electron transfer to FAD through the Trp-

triad residues. However, this hypothesis has been challenged by several genetics studies, whereby mutations of the Trp-triad residues abolish photoreduction of the mutant proteins in vitro without abolishing their physiological activities *in vivo* [12, 13, 30]. For example, we have previously reported that mutations altered in any one of the Trp-triad residues of *Arabidopsis* CRY1 or CRY2 completely abolished FAD photoreduction in vitro, but the mutant proteins remain photophysiologically active in vivo [12, 13]. It has been proposed that some small molecules in the cell, such as ATP, could bind to cryptochromes and rescue both the photoreduction activity and the physiological activities of a Trp-triad mutant [24]. However, a non-universally conserved tryptophan residue has been reported recently for an alternative electron transport pathway governing activities of the *Drosophila* dCRY mutations [31], and ATP fails to rescue the photoreduction activity of every Trp-triad single mutant of *Arabidopsis* CRY1 [13]. Therefore, we further investigated the Trp-triad hypothesis by asking whether the double and triple mutants of CRY2 altered in two or all three of the Trp-triad residues (Table 2-1) might abolish the activity of CRY2 in vivo. In this experiment, the wild-type CRY2 and the W-to-A (redox inactive) or W-to-F (redox inactive but structurally more similar to W) double or triple Trp-triad mutants of CRY2 were constitutively expressed as the GFP-fusion proteins in the *cry1cry2* mutant background. Although a transgenic study using the CRY2 native promoter would be the optimal, we used the constitutive promoter to be consistent with that used in the previous studies [12, 13]. We have previously shown that the GFP-CRY2 fusion protein is active in all photophysiological and photobiochemical responses tested [12, 13, 30]. The transgenic plants were analyzed for the CRY2 protein abundance and three best known photophysiological

activities of CRY2, including blue-light inhibition of hypocotyl elongation, blue-light stimulation of cotyledon unfolding, and promotion of floral initiation (Fig. 2-1, Fig. S2-1 and Table 2-1). It is interesting that double or triple mutations of the Trp-triad residues often eliminate the gain-of-function hyperactivity of the single mutations affecting residues W374 and W397. For example, plants expressing the single Trp-triad mutants W374A or W397A exhibited constitutive hypocotyl inhibition and floral acceleration phenotype, whereas the double mutant 2WA3 (W374A and W397A) exhibited blue light-dependent hypocotyl inhibition but constitutive floral acceleration phenotype (Fig. 2-1, Fig. S2-1 and Table 2-1). This observation suggests that the hyperactivities of the single mutants may result from structural changes, which may or may not be suppressed by the additional mutations. Remarkably, all double and triple Trp-triad mutants of CRY2 (Table 2-1), including those expressed at the levels markedly lower than that of the control GFP-CRY2, were able to rescue, to various extent, the defective phenotypes of the *cry1cry2* mutant parent, suggesting that all double and triple Trp-triad mutants of CRY2 remained physiologically active (Fig. 2-1 and Fig. S2-1). For example, a continuous imaging-based kinetic analysis clearly demonstrated that the three double mutants altered in any two of the three Trp-triad residues (2WA1 or CRY2^{W321A,W374A}, 2WA2 or CRY2^{W321A,W397A}, and 2WA3 or CRYs^{W374A,W397A}) and the triple mutant (3WA or CRY2^{W321A,W374A,W397A}) altered in all three Trp-triad residues were active in mediating blue-light inhibition of hypocotyl elongation, such that they all rescued the blue light-specific long-hypocotyl phenotype of the *cry1cry2* parent (Fig. 2-1B). Moreover, transgenic *cry1cry2* plants expressing the triple mutant (3WA) protein at the level comparable to that of the wild-type GFP-CRY2 control (Fig. 2-1A) fully rescued the

parental late-flowering phenotype (Fig. 2-1, E-G). These results confirm that the universally conserved Trp-triad residues are not essential to photophysiological functions of CRY2 in vivo, although they are essential to the photoreduction of CRY2 in vitro [12].

2.3.2 The UCRs required for stable protein expression of *Arabidopsis* CRY2 are not equally required for stable protein expression of human hCRY1

The observation that none of the three universally conserved Trp-triad residues of CRY2 are universally required for all three physiological functions examined seems counterintuitive, because they are universally conserved residues that are commonly considered to be important to preserve integrity of the appropriate conformation of members of a protein family. We hypothesize that the structure elements commonly preserved for members of a protein family may not be required for all functions of individual members under all experimental conditions tested, but such "partial functional requirement" is sufficient to prevent accumulation of any mutation in nature during evolution. Cryptochromes are ancient proteins evolutionarily conserved in all major lineages including plants and human [4], a systematic analysis of mutations of all or most universally conserved residues or UCRs of cryptochromes, such as *Arabidopsis* CRY2, would allow us to test this hypothesis. Based on multiple sequence alignment analyses, we identified 57 UCRs of *Arabidopsis* CRY2, which are defined for the present study as the invariable amino acids that are conserved in the same position of cryptochromes of *Arabidopsis* and human (Fig. 2-2A, Fig S2-2 and Table 2-3, Table 2-4). Those 57 residues are also universally conserved among an arbitrarily selected

group of 24 cryptochromes from three plant families, one algae family, and four animal families (Fig S2-2 and Table 2-2).

None of the 57 UCRs of CRY2 are alanine, so we changed each of them individually to alanine by site-directed mutagenesis [32], and prepared transgenic plants each constitutively expressing one site-specific CRY2 mutants as GFP-CRY2 fusion protein in the *cry1cry2* mutant background. Among the 57 UCRs of CRY2, we successfully obtained transgenic lines expressing 51 UCR mutants (Fig. 2-2A, Fig. S2-2 and Table S2-3). Among these, stable protein expression in plants were observed for 61% (31/51) CRY2 UCR mutants examined, whereas 39% (20/51) failed to stably express detectable amount of mutant proteins in all five independent transgenic lines of each mutant examined. Because all 20 CRY2 UCR mutant genes expressed mRNAs in plants at the level comparable to that of the "wild-type" GFP-CRY2 control (Fig. 2-2B, Fig. S2-3, A-C and Table 2-4), these UCRs are most likely required for translation or protein stability of *Arabidopsis* CRY2 in vivo. To test whether these UCRs are similarly required for the stable protein expression of another cryptochrome, we analyzed protein expression of human cryptochrome 1 (hCRY1) mutated in the equivalent UCRs. Interestingly, we found that all the hCRY1 UCR mutant proteins tested expressed at the levels comparable to that of the "wild-type" hCRY1 control in HEK293T (Human Embryonic kidney) cells (Fig. 2-2C and Fig. S2-3D, E), which is in stark contrast to their *Arabidopsis* counterparts that expressed mRNA but not protein in plant cells (Fig. 2-2B). The mouse mutant mCRY1^{F257A}, which is an UCR mutant equivalent to *Arabidopsis* CRY2^{F253A} that failed to stably express the protein in plants (Fig. S2-3, A and B) has also been reported by others to stably express the mutant protein in both HEK293 and

MEF (Mouse Embryonic Fibroblasts) cell lines [16]. We concluded that the UCRs of cryptochromes are not universally required for stable protein expression of cryptochromes.

2.3.3 UCRs are not universally required for photophysiological activities of CRY2

CRY2 is nuclear protein that undergoes blue light-dependent dimerization [33], phosphorylation [34], photobody formation [35, 36], ubiquitination and degradation [37] in nucleus. We examined the subcellular localization of the stably expressed CRY2 UCR mutant proteins and found all of them still locate in the nucleus (Fig. S2-4). We next tested the photophysiological activities of the 31 CRY2 UCR mutant proteins that are stably expressed in plants. Like enzymes, the cellular concentration of signaling proteins, such as photoreceptors, is expected to determine the total activity of the photoreceptor and photo-responsiveness of the plants [12]. Because it is technically difficult, if not impossible, to obtain different mutant lines that express the identical amount of different CRY2 mutant proteins by either reverse genetics or forward genetics methods, we determined the relative specific-activity of individual CRY2 mutants, based on the standard curves constructed using transgenic lines expressing the “wild-type” GFP-CRY2 protein at different levels (Fig. 2-3). Specifically, we first screened and selected five transgenic lines, referred to as L1, L2, L3, L4, and L5, that express the “wild-type” GFP-CRY2 fusion protein at gradually increased levels, with L5 expressing GFP-CRY2 at the highest level. We quantified the relative abundance of the GFP-CRY2 protein in those five control lines by the quantitative fluorescence immunoblot assay using an Odyssey imager (LI-COR Biotechnology). We next analyzed

photomorphogenic phenotypes of the transgenic lines L1 to L5, including blue-light inhibition of hypocotyl elongation, blue-light stimulation of cotyledon unfolding, and photoperiodic promotion of flowering (Fig. 2-3, A-C). We then constructed the standard curves, whereby the relative protein abundance of the "wild-type" GFP-CRY2 protein in the L1, L2, L3, L4, and L5 transgenic lines were plotted against the relative light responsiveness of the respective transgenic lines. In the three standard curves shown in Fig. 2-3 (A-C), individual light responses of plants are used as the proxy of the relative photophysiological activities of the "wild-type" GFP-CRY2 expressed at the levels measured in the transgenic lines L1 to L5, with both the protein abundance and CRY2 activity of the L5 line set as 100%. As expected, the "wild-type" GFP-CRY2 protein shows strong correlations between the protein abundance and the photophysiological activities inferred from the photomorphogenic phenotypes of the lines L1 to L5, including blue light inhibition of hypocotyl elongation ($r=0.9636$, $p<0.01$), blue light stimulation of cotyledon unfolding ($r=0.8585$, $p<0.01$), and CRY2 promotion of floral initiation ($r=0.9701$, $p<0.001$) (Fig. S2-5, A-C). This correlation is better observed at relatively lower levels of the CRY2 protein (Fig. 2-3A-C), which is consistent with the expectation that the total activity of CRY2 is apparently saturable. All three standard curves showed saturation of activities of GFP-CRY2 at the approximate protein abundance between that of the transgenic lines L4 and L5, indicating the appropriate sensitivity ranges for our analyses. The "wild-type" GFP-CRY2 protein has been considered similar to the endogenous CRY2, but its activities have never been quantitatively compared with that of the endogenous CRY2 [12, 35]. We plot the relative abundance and activities of the endogenous CRY2 to the standard curves of

GFP-CRY2 and compared the relative specific-activity of the two proteins. Fig. 2-3 (A-C) shows that the endogenous CRY2 protein expresses at the level approximately 40% that of the "wild-type" GFP-CRY2 of the transgenic line L1, which expressed the lowest level of GFP-CRY2 among the standard lines L1 to L5 (Fig. 2-3, A-C and Table 2-3). However, the endogenous CRY2 has the similar photophysiological activity (~100%) mediating blue light inhibition of hypocotyl growth as that of GFP-CRY2 of the L1 line (Fig. 2-3A and Table 2-3), demonstrating that the relative specific-activity, or the protein abundance-adjusted photophysiological activity, of the "wild-type" GFP-CRY2 is about 40% (40/100) that of the endogenous CRY2 mediating this photoresponse. Similarly, we estimated that the relative specific-activity of GFP-CRY2 mediating blue light promotion of cotyledon unfolding or CRY2 promotion of floral initiation are approximately 20% (40/200) or 16% (40/250) that of the endogenous CRY2 (Fig. 2-3, B-C and Table 2-3). These results suggest that the endogenous CRY2 is approximately 2.5- to 6.25-fold more active than that of the "wild-type" GFP-CRY2. The lower relative specific-activity of GFP-CRY2 in comparison to that of the endogenous CRY2, which is likely caused by the structure disturbances of GFP fusion, minimizes a potential bias in our study due to saturation of activities of the overexpressed GFP-CRY2 mutant proteins tested.

We next estimated the relative specific-activity of the individual CRY2 mutants expressed as the GFP-CRY2 fusion proteins, assuming that the GFP fusion has similar effects on the photophysiological activities of the mutant GFP-CRY2 proteins as that on the "wild-type" GFP-CRY2. In this experiment, the protein abundance of the mutant GFP-CRY2 and the light responsiveness of the respective transgenic lines were

measured as described above and plotted to the standard curves of the “wild-type” GFP-CRY2 (Fig. 2-3, A-C). The relative specific-activity of all mutants were classified according to their relative positions in the individual standard curves of the “wild-type” GFP-CRY2 proteins. The photophysiological activities of the 31 CRY2 UCR mutant proteins are classified as hypermorph (HYPER) for those positioned higher than the upper limits of the 95% prediction bands of the standard curves (Fig. 2-3, A-F), wild-type-like (WTL) for those positioned within 95% prediction bands of the standard curves, hypomorph (HYPO) for those positioned below the lower limits of the 95% prediction bands of the standard curves but higher than 20% of the standard curves, or loss-of-function (LOF) for those positioned below 20% of the standard curves.

All 31 stably expressed CRY2 mutant proteins exhibited defects in at least one of the three photophysiological activities examined. However, only two mutants, D387A and G427A, lost all three photophysiological activities examined, and they also lost the blue light-induced homodimerization activity (Fig. S2-6A). 58% (18/31) stably expressed CRY2 mutant proteins exhibited the wild-type-like activity mediating at least one of the three photophysiological activities measured (Fig. 2-3 and Table 2-3). Figure 2-3 (D-F) shows that among the 31 stably expressed CRY2 mutants analyzed, 9.7%, 35.5%, or 45.2% are classified as wild-type-like that exhibited the relative specific-activity similar to that of the “wild-type” GFP-CRY2 for the blue-light inhibition of hypocotyl elongation, blue-light stimulation of cotyledon unfolding, and promotion of flowering, respectively (Fig. 2-3, D-F). It is interesting that about four times as many CRY2 mutants showed wild-type-like activity promoting floral initiation (45.2%) or cotyledon unfolding (35.5%) in comparison to that of the mutants retained wild-type-like

activity inhibiting hypocotyl elongation (9.7%). One possible interpretation of this observation is that different structural elements determined by different UCRs of CRY2, albeit their universal conservativity, contribute to different physiological activities of the photoreceptor, such that individual structure disturbance resulting from different mutations manifest differently in different activities of CRY2. In summary, our results demonstrate that among the 51 UCRs of CRY2 analyzed, 20 may determine protein stability, 2 are universally required for all functions of CRY2, whereas 29 are not universally required for all activities of CRY2 examined under the experimental conditions used.

2.3.4 UCRs are not universally required for the photobiochemical activity of CRY2

It is intuitive that UCRs of a protein family should be universally important for the members of the respective protein family, otherwise mutations would have accumulated at those positions throughout evolution in at least some members of some lineages. It is somewhat unexpected that more than half (29/51) of the UCRs of CRY2 affect only a subset of the functions of CRY2. We next analyzed how mutations of the UCRs affecting the photobiochemical activity of the CRY2 protein. CRY2 undergoes blue light-dependent phosphorylation, ubiquitination, and degradation, presumably leading to the negative feedback regulation of CRY2 activity and plant photosensitivity [37, 38]. Since different transgenic lines expressing different CRY2 mutants at different levels, we first examined whether the difference in protein abundance might affect the blue light-induced degradation of the “wild-type” GFP-CRY2 protein in three transgenic lines

expressing the “wild-type” GFP-CRY2 protein at different levels (Fig. S2-7, A-D). Figure S2-7D showed that the “wild-type” GFP-CRY2 protein expressed at different levels have the similar half-life ($^{60}T_{1/2}$, see Fig. 2-3J) of approximately 30-40 min when etiolated seedlings were exposed to blue light of the indicated fluence rate, which is also similar to the half-life of the endogenous CRY2 reported previously (Fig. S2-7, A-D) [37]. These results suggest that the rate of blue light-induced proteolysis of CRY2 is not significantly affected by fusion to GFP or the absolute abundance of CRY2 protein, at least within the ranges of protein abundance and blue-light intensities examined. We next analyzed the half-life of all 31 stably expressed GFP-CRY2 mutant proteins in etiolated seedlings exposed to blue light. Similar to the photophysiological activities of CRY2, we classified the blue light-induced proteolysis activity of the CRY2 UCR mutant proteins into four groups for the experimental conditions of 2-hour blue light treatment at the fluence rate of 60 μmol m^{-2} s^{-1} (Fig. 2-3, H-J. Fig. 2-4A, Fig. S2-7, E and F and Table 2-3), including hypermorphic (HYPER, e.g. P416A) for those degraded faster than the “wild-type” GFP-CRY2 in response to blue light; hypomorphic (HYPO, e.g. L348A) for those degraded slower than the “wild-type” GFP-CRY2 but achieved at least 50% of degradation within 2 hours of blue light exposure; loss-of-function (LOF, e.g. D387A) for those failed to achieve 50% of degradation within 2 hours of blue light exposure; and wild-type-like (WTL, e.g. S257A) for those degraded at comparable rates of the “wild-type” GFP-CRY2.

The two mutations, D387A ($CRY2^{D387A}$) and G427A ($CRY2^{G427A}$) that lost all their photophysiological activities tested, also showed no photobiochemical activity in light-dependent proteolysis (Table 2-3). D387A is a known chromophore-less mutant that

fails to bind to the FAD chromophore [15, 39], which is presumably required for all light-dependent activities of any cryptochrome family proteins. G427A is located far away from the FAD-binding pocket in the modeled structure of CRY2, and it is universally conserved in the photolyase/cryptochrome of eukaryotes (Fig. 2-2A, Fig. S2-2 and Table 2-3). Why the mutation at G427 has the same detrimental effect to all CRY2 activities examined as that of the FAD-less D387A mutant remains unclear. Also analogous to the photophysiological activity, a large number (41.9%, 13/31) of the stably expressible CRY2 UCR mutant proteins (e.g. S257A) analyzed exhibited the half-life ($^{60}T_{1/2}$) similar to that of "wild-type" GFP-CRY2 (Fig. 2-3I, Fig. S2-7, E and F and Table 2-3). This result is consistent with the observation that UCRs are not required for all activities of CRY2. Moreover, the altered activity of light-induced proteolysis of most CRY2 mutants, except D387A and G427A, do not correlate with the other three photophysiological activities examined (Fig. S2-5, D-F and Fig. S2-5B). For example, the CRY2^{P416A} mutant, which showed wild-type-like activity in promoting flowering and lower activity in mediating blue-light inhibition of hypocotyl elongation or blue light stimulation of cotyledon unfolding, degrades faster than the "wild-type" GFP-CRY2 in response to blue light (Table 2-3). In contrast, the CRY2^{L370A} mutant, which is classified as hypomorphic for the blue light-induced proteolysis, appear wild-type-like or hypermorphic for all the three photophysiological activities examined (Table 2-3). Taken together, our analyses of the photophysiological and photobiochemical activities of the CRY2 UCR mutants may be interpreted by a hypothesis that the structure elements determined by UCRs are required for specific functions but not universally required for all functions of members of a protein family.

2.3.5 Neighboring UCRs have similar effects on the CRY2 function and regulation

We tested the above hypothesis by examination of whether neighboring UCRs, which are presumably associated with the same structure elements, may determine the same functions of CRY2. We analyzed the functional effects of two CRY2 UCR mutants altered in the residues P448 and Y467. Because these two UCRs appear to locate in the close vicinity of each other (~2.4 Å for the closest atoms) (Fig. 2-2A) and therefore likely associated with the same structural element. We analyzed the effects of the mutations of P448 and Y467 on different functions of CRY2 under different experimental conditions. First, the P448A (GFP-CRY2^{P448A}) and Y467A (GFP-CRY2^{Y467A}) mutants exhibited wild-type-like activity in blue light-induced proteolysis in etiolated seedlings exposed to blue light (Fig. 2-4A and Table 2-3). Second, both mutants exhibited similarly complex fluence rate-dependent activity mediating the blue-light inhibition of hypocotyl growth (Fig. 2-4, B and C). When grown under blue light with the fluence rate lower than 15-20 μmol m^{-2} s^{-1}, both mutants developed hypocotyls slightly longer than that of the GFP-CRY2 control (Fig. 2-4C). However, seedlings expressing either mutants developed hypocotyls indistinguishable from that of the GFP-CRY2 control when grown under blue light with the fluence rate higher than 15-20 μmol m^{-2} s^{-1} (Fig. 2-4, B and C). Because seedlings expressing the P448A and Y467A mutant proteins at the similar level, which are 10-20% that of the GFP-CRY2 control (Table 2-3), these two mutants appear to have higher relative specific-activity than that of the “wild-type” GFP-CRY2 protein mediating light inhibition of hypocotyl elongation in

at least high light. Third, the P448A and Y467A mutants exhibited similar wavelength-dependent activities promoting floral initiation (Fig. 2-4, D-I). The *cry2* mutant is known to exhibit delayed flowering in white light or blue-plus-red light but normal flowering-time when grown under monochromatic blue light, whereas the *cry1cry2* double mutant exhibits delayed flowering in both white light and monochromatic blue light [11, 40]. These wavelength-dependent flowering-time phenotypes have been interpreted to result from two different modes of actions of CRY2 in promoting floral initiation: a phytochrome B (PHYB)-dependent pathway and a PHYB-independent pathway [11, 40–42] In the PHYB-dependent pathway, CRY2 exerts blue light-dependent inhibition of the red light-dependent suppression of flowering by PHYB, such that this function of CRY2 is dependent on both blue light and red light. In the PHYB-independent pathway, CRY2 acts redundantly with CRY1 to promote floral initiation directly, such that this function of CRY2 is dependent on blue light but not red light [40]. Plants expressing the P448A and Y467A mutants flowered later than the wild-type plants when they were grown in white light comprised of both blue and red wavelengths, suggesting both mutants are impaired in the PHYB-dependent activity of CRY2 (Fig. 2-4, D-F and Table 2-3). However, plants expressing the “wild-type” GFP-CRY2, P448A, or Y467A mutant proteins all rescued the later-flowering phenotype of *cry1cry2* when grown in continuous blue light (Fig. 2-4, G and H), indicating that neither mutant is compromised in their activity mediating PHYB-independent promotion of floral initiation. Therefore, the structure elements determined by the neighboring P448 and Y467 residues are required for the PHYB-dependent but not PHYB-independent activity of CRY2 (Fig. 2-4 I). Taken together, these results are consistent with the hypothesis that the structure elements

determined by UCRs are required for specific functions of CRY2 but not universally required for all functions of CRY2.

2.4 Discussion

UCRs are commonly considered critical for the structural integrity common to all members of a protein family. To our knowledge, this notion has not been systematically tested experimentally. In the present study, we analyzed the in vivo relative specific activities of UCR mutations of the *Arabidopsis* blue light receptor CRY2, using a standard curve-based quantitative approach. Our result that all UCR mutations of CRY2 exhibited at least a minor impairment in at least one of the four physiological or biochemical activities examined is consistent with the expectation that UCRs are evolutionarily conserved for functional reasons. On the other hand, it is interesting that none of the UCRs required for stable protein expression of *Arabidopsis* CRY2 in plants is required for stable protein expression of human hCRY1 in a human cell line (Fig. 2-2). Moreover, 74% (23/31) of the stably expressible CRY2 mutant proteins exhibited the wild-type-like activity mediating at least one of the 4 photophysiological and photobiochemical responses examined (Fig. 2-3 and Table 2-3). These results and our follow-up analyses (Fig. 2-4) demonstrate that UCRs of CRY2 are not universally required for protein stability or all functions of cryptochromes. UCRs of closely related cryptochromes may also exert different effects on the same function of the closely related family members. For example, the *Arabidopsis* CRY2 UCR mutant, $CRY2^{W331A}$, exhibits the loss-of-function phenotype for the blue-light inhibition of hypocotyl growth response in the present study (Table 2-3), whereas the $CRY2^{W331A}$ equivalent mutant of

Arabidopsis CRY1, CRY1^{W334A}, exhibited wild-type-like activity for the same photoresponse in our previous study [13]. The mechanisms underlying different functions of the equivalent residues in different cryptochromes may include different post-translational protein modifications. This may be illustrated by the comparison of protein phosphorylation and functions of the *Arabidopsis* CRY2^{S257A} mutant examined in the present study with that of the corresponding mouse mCRY1^{S261A} mutant reported recently [18]. *Arabidopsis* CRY2 or mouse mCRY1 are phosphorylated in at least 24 or 27 residues, respectively [18, 34]. None of the phosphorylated residues of *Arabidopsis* CRY2 is universally conserved. But two of the 27 phosphosites of mouse mCRY1, S252 and S261 that correspond to the unphosphorylated S248 and S257 of *Arabidopsis* CRY2, are universally conserved (Fig. S2-2 and Table 2-3). The *Arabidopsis* CRY2^{S248A} or CRY2^{S257A} mutants, which are not expected to directly impair CRY2 phosphorylation [34], fully or partially rescued three photophysiological phenotypes of the *cry1cry2* mutant plants, respectively (Fig. 2-3 and Table 2-3). In contrast, the mCRY1^{S261A} mutant that directly impaired phosphorylation of mCRY1 failed to rescue the arrhythmic phenotype of the *mCry1mCry2* knockout mice [18]. On the other hand, the mCRY1^{S252D} mutant altered in the other phosphorylated UCR exhibits short period and lower amplitude [18]. These results demonstrate that UCRs of different cryptochromes can differentially affect protein phosphorylation to impact the functions of respective cryptochromes differently.

Results of our studies demonstrate that the structure elements determined by UCRs common to different members of a protein family are not universally required for the protein stability, post-translational modification, and functions of the individual family

members of cryptochromes. We hypothesize that common structure elements associated with the UCRs may evolve to play different subsets of functions in different members of a protein family, and those diverse subsets of functions all contribute to the long-term fitness of the host organisms, sanctioning their universal conservation in evolution. It is conceivable that cryptochromes of different evolutionary origins, which are believed to evolve independently from the ancestral DNA photolyases, may adopt the same folds for different but functionally essential purposes specific to individual members of a protein family. For example, plant cryptochromes may rely on certain UCRs for the light-dependent homodimerization whereas metazoan CRYs might rely on the similar UCRs for light-independent interaction with transcription activators. However, how those UCRs associated with different functions of different members of the protein family are universally conserved remains to be elucidated. Moreover, UCRs are apparently not the only structural elements that are essential for function. Mutations of many non-UCRs are also known to affect the same function impacted by UCRs of cryptochromes in both *Arabidopsis* and mouse, although functionally defective UCR mutants appear in higher percentages of total UCRs than that of the functionally defective non-UCR mutants among the total non-UCRs in both types of cryptochromes (Fig. 2-2D and Table 2-5). Therefore, how are the UCRs distinguished from non-UCRs during evolution needs to be further investigated to better understand the evolutionary history and the structure-function relationship of the cryptochrome family of proteins that play important functions in plant development and human health.

2.5 Materials and Methods

2.5.1 Multiple sequence alignment and structure simulation

Species, residues forming PHR domain of each protein used for multiple sequence alignment, and NCBI protein accession numbers (Fig. S2-2) were listed in Table 2-2, respectively. Multiple sequence alignment was conducted using T-Coffee [43]. The resulting Clustalw_aln files were uploaded onto ESPript 3.0 [44] to generate a black and white version of the alignment, and then manually edited in Adobe Photoshop CC 2017 to add blue shades at desired positions.

The CRY2 structure was simulated using SWISS-MODEL [45] from CRY2 full-length protein sequence based on crystal structure of CRY1 (PDB: 1U3C) [46].

2.5.2 Plasmid construction and plant materials

All *Arabidopsis* plant lines used in this study were in Columbia (Col) background. The wild type plants used in this study are *rdr6-11* [47]. The Ti plasmid pFGFP [34] was modified from pCambia3301. The coding sequence (CDS) of wild-type or site-specific mutants of CRY2 were PCR amplified and incorporated into the BamHI site of pFGFP using In-Fusion Cloning Kit (Clontech). The resulting constructs were P_{ACTIN2}::FLAG-EGFP-CRY2::T_{35S} and were introduced into the *cry1cry2rdr6* plants by standard floral dip method [48]. The *cry1cry2rdr6* lines were acquired by crossing *cry1-304* [40], *cry2-1* [11] and *rdr6-11* (which suppresses gene silencing) [47]. The transgenic T1, T2 and T3 populations were screened and maintained on compound soil sub-irrigated with the Basta solution [48].

The pQCMV-FLAG-EGFP plasmid was modified from pEGFP-N1 vectors (Clontech) by A) inserting DNA sequences for a Kozak motif, a FLAG epitope tag, and a flexible protein linker (PAPAP) (gccaccATGGACTACAAAGACGATGACGACAAGgctACTAGTgccCCTAGGgctCCAGCTCCAGCTCCA) between SacI site (GAGCTC, "GAG" remained in the new pQCMV-FLAG-EGFP plasmid, whereas "CTC" was removed) and start codon of EGFP [33]; B) inserting DNA sequences for a flexible protein linker (PAPAP) and multiple cloning site (MCS) containing KpnI and SacI recognition site (tccggaCCAGCTCCAGCTCCAgctGGTACCgctGAGCTCgct) right before the stop codon (TAA) of EGFP. The CDS of wild-type or site-specific mutants (F9A, R10A, L13A, D17A, P19A, L59) of human hCRY1 were PCR amplified and cloned into the KpnI site of pQCMV-FLAG-EGFP. The resulting constructs were P_{CMV}::FLAG-EGFP-hCRY1::T_{SV40}.

The pCMV plasmid was described previously [34]. The CDS of wild-type or site-specific mutants (F9A, R10A, L13A, R14, D17A, P19A, L59, D110A, L132A, F257A, R293A, D341A, H354A, R358A, F381A, D389A, D423A, P424A, P440A, W448A) of human hCRY1 were PCR amplified and cloned into the BamHI site of pCMV plasmid. The resulting constructs were P_{CMV}::MYC-hCRY1::$T_{\beta\text{-globin}}$.

2.5.3 Plant growth conditions and physiological analyses

For hypocotyl inhibition assays in darkness or blue light that were not used for growth kinetics analysis, seeds were sterilized and sown onto fresh-made MS Agar (0.8%) plates, subjected to 4˚C cold treatment in darkness for 4 days, exposed to white

light at room temperature for 24 hours, and then put into indicated light conditions at room temperature for 5 days. The resulting seedlings were sandwiched between two plastic sheets (one transparent, the other black), scanned and measured by Fiji (NIH).

For image based hypocotyl growth kinetics analyses [49], seeds were sterilized and sown onto MS Agar (0.8%) plates, subjected to 4˚C cold treatment for 4 days and exposed to white light at room temperature for 24 hours. The imbibed seeds were then transferred onto 100mm x 100mm squire MS Agar (0.8%) plates with grids. The plates were placed vertically under blue light (15 μmol m^{-2} s^{-1}) and images were captured each hour for the next seven days, by using a CCD camera (Jinghang JHSM500B) equipped with a prime macro lens. Image acquisition was controlled by a custom-designed software. Images captured between 48-144 hours post exposure of blue light were manually measured by using Fiji (NIH) to get hypocotyl length. Three seedlings were measured for each genotype. Absolute activities (AA) were acquired as reciprocal of the slope of a linear regression of growth kinetics of 48-96 hours. Relative specific activity of hypocotyl inhibition of CRY2 (Fig. 2A) was calculated by the following formula:

$$Relative\ specific\ activity\ (\%) = \frac{AA_{mutant} - AA_{cry1cry2}}{AA_{GFP-CRY2\ L5} - AA_{cry1cry2}} \quad (1)$$

For cotyledon unfolding assay in darkness, seeds were sterilized and sown onto MS Agar (0.8%) plates, subjected to 4˚C cold treatment for 4 days and exposed to white light at room temperature for 24 hours, and then put into corresponding light conditions for 5 days before analysis. Seedlings were carefully sandwiched between adhesive sides of transparent tapes without disturbing cotyledon unfolding angles and then taped onto black paper for scanning. More than 20 seedlings were measured for each genotype. For cotyledon unfolding assay in blue light, plants were similarly prepared as

in hypocotyl growth kinetics analyses. Cotyledon unfolding angles were measured by drawing lines between the shoot apical meristem (SAM) and tips of cotyledons in Fiji (NIH) as previously described [50]. Three seedlings were measured for each genotype. Cotyledon unfolding activities of CRY2 shown in Fig. 2-2B were calculated by the following formula:

$$Relative\ specific\ activity\ (\%) = \frac{Angle_{mutant} - Angle_{cry1cry2}}{Angle_{GFP-CRY2\ L5} - Angle_{cry1cry2}} \quad (2)$$

For measuring flowering time in blue light, seeds were sown in soil, subjected to 4˚C cold treatment for 4 days, exposed to white light at room temperature for 24 hours, and then put into blue light (70-80 μmol•m^{-2}•s^{-1}) as previously described [40]. For measuring flowering time in the long-day (16 h day/ 8 h night) or short-day (16 h day/ 8 h night) period, seeds were sown in soil, subjected to 4˚C cold treatment for 4 days, and then moved into corresponding light conditions. Days to flowering and leaf number were counted daily. The day when there was at least 1 cm of inflorescence with visible floral meristem at top was regarded as day of flowering. Only rosette leaves were counted. Days to flowering (Days) in the long-day period were used to calculate floral promotion activities shown in Fig. 2-2C. The floral initiation activities were calculated by the following formula:

$$Relative\ specific\ activity\ (\%) = 1 - \frac{Days_{mutant} - Days_{GFP-CRY2\ L5}}{Days_{cry1cry2} - Days_{GFP-CRY2\ L5}} \quad (3)$$

Standard curves were created by fitting data points of wild-type GFP-CRY2 L1–L5 into the hyperbola non-linear regression model in GraphPad Prism version 8.0.2 (159) for Mac OS X (GraphPad Software, San Diego, California USA, www.graphpad.com). 95% prediction curve were automatically calculated by GraphPad Prism.

Normalized activities used in Fig. 2-3G were calculated by the following formula:

$$Normalized\ activity\ (\%) = \frac{Relative\ specific\ activity_{mutant}}{Relative\ specific\ activity_{GFP-CRY2}} \quad (4)$$

where Relative specific activity$_{GFP\text{-}CRY2}$ was determined case by case: first, protein abundance of the respective mutants was identified. The protein abundance was then introduced into the formulas of standard curves to get the respective Relative specific activity$_{GFP\text{-}CRY2}$.

Long-day (16 h day/8 h night) and short-day (8 h day/16 h night) photoperiod-treated plants were grown in walk-in growth chambers at 22˚C, 65% relative humidity under cool white fluorescent tubes. Light-emitting diode (LED) was used to obtain monochromatic blue light (peak 450 nm; half-bandwidth of 20 nm).

2.5.4 Immunoblot and blue-light induced proteolysis of plant samples

To prepare protein extracts, plant materials were dipped into liquid N_2 and homogenized by TissueLyser (QIAGEN). The resulting plant tissue powders were added 0.8x volume of powder of protein extraction buffer (120mM Tris-HCl pH6.8; 100mM EDTA pH 8.0, 4% w/v SDS, 10% v/v 2-Mercaptoethanol, 5% Glycerol and 0.01% Bromophonol Blue), boiled for 8 minutes, and then centrifuged with table top centrifuges at top speed for 10 minutes. The resulting protein extract supernatant was separated by home-made 10% (for checking protein abundance) or 8% (for assaying proteolysis) SDS-PAGE gels and transferred to Pure Nitrocellulose Blotting Membranes (BioTrace NT, Pall Life Sciences) using wet electroblotting system (Bio-Rad, Hercules, CA, USA). Ponceau S Red solution (0.1%(w/v) Ponceau S; 5%(v/v) acetic acid) was used to stain transferred membranes to gauge transferring efficiency. The membranes

were then cut horizontally along ~70kD for separate incubation with primary and secondary antibodies. For immunoblot signals captured by the Odyssey® CLx Infrared Imaging System (LI-COR Inc, Lincoln, NE, USA), membranes were blocked with 0.5% Casein in PBS solution, blotted with primary antibodies in 0.5% Casein in PBST solution, and then blotted with fluorescent secondary antibodies (ThermoFisher, A11357, A11369) in 0.5% Casein in PBST solution. Images captured by Odyssey® CLx System (LI-COR) were processed with Image Studio Lite software (LI-COR) and organized with Adobe Photoshop CC 2017. Primary antibodies used in this study were: rabbit-anti-CRY2 (1:3000, home-made) [11], mouse-anti-ACTIN11 (1:3000, 26F7, Ab-mart, Inc., Berkeley Heights, NJ, USA) and rabbit-anti-HSP90 (1:3000, sc-33755, Santa Cruz Biotechnology, Inc., Dallas, TX, USA). Secondary antibodies used here were: goat-anti-rabbit IgG (1:15000, A11369, Thermo Fischer Scientific, Inc., Waltham, MA) and goat-anti-mouse IgG (1:15000, A11357, Thermo Fischer Scientific), both conjugated to Alexa Fluro® 790. For immunoblot signals captured by exposure to X-ray film, membranes were blocked with 5% skimmed milk in PBST solution, blotted with primary antibodies in PBST solution, and then blotted with secondary antibodies in PBST solution. After blotted with secondary antibodies, the membranes were incubated in the home-made ECL solution (Solution A: 100mM Tris-HCl pH8.5; 0.2mM coumaric acid; Solution B: 100mM Tris-HCl pH 8.5; 1.25mM luminol; Right before use, mix 3ml Solution A with 3 mL Solution B and add 2μl 30% H_2O_2) and exposed to X-ray films. The resulting films were scanned and organized by Adobe Photoshop CC 2017. Primary antibodies were the same as above. Secondary antibodies used here were donkey-anti-rabbit IgG (1:10000, NA9340-1ML, GE Healthcare, Chicago, IL, USA) and

sheep-anti-mouse IgG (1:10000, NA9310-1ML, GE Healthcare), both conjugated to HRP.

To quantify protein abundance of CRY2, fluorescent signals captured by Odyssey® CLx System (LI-COR) were quantified by an internal method of Image Studio Lite software (LI-COR). The resulting signals were used to calculate protein abundance by the following formula:

$$\text{Protein abundance (\%)} = \frac{\text{CRY2}_{\text{mt}}/\text{ACTIN}_{\text{mt}} - \text{CRY2}_{cry1cry2}/\text{ACTIN}_{cry1cry2}}{\text{CRY2}_{\text{GFP-CRY2 L5}}/\text{ACTIN}_{\text{GFP-CRY2 L5}} - \text{CRY2}_{cry1cry2}/\text{ACTIN}_{cry1cry2}} \quad (5)$$

For the 20 CRY2 mutants that are transcribed but fail to accumulate the mutant protein in plants (Fig. 2-2C and Fig. S2-3, A-C), the results were verified in at least 6 independent transgenic lines except GFP-CRY2^{D112A} (D112A), GFP-CRY2^{F253A} (F253A), and GFP-CRY2^{W449A} (W449A). Expression of proteins was detectable in D112A, F253A and W449A lines with abundance lower than 5% of that of the L5 of wild-type GFP-CRY2 line, and were thus categorized into the "lack of protein" group.

Blue-light induced proteolysis curves were plotted to degradation curve:

$$Y = 100 \times e^{-kt} \quad (6)$$

where Y (%) is percentage of initial signal, *e* is Euler's number, k (%/min) is rate of degradation, t (min) is independent variable time. Half-life ($^{60}t_{1/2}$)was calculated by the formula:

$$^{60}t_{1/2} = ln2/k \quad (7)$$

The blue-light dependent proteolysis activities were calculated by the following formula:

$$Proteolysis\ activity\ (\%) = \frac{1/t_{1/2\ mutant} - 1/120}{1/t_{1/2\ GFP-CRY2} - 1/120} \quad (8)$$

2.5.5 Human cell culture, transfection, protein expression and co-immunoprecipitation assay

Human embryonic kidney (HEK) 293T cells were routinely cultured in Dulbecco's modified Eagle's medium (DMEM, 10-013-CM, Corning, New York, NY, USA) supplemented with 10% (v/v) FBS, 100 IU penicillin and 100mg/L streptomycin, in humidified 5% (v/v) CO_2 incubator at 37°C.

For protein expression assays, HEK293 were seeded at a density of ~3×10^5 cells per well of a 6-well plate and transfected using a calcium phosphate precipitation protocol as previously described [33]. 2.5 μg of wild-type or mutant P_{CMV}::FLAG-EGFP-hCRY1::T_{SV40}, or P_{CMV}::MYC-hCRY1::$T_{β\text{-}globin}$ plasmids were co-transfected with 2 μg of P_{CMV}::FLAG-EGFP::T_{SV40} plasmids (control. Cells were harvested 36~48 hours after transfection and lysed in 3 volumes of 1% Brij buffer (1% Brij-35, 50 mM Tris-HCl pH 8.0, 150 mM NaCl, 1mM EDTA and 1x Protease inhibitor cocktail). The cells were kept on ice for 30 min, followed by centrifugation at 14,000g for 10 min at 4°C. The supernatant was mixed with equal volume of 2 x SDS-PAGE Sample Buffer and heated at 100°C for 3 min. The protein samples were separated by 10% SDS-PAGE and analyzed by immunoblot using the Odyssey CLx Imaging System (LI-COR) as described above. The primary antibodies used in this assay were rabbit-anti-FLAG (1:3000, F7425, Sigma- Aldrich Corp., St. Louis, MO, USA). The secondary antibodies were as listed above. The expression levels of wild-type and mutant hCRY1 were

normalized against the expression of GFP, and converted to relative expression units (REU) by dividing with the mean (n=3) of wild type hCRY1 expression level.

Co-immunoprecipitation was conducted as described in [33]. Basically, 36 hours after transfection, the HEK293T cells expressing the indicated proteins were treated with blue light (100 μmol m^{-2} s^{-1}) or darkness for 120 minutes. After light treatment, the cells were harvested and washed with PBS, and then lysed with 1% Brij buffer. FLAG-affinity beads (F2426, Sigma) were added to cell lysate, and incubated with gentle rocking at 4˚C in darkness for 2 hours. After incubation, the beads were washed with 1% Brij buffer for three times. The proteins were eluted by competition with 30 μl of 500 μg/ml of 3xFLAG peptide with shaking of 1400 rpm at 4˚C for 30 minutes. The eluted proteins were then analyzed by fluorescent immunoblot.

2.5.6 Fluorescence microscopy

Seeds were sterilized and sown onto fresh-made MS Agar (0.8%) plates, subjected to 4˚C cold treatment in darkness for 1 day, and then put into the long-day (16 h day/ 8 h night) period of white light at room temperature for 2 days. The resulting seedlings were directly imaged using a Zeiss LSM 700 confocal microscope.

2.6 References

1. Ahmad, M., and Cashmore, A.R. (1993). HY4 gene of A. thaliana encodes a protein with characteristics of a blue-light photoreceptor. Nature 366, 162–166.
2. Lin, C., Robertson, D.E., Ahmad, M., Raibekas, A.A., Jorns, M.S., Dutton, P.L., and Cashmore, A.R. (1995). Association of flavin adenine dinucleotide with the Arabidopsis blue light receptor CRY1. Science 269, 968–970.
3. Cashmore, A.R. (2003). Cryptochromes: Enabling plants and animals to determine circadian time. Cell 114, 537–543.
4. Sancar, A. (2003). Structure and Function of DNA Photolyase and Cryptochrome Blue-Light Photoreceptors. Chem. Rev. 103, 2203–2238.
5. Lin, C., and Shalitin, D. (2003). Cryptochrome structure and signal transduction. Annu. Rev. Plant Biol. 54, 469–496.
6. Cashmore, A.R., Jarillo, J.A., Wu, Y.J., and Liu, D. (1999). Cryptochromes: blue light receptors for plants and animals. Science 284, 760–5.
7. Mirny, L.A., and Shakhnovich, E.I. (1999). Universally conserved positions in protein folds: reading evolutionary signals about stability, folding kinetics and function. J. Mol. Biol. 291, 177–196.
8. Valencia, A., Chardin, P., Wittinghofer, A., and Sander, C. (1991). The ras protein family: evolutionary tree and role of conserved amino acids. Biochemistry 30, 4637–4648.
9. Landau, M., Mayrose, I., Rosenberg, Y., Glaser, F., Martz, E., Pupko, T., and Ben-Tal, N. (2005). ConSurf 2005: the projection of evolutionary conservation scores of residues on protein structures. Nucleic Acids Res. 33, W299–W302.

10. Gu, N.-N., Zhang, Y.-C., and Yang, H.-Q. (2012). Substitution of a conserved glycine in the PHR domain of Arabidopsis CRYPTOCHROME 1 confers a constitutive light response. Mol. Plant 5, 85–97.

11. Guo, H., Yang, H., Mockler, T.C., and Lin, C. (1998). Regulation of flowering time by Arabidopsis photoreceptors. Science 279, 1360–1363.

12. Li, X., Wang, Q., Yu, X., Liu, H., Yang, H., Zhao, C., Liu, X., Tan, C., Klejnot, J., Zhong, D., and Lin, C. (2011). Arabidopsis cryptochrome 2 (CRY2) functions by the photoactivation mechanism distinct from the tryptophan (trp) triad-dependent photoreduction. Proc. Natl. Acad. Sci. U. S. A. 108, 20844–9.

13. Gao, J., Wang, X., Zhang, M., Bian, M., Deng, W., Zuo, Z., Yang, Z., Zhong, D., and Lin, C. (2015). Trp triad-dependent rapid photoreduction is not required for the function of Arabidopsis CRY1. Proc. Natl. Acad. Sci. U. S. A. 112, 9135–9140.

14. Taslimi, A., Zoltowski, B., Miranda, J.G., Pathak, G.P., Hughes, R.M., and Tucker, C.L. (2016). Optimized second-generation CRY2–CIB dimerizers and photoactivatable Cre recombinase. Nat. Chem. Biol. 12, 425–430.

15. Stanewsky, R., Kaneko, M., Emery, P., Beretta, B., Wager-Smith, K., Kay, S.A., Rosbash, M., and Hall, J.C. (1998). The cryb mutation identifies cryptochrome as a circadian photoreceptor in Drosophila. Cell 95, 681–692.

16. Rosensweig, C., Reynolds, K.A., Gao, P., Laothamatas, I., Shan, Y., Ranganathan, R., Takahashi, J.S., and Green, C.B. (2018). An evolutionary hotspot defines functional differences between CRYPTOCHROMES. Nat. Commun. 9, 1138.

17. McCarthy, E. V, Baggs, J.E., Geskes, J.M., Hogenesch, J.B., and Green, C.B.

(2009). Generation of a novel allelic series of cryptochrome mutants via mutagenesis reveals residues involved in protein-protein interaction and CRY2-specific repression. Mol. Cell. Biol. 29, 5465–76.

18. Ode, K.L., Ukai, H., Susaki, E.A., Narumi, R., Matsumoto, K., Hara, J., Koide, N., Abe, T., Kanemaki, M.T., Kiyonari, H., and Ueda, H.R. (2017). Knockout-Rescue Embryonic Stem Cell-Derived Mouse Reveals Circadian-Period Control by Quality and Quantity of CRY1. Mol. Cell 65, 176–190.
19. Aubert, C., Vos, M.H., Mathis, P., Eker, a P., and Brettel, K. (2000). Intraprotein radical transfer during photoactivation of DNA photolyase. Nature 405, 586–590.
20. Li, Y.F., Heelis, P.F., and Sancar, A. (1991). Active site of DNA photolyase: tryptophan-306 is the intrinsic hydrogen atom donor essential for flavin radical photoreduction and DNA repair in vitro. Biochemistry 30, 6322–6329.
21. Chaves, I., Pokorny, R., Byrdin, M., Hoang, N., Ritz, T., Brettel, K., Essen, L.-O., van der Horst, G.T.J., Batschauer, A., and Ahmad, M. (2011). The Cryptochromes: Blue Light Photoreceptors in Plants and Animals. Annu. Rev. Plant Biol. 62, 335–364.
22. Zeugner, A., Byrdin, M., Bouly, J.-P., Bakrim, N., Giovani, B., Brettel, K., and Ahmad, M. (2005). Light-induced electron transfer in Arabidopsis cryptochrome-1 correlates with in vivo function. J. Biol. Chem. 280, 19437–40.
23. Solov'yov, I.A., Domratcheva, T., Moughal Shahi, A.R., and Schulten, K. (2012). Decrypting cryptochrome: revealing the molecular identity of the photoactivation reaction. J. Am. Chem. Soc. 134, 18046–18052.
24. Engelhard, C., Wang, X., Robles, D., Moldt, J., Essen, L.-O., Batschauer, A., Bittl,

R., and Ahmad, M. (2014). Cellular metabolites enhance the light sensitivity of Arabidopsis cryptochrome through alternate electron transfer pathways. Plant Cell 26, 4519–31.

25. Ahmad, M. (2016). Photocycle and signaling mechanisms of plant cryptochromes. Curr. Opin. Plant Biol. 33, 108–115.
26. Banerjee, R., Schleicher, E., Meier, S., Viana, R.M., Pokorny, R., Ahmad, M., Bittl, R., and Batschauer, A. (2007). The signaling state of Arabidopsis cryptochrome 2 contains flavin semiquinone. J. Biol. Chem. 282, 14916–22.
27. Langenbacher, T., Immeln, D., Dick, B., and Kottke, T. (2009). Microsecond light-induced proton transfer to flavin in the blue light sensor plant cryptochrome. J. Am. Chem. Soc. 131, 14274–14280.
28. Müller, M., and Carell, T. (2009). Structural biology of DNA photolyases and cryptochromes. Curr. Opin. Struct. Biol. 19, 277–285.
29. Müller, P., Bouly, J.-P., Hitomi, K., Balland, V., Getzoff, E.D., Ritz, T., and Brettel, K. (2015). ATP Binding Turns Plant Cryptochrome Into an Efficient Natural Photoswitch. Sci. Rep. 4, 5175.
30. Gegear, R.J., Foley, L.E., Casselman, A., and Reppert, S.M. (2010). Animal cryptochromes mediate magnetoreception by an unconventional photochemical mechanism. Nature 463, 804–807.
31. Lin, C., Top, D., Manahan, C.C., Young, M.W., and Crane, B.R. (2018). Circadian clock activity of cryptochrome relies on tryptophan-mediated photoreduction. Proc. Natl. Acad. Sci. U. S. A. 115, 3822–3827.
32. Zhu, B., Cai, G., Hall, E.O., and Freeman, G.J. (2007). In-Fusion™ assembly:

seamless engineering of multidomain fusion proteins, modular vectors, and mutations. Biotechniques 43, 354–359.

33. Wang, Q., Zuo, Z., Wang, X., Gu, L., Yoshizumi, T., Yang, Z., Yang, L., Liu, Q., Liu, W., Han, Y.-J., Kim, J.-I., Liu, B., Wohlschlegel, J.A., Matsui, M., Oka, Y., and Lin, C. (2016). Photoactivation and inactivation of Arabidopsis cryptochrome 2. Science 354, 343–347.

34. Liu, Q., Wang, Q., Deng, W., Wang, X., Piao, M., Cai, D., Li, Y., Barshop, W.D., Yu, X., Zhou, T., Liu, B., Oka, Y., Wohlschlegel, J., Zuo, Z., and Lin, C. (2017). Molecular basis for blue light-dependent phosphorylation of Arabidopsis cryptochrome 2. Nat. Commun. 8, 15234.

35. Yu, X., Sayegh, R., Maymon, M., Warpeha, K., Klejnot, J., Yang, H., Huang, J., Lee, J., Kaufman, L., and Lin, C. (2009). Formation of nuclear bodies of Arabidopsis CRY2 in response to blue light is associated with its blue light-dependent degradation. Plant Cell 21, 118–130.

36. Zuo, Z.C., Meng, Y.Y., Yu, X.H., Zhang, Z.L., Feng, D.S., Sun, S.F., Liu, B., and Lin, C.T. (2012). A study of the blue-light-dependent phosphorylation, degradation, and photobody formation of Arabidopsis CRY2. Mol. Plant 5, 726–733.

37. Yu, X., Klejnot, J., Zhao, X., Shalitin, D., Maymon, M., Yang, H., Lee, J., Liu, X., Lopez, J., and Lin, C. (2007). Arabidopsis cryptochrome 2 completes its posttranslational life cycle in the nucleus. Plant Cell 19, 3146–56.

38. Shalitin D; Yang H; Mockler TC; Maymon M; Guo H; Whitelam GC; Lin C (2002). Regulation of Arabidopsis cryptochrome 2 by blue-light- dependent

phosphorylation. Nature 417, 763–767.

39. Liu, H., Yu, X., Li, K., Klejnot, J., Yang, H., Lisiero, D., and Lin, C. (2008). Photoexcited CRY2 interacts with CIB1 to regulate transcription and floral initiation in *Arabidopsis*. Science 322, 1535–1539.

40. Mockler, T.C., Guo, H., Yang, H., Duong, H., and Lin, C. (1999). Antagonistic actions of Arabidopsis cryptochromes and phytochrome B in the regulation of floral induction. Development 126, 2073–2082.

41. Zuo, Z., Liu, H., Liu, B., Liu, X., and Lin, C. (2011). Blue Light-Dependent Interaction of CRY2 with SPA1 Regulates COP1 activity and Floral Initiation in Arabidopsis. Curr. Biol. 21, 841–847.

42. Valverde, F., Mouradov, A., Soppe, W., Ravenscroft, D., Samach, A., and Coupland, G. (2004). Photoreceptor regulation of CONSTANS protein in photoperiodic flowering. Science 303, 1003–1006.

43. Notredame, C., Higgins, D.G., and Heringa, J. (2000). T-coffee: a novel method for fast and accurate multiple sequence alignment. J. Mol. Biol. 302, 205–217.

44. Robert, X., and Gouet, P. (2014). Deciphering key features in protein structures with the new ENDscript server. Nucleic Acids Res. 42, W320–W324.

45. Waterhouse, A., Bertoni, M., Bienert, S., Studer, G., Tauriello, G., Gumienny, R., Heer, F.T., de Beer, T.A.P., Rempfer, C., Bordoli, L., Lepore, R., and Schwede, T. (2018). SWISS-MODEL: homology modelling of protein structures and complexes. Nucleic Acids Res. 46, W296–W303.

46. Brautigam, C.A., Smith, B.S., Ma, Z., Palnitkar, M., Tomchick, D.R., Machius, M., and Deisenhofer, J. (2004). Structure of the photolyase-like domain of

cryptochrome 1 from Arabidopsis thaliana. Proc. Natl. Acad. Sci. 101, 12142–12147.

47. Peragine, A., Yoshikawa, M., Wu, G., Albrecht, H.L., and Poethig, R.S. (2004). SGS3 and SGS2/SDE1/RDR6 are required for juvenile development and the production of trans-acting siRNAs in Arabidopsis. Genes Dev. 18, 2368–79.

48. Clough, S.J. (2005). Floral dip: Agrobacterium-mediated germ line transformation. Transgenic Plants, 091–102.

49. Wang, X., Wang, Q., Han, Y.-J., Liu, Q., Gu, L., Yang, Z., Su, J., Liu, B., Zuo, Z., He, W., Wang, J., Liu, B., Matsui, M., Kim, J.-I., Oka, Y., and Lin, C. (2017). A CRY-BIC negative-feedback circuitry regulating blue light sensitivity of Arabidopsis. Plant J. 92, 426–436.

50. Neff, M.M., and Chory, J. (1998). Genetic interactions between phytochrome A, phytochrome B, and cryptochrome 1 during Arabidopsis development. Plant Physiol. 118, 27–35.

51. Platten, J.D., Foo, E., Elliott, R.C., Hecht, V., Reid, J.B., and Weller, J.L. (2005). Cryptochrome 1 contributes to blue-light sensing in pea. Plant Physiol. 139, 1472–82.

52. Lamia, K.A., Sachdeva, U.M., DiTacchio, L., Williams, E.C., Alvarez, J.G., Egan, D.F., Vasquez, D.S., Juguilon, H., Panda, S., Shaw, R.J., Thompson, C.B., and Evans, R.M. (2009). AMPK regulates the circadian clock by cryptochrome phosphorylation and degradation. Science 326, 437–40.

53. Ruckle, M.E., DeMarco, S.M., and Larkin, R.M. (2007). Plastid signals remodel light signaling networks and are essential for efficient chloroplast biogenesis in

Arabidopsis. Plant Cell 19, 3944–60.

54. Czarna, A., Berndt, A., Singh, H.R., Grudziecki, A., Ladurner, A.G., Timinszky, G., Kramer, A., and Wolf, E. (2013). Structures of Drosophila Cryptochrome and Mouse Cryptochrome1 Provide Insight into Circadian Function. Cell 153, 1394–1405.

55. Hitomi, K., DiTacchio, L., Arvai, A.S., Yamamoto, J., Kim, S.-T., Todo, T., Tainer, J.A., Iwai, S., Panda, S., and Getzoff, E.D. (2009). Functional motifs in the (6-4) photolyase crystal structure make a comparative framework for DNA repair photolyases and clock cryptochromes. Proc. Natl. Acad. Sci. U. S. A. 106, 6962–7.

56. Shalitin, D., Yu, X., Maymon, M., Mockler, T., and Lin, C. (2003). Blue light-dependent in vivo and in vitro phosphorylation of Arabidopsis cryptochrome 1. Plant Cell 15, 2421–2429.

57. Ahmad, M., Lin, C.T., and Cashmore, A.R. (1995). Mutations Throughout an Arabidopsis Blue-Light Photoreceptor Impair Blue-Light-Responsive Anthocyanin Accumulation and Inhibition of Hypocotyl Elongation. Plant J. 8, 653–658.

58. Ahmad, M., Jarillo, J.A., Smirnova, O., and Cashmore, A.R. (1998). The CRY1 Blue Light Photoreceptor of Arabidopsis Interacts with Phytochrome A In Vitro. Mol. Cell 1, 939–948.

59. Botto, J.F., Alonso-Blanco, C., Garzarón, I., Sánchez, R.A., and Casal, J.J. (2003). The Cape Verde Islands allele of cryptochrome 2 enhances cotyledon unfolding in the absence of blue light in Arabidopsis. Plant Physiol. 133, 1547–56.

60. El-Din El-Assal, S., Alonso-Blanco, C., Peeters, A.J.M., Raz, V., and Koornneef,

M. (2001). A QTL for flowering time in Arabidopsis reveals a novel allele of CRY2. Nat. Genet. 29, 435–440.

61. Eckel, M., Steinchen, W., and Batschauer, A. (2018). ATP boosts lit state formation and activity of Arabidopsis cryptochrome 2. Plant J. 96, 389–403.

62. Exner, V., Alexandre, C., Rosenfeldt, G., Alfarano, P., Nater, M., Caflisch, A., Gruissem, W., Batschauer, A., and Hennig, L. (2010). A gain-of-function mutation of Arabidopsis cryptochrome1 promotes flowering. Plant Physiol. 154, 1633–45.

63. Taslimi, A., Vrana, J.D., Chen, D., Borinskaya, S., Mayer, B.J., Kennedy, M.J., and Tucker, C.L. (2014). An optimized optogenetic clustering tool for probing protein interaction and function. Nat. Commun. 5, 4925.

2.7 Figures

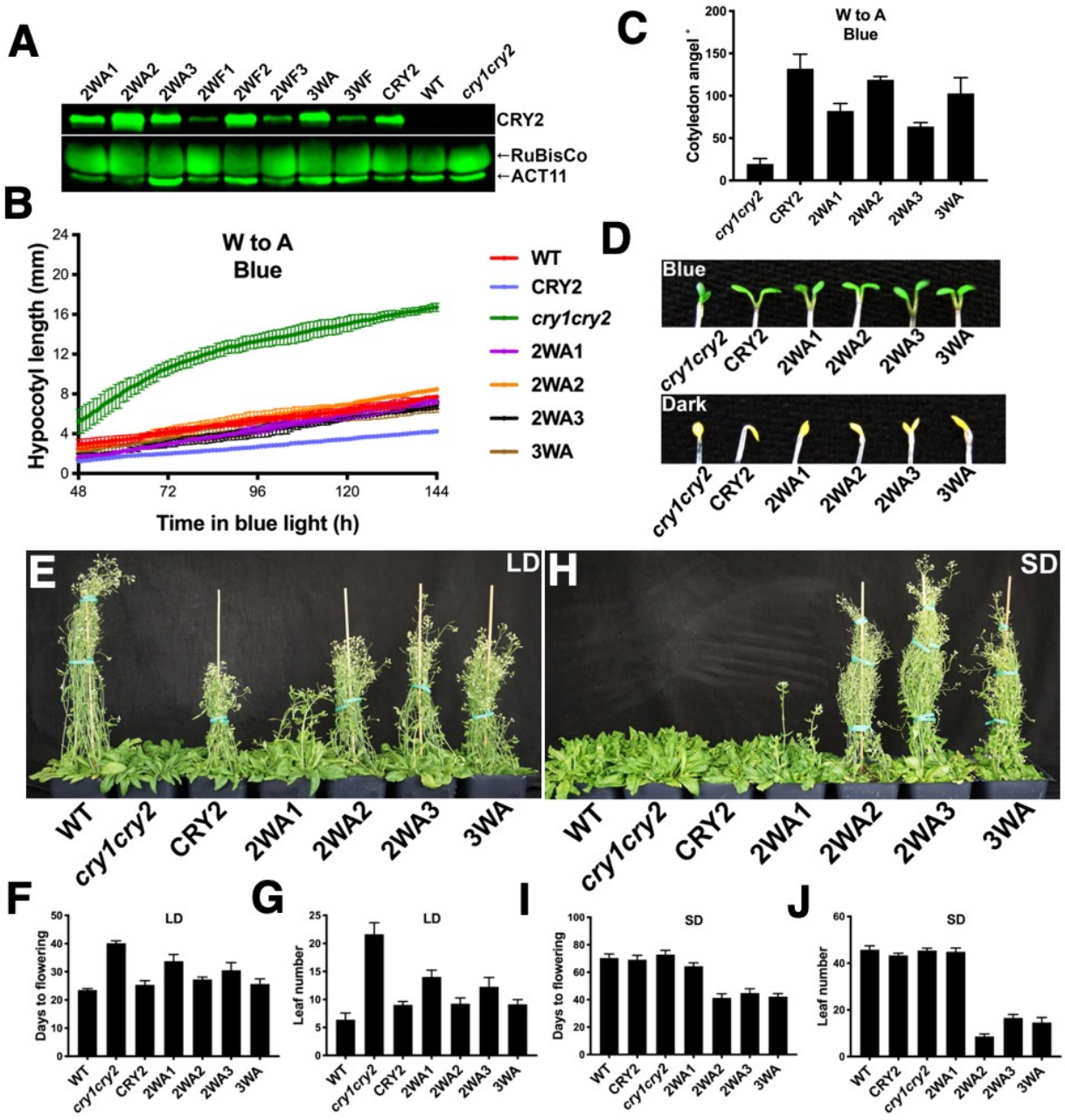

Fig. 2-1. Analyses of double and triple mutants of the Trp-triad residues of CRY2.

A) Immunoblots showing expression of double and triple Trp-triad mutants of the GFP-CRY2 fusion protein.

B) Kinetics analysis of elongation of seedlings germinated and grown under the blue light (15 μmol m^{-2} s^{-1}). Seedlings are imaged 48 hours after germination at the frequency of 1 image per hour for another 96 hours (n=3).

C) Angles between the two cotyledons were measured from the images taken at 114 hours after germination in B (n=3).

D) The cotyledon unfolding phenotype of 6-day-old seedlings grown in blue light (20 μmol m^{-2} s^{-1})(upper) or darkness (lower).

E-J) Images of 40 (E)- or 60 (H)-day-old plants grown in LD (16 h day/ 8 h night) or SD (8 h day/16 h night). Days to flowering (F, I) and rosette leaf number (G, J) at flowering are shown (n≥8). The wild-type (WT) and transgenic plants constitutively expressing the "wild-type" GFP-CRY2 or the double (2WA1, 2WA2, 2WA3) or triple (3WA) mutants of GFP-CRY2 fusion proteins in the *cry1cry2* mutant background are indicated. See Table 2-1 for more detailed information. Bars in B, C, F, G, I and indicates SD of the mean.

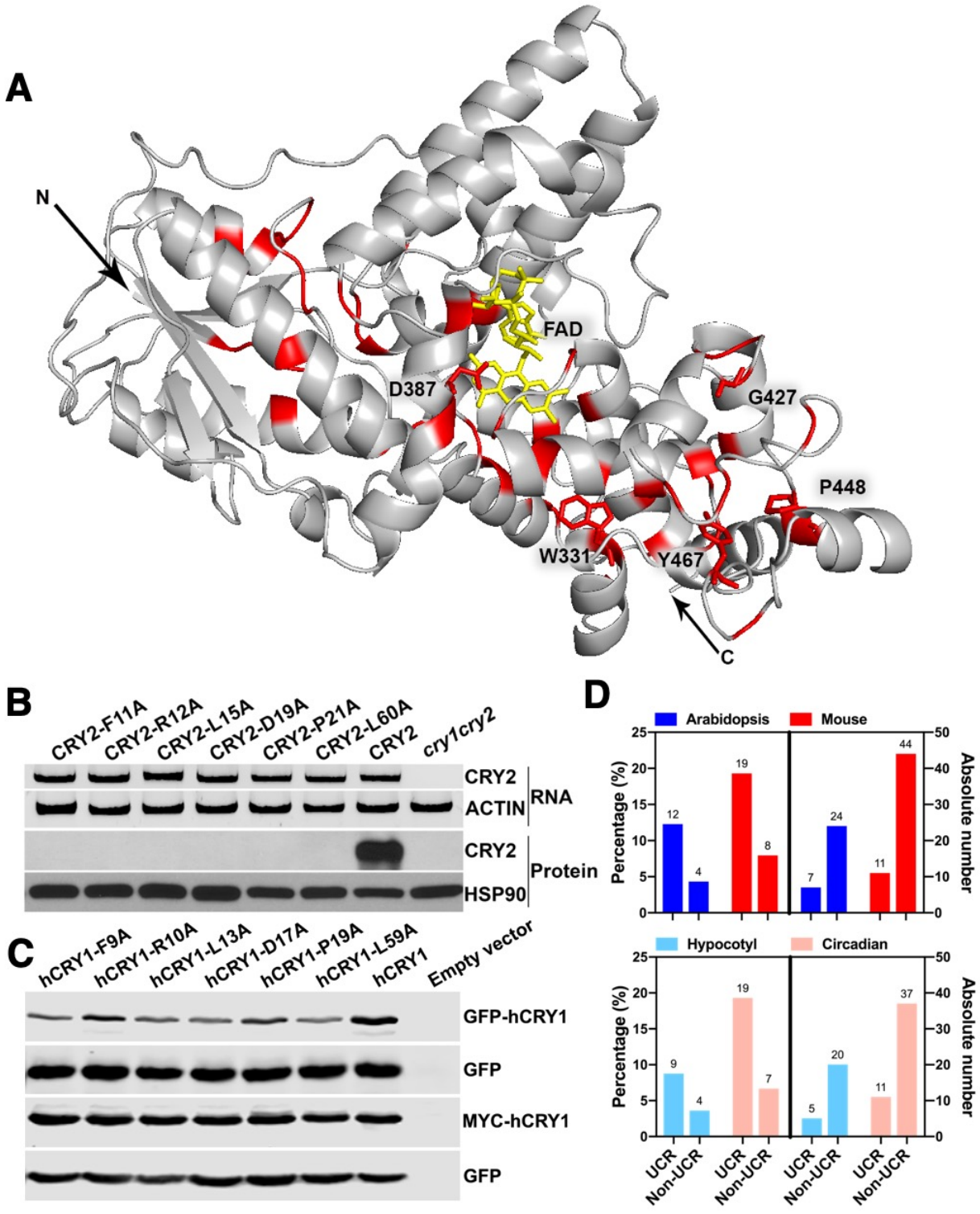
A
N
FAD
D387
G427
P448
W331
Y467
C
B
CRY2-F11A
CRY2-R12A
CRY2-L15A
CRY2-D19A
CRY2-P21A
CRY2-L60A
CRY2
cry1cry2
CRY2
ACTIN
RNA
CRY2
HSP90
Protein
C
hCRY1-F9A
hCRY1-R10A
hCRY1-L13A
hCRY1-D17A
hCRY1-P19A
hCRY1-L59A
hCRY1
Empty vector
GFP-hCRY1
GFP
MYC-hCRY1
GFP
D
Arabidopsis
Mouse
Hypocotyl
Circadian
Percentage (%)
Absolute number
UCR
Non-UCR

Fig. 2-2: Analyses of UCRs and non-UCRs of plant and mammalian cryptochromes.

A) Structure model of the PHR domain of *Arabidopsis* CRY2. FAD (yellow), UCRs (red), the N-terminus, the C-terminus (arrows), and 5 UCRs discussed in the text are indicated.

B) RT-PCR (upper two panels) and immunoblots (lower two panels) showing representative *Arabidopsis* CRY2 UCR mutants that stably express mRNA but not recombinant protein in transgenic plants of the indicated constructs.

C) Immunoblot showing stable protein expression of hCRY1 UCR mutant proteins, each being altered at the residue equivalent to the corresponding *Arabidopsis* CRY2 UCR mutants shown in B. Samples were prepared from whole cell lysates of HEK293T cells co-transfected by two plasmids: the sample plasmid encoding the indicated GFP-hCRY1 (upper two panels) or MYC-hCRY1 (lower two panels) protein mutated at the indicated UCR and the control plasmid encoding GFP as the transfection and immunoblot controls.

D) A comparison of the functions of UCR and non-UCR. The functionally defective Arabidopsis (blue) and mouse (red) cryptochromes previously examined based on protein functions (top, all functions, bottom, specific functions) but not sequence conservations were assigned to the UCR or Non-UCR groups according to Fig. S2-2. Right panels: absolute incidences of mutation independently reported. Left panels: percentage of the functionally defective mutations of the total number of the respective group of residues (UCR or non-UCRs). See Table 2-5 for details.

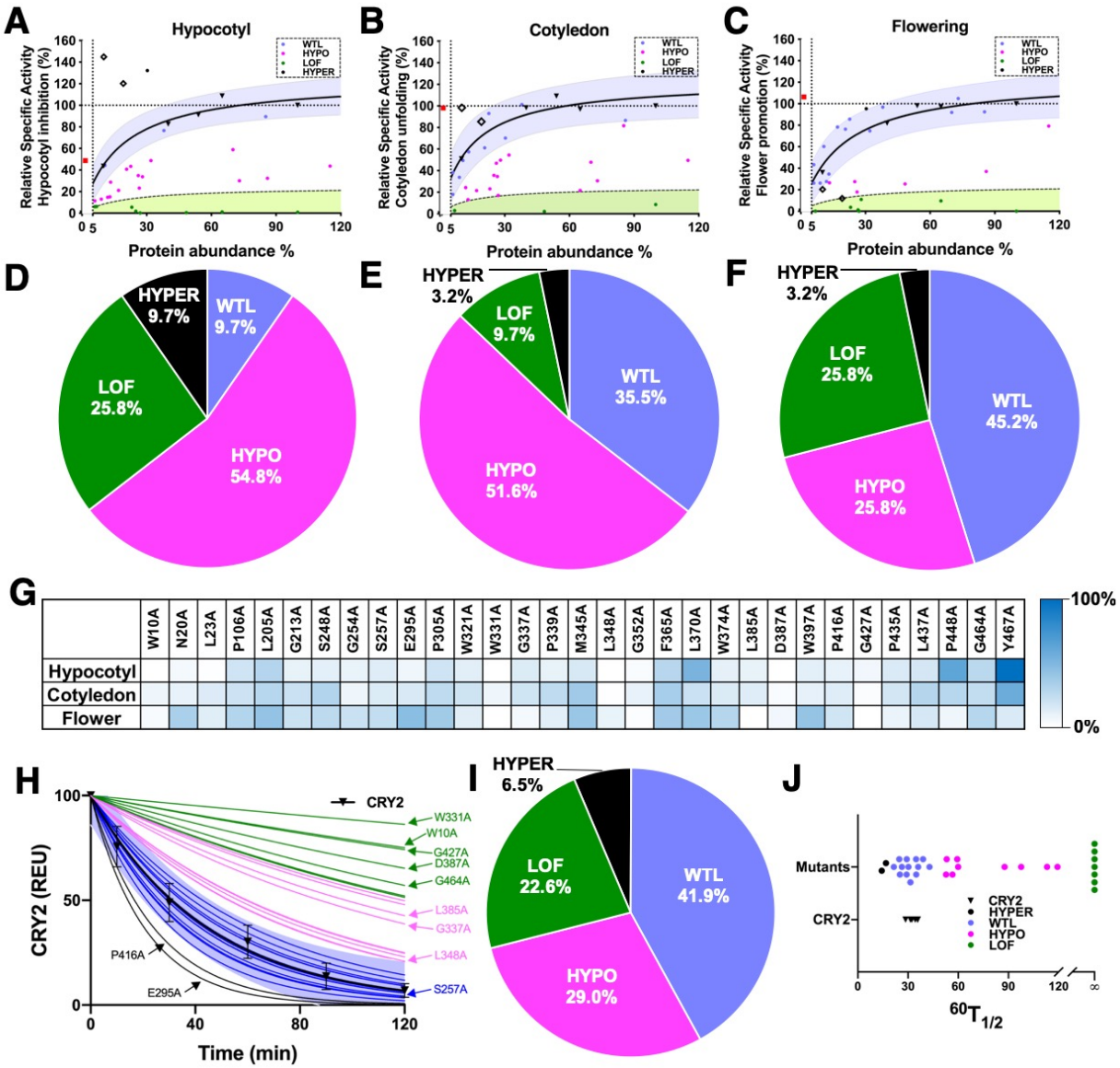

Fig. 2-3: Systematic analyses of the activities of the CRY2 UCR mutant proteins.

A-C) Using the standard curve approach to determine the relative specific photophysiological activities of the CRY2 UCR mutant. Blue shades represent regions within 95% prediction bands of the standard curves. Green shades indicate regions lower than 20% of the standard curve (black dashed lines). Round dots, red squires

and hollow squires indicate CRY2 UCR mutants, endogenous CRY2, $CRY2^{P448A}$ and $CRY2^{Y467A}$ (further analyzed in Fig. 4), respectively.

D-F) Classification of CRY2 UCR mutants for hypocotyl inhibition (D), and cotyledon unfolding (E), floral promotion (F), respectively.

G) Heat map showing normalized specific activities of all CRY2 UCR mutants.

H-J) The blue light-induced photobiochemical activity of the CRY2 UCR mutant proteins. 6-day-old etiolated seedlings were transferred to blue light (60 μmol m^{-2} s^{-1}) for the indicated time before sample collection for immunoblot analysis. The levels of CRY2 in different samples were quantified by fluorescent immunoblots (Odyssey® CLx Infrared Imaging System, LI-COR Inc, Lincoln, NE, USA), normalized to signals of the dark samples of the respective genotypes, and presented as REU (Relative Expression Unit). (H) Blue shades represent regions within 95% prediction bands of the standard proteolytic time-course of the "wild-type" GFP-CRY2. Color of thin lines match with color in *i*. SD of the mean (n≥6) are shown. (I) Classification of the photobiochemical activities of the CRY2 UCR mutant proteins. (J) Distribution of $^{60}T_{1/2}$.

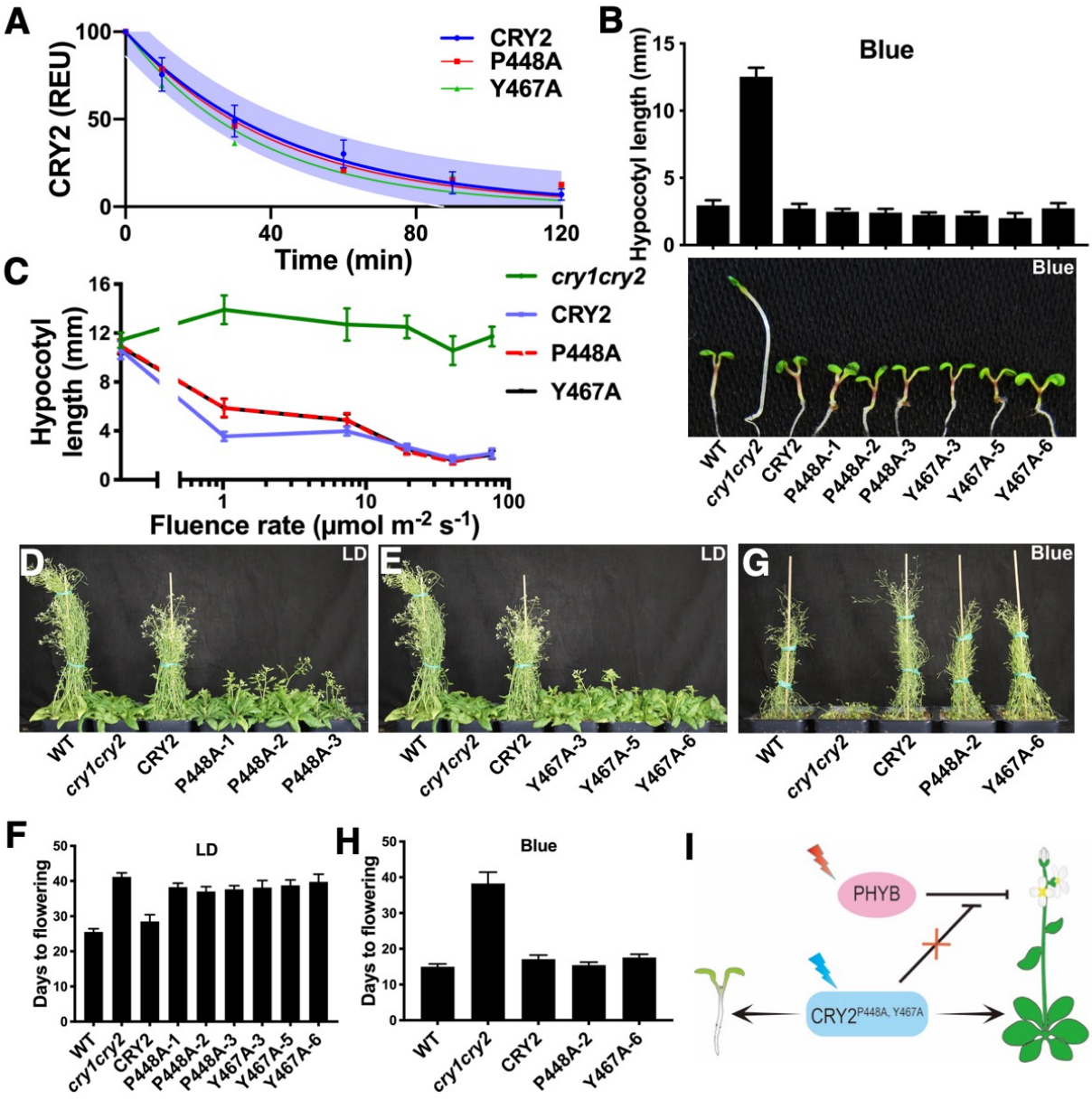

Fig. 2-4: Functional analyses of neighboring UCRs of CRY2.

A) Blue-light dependent proteolysis of the "wild-type" GFP-CRY2, GFP-CRY2^{P448A} and GFP-CRY2^{Y467A} UCR mutant proteins. The assay was carried out as described in Fig. 3H.

B) representative 6-day-old seedlings of the indicated genotypes grown in continuous blue light (20 μmol m^{-2} s^{-1}). Bar indicates SD of the mean (n≥20) are shown.

C) Hypocotyl lengths of 6-day-old seedlings grown in dark or continuous blue light with fluence rates of 1 to 80 μmol m^{-2} s^{-1}. Bar indicates SD of the mean (n≥20).

D-E) Plants of indicated genotypes grown in LD photoperiods (16 h day/ 8 h night) for 40 days.

F) Days to flowering of the indicated genotypes grown in LD. Bars indicates SD of the mean (n≥8).

G) Plants of indicated genotypes grown in continuous blue light (70-80 μmol m^{-2} s^{-1}) for 35 days.

H) Days to flowering of the indicated genotypes grown in LD. Bars indicates SD of the mean (n≥8).

I) A hypothetic model depicting CRY2$^{P448A, Y467}$-mediated regulation of de-etiolation and flowering time.

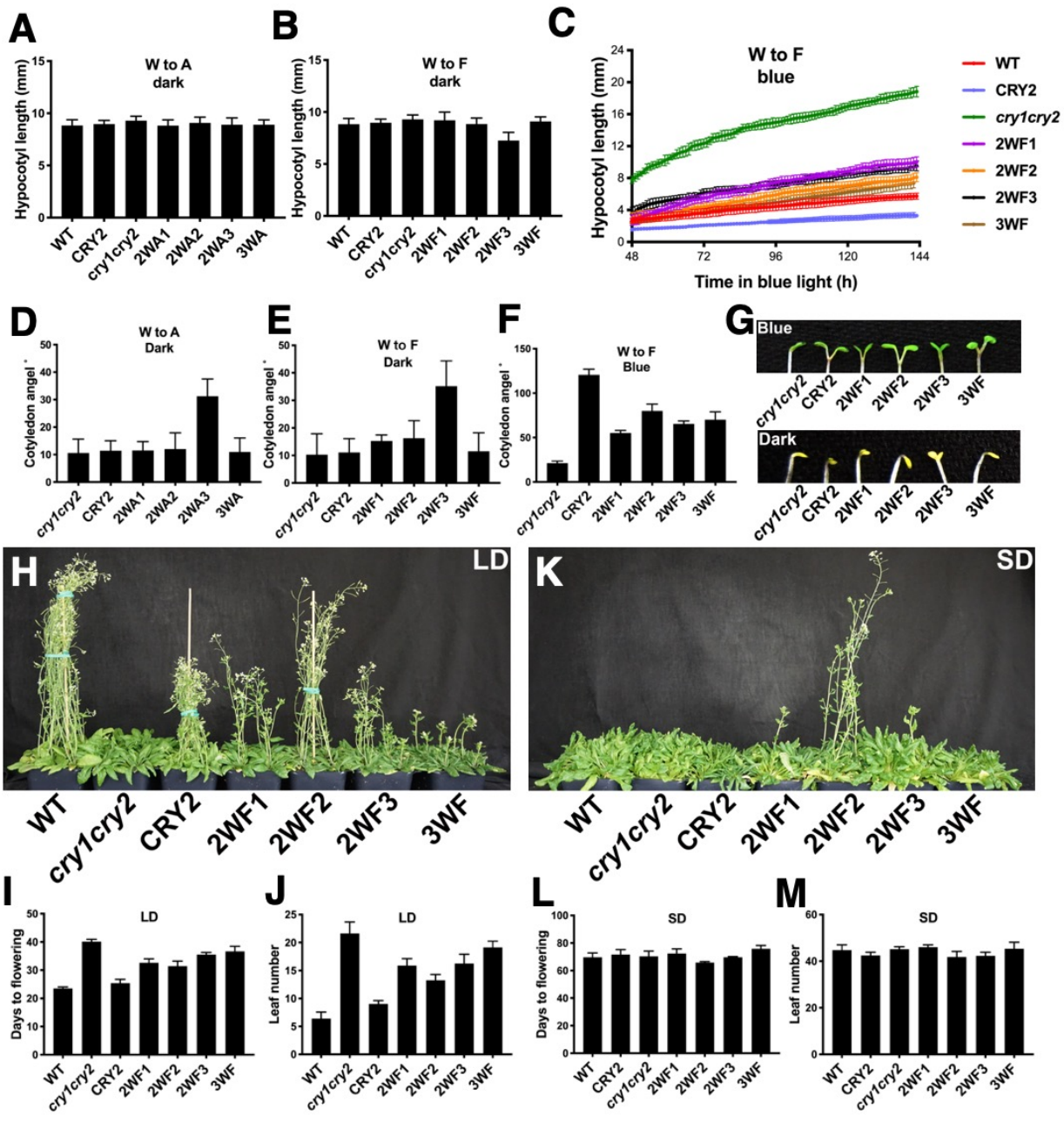

Fig. S2-1: The double and triple mutants of the universally conserved Trp-triad residues of CRY2 remained photobiologically active *in vivo.*

A-B) Hypocotyl length of six-day-old seedlings expressing the respective wild-type (GFP-CRY2) or double and triple mutants of the Trp-triad residues of CRY2, and

cry1cry2 mutant or wild-type (WT) seedlings grown in continuous blue light (20 μmol m^{-2} s^{-1}). The residue replacements involved in each mutant were listed in table S1. Bar indicates SD of the mean (n≥20).

C) The assay was carried out as described in Fig. 2-1B.

D-E) Cotyledon unfolding angels of six-day-old seedlings of the indicated genotypes grown in dark were manually measured using Fiji (NIH). Bar indicates SD of the mean (n≥20).

F) Angles between the two cotyledons, indicating the cotyledon unfolding phenotype, were measured from the images taken for Fig. S2-1C at 114 hours after germination. Bars indicates SD of the mean (n=3).

G) The cotyledon unfolding phenotype. Representative images of 6-day-old seedlings grown in blue light (20 μmol•m^{-2}•s^{-1}) (upper) or darkness (lower) are shown.

H-M) Images of 40 (H) or 60 (K) -day-old plants grown in in long-day (16 h day/ 8 h night) or short-day photoperiod (8 h day/ 16 h night). Days to flowering (I, L) and rosette leaf number (J, M) at flowering of the respective genotypes are shown. Bars indicates SD of the mean (n≥8). The wild-type (WT) and transgenic plants constitutively expressing the “wild-type” GFP-CRY2 (CRY2) fusion protein or three different double (2WF1, 2WF2, 2WF3) or triple (3WF) mutants of the Trp-triad residues of CRY2 as the GFP-CRY2 fusion proteins in the *cry1cry2* mutant background are indicated. See Table S2-1 for more detailed information.

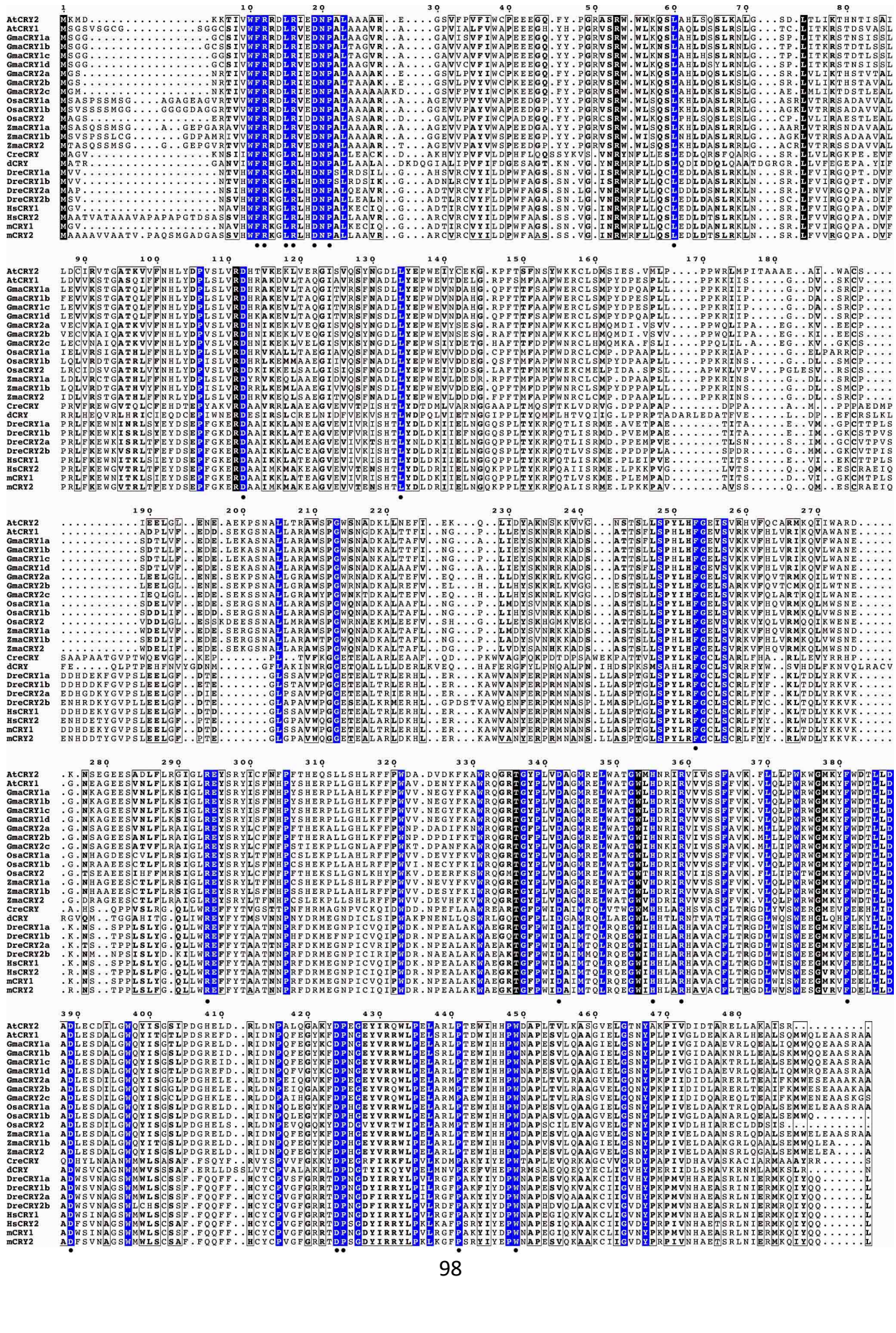

Fig. S2-2: Amino acid multiple sequence alignment of representative eukaryotic cryptochromes. The 57 universally conserved residues are labelled with blue (analyzed in the current study) or black background (not analyzed in the current study, excluding the start codon); the residues whose the alanine replacement were classified as “lack of protein” (Fig. 2-3 and Table 2-4) are labelled with round dots. NCBI accession number and species of proteins used in the alignment are listed in Table 2-2.

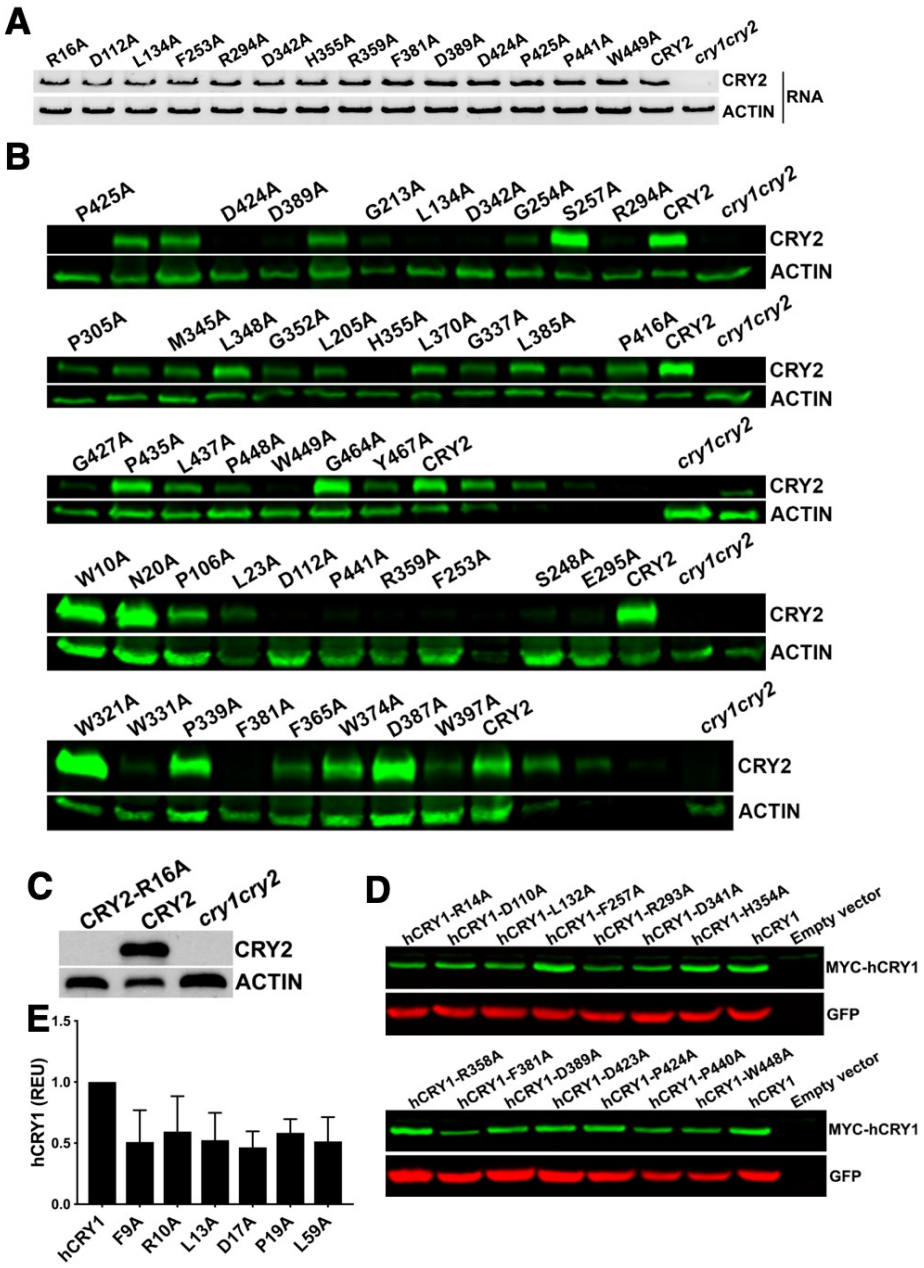

Fig. S2-3: RNA and protein expression of mutants of *Arabidopsis* CRY2 and hCRY1.

A) RT-PCR analyses of mRNA expression of wild-type or mutant of *GFP-CRY2* (upper) and *ACTIN* (lower) genes of seven-day-old seedlings grown in the long-day (16 h day/ 8 h night) period in the genotypes indicated.

B) Immunoblots showing wild-type or UCR mutant of GFP-CRY2 fusion proteins expressed in transgenic plants. Samples were extracted from seven-day-old seedlings grown in the long-day (16 h day/8 h night) period, fractionated in SDS/PAGE (10%), blotted, probed with anti-CRY2 antibody (CRY2) and anti-ACTIN antibody (ACTIN, control), then probed with fluorophore-conjugated secondary antibodies, and detected by the Odyssey CLx System (LI-COR).

C) Immunoblots showing wild-type or the $CRY2^{R16A}$ mutant of GFP-CRY2 fusion proteins expressed in transgenic plants. Samples were prepared, blotted, and probed primary antibodies as described in B, then probed with HRP (horseradish peroxidase)-conjugated secondary antibodies and detected by the ECL (Enhance Chemiluminescence) method.

D) Immunoblot showing stable protein expression of hCRY1 UCR mutant proteins, each being altered at the residue equivalent to the corresponding Arabidopsis CRY2 UCR “lack of protein) mutants. Samples were prepared from whole cell lysates of HEK293T cells co-transfected by two plasmids: the sample plasmid encoding the indicated MYC-hCRY1 protein mutated at the indicated UCR and the control plasmid encoding GFP as the transfection and immunoblot controls.

E) Quantification of protein expression of hCRY1 in HEK293T cells from Fig. 2-2C. GFP-hCRY1 Signals were acquired from fluorescent immunoblot (LI-COR) and quantified by an internal method of Image Studio Lite software (LI-COR), and normalized to GFP signals, and presented as hCRY1 relative expression unit (REU). Bars indicates SD of the mean (n=3).

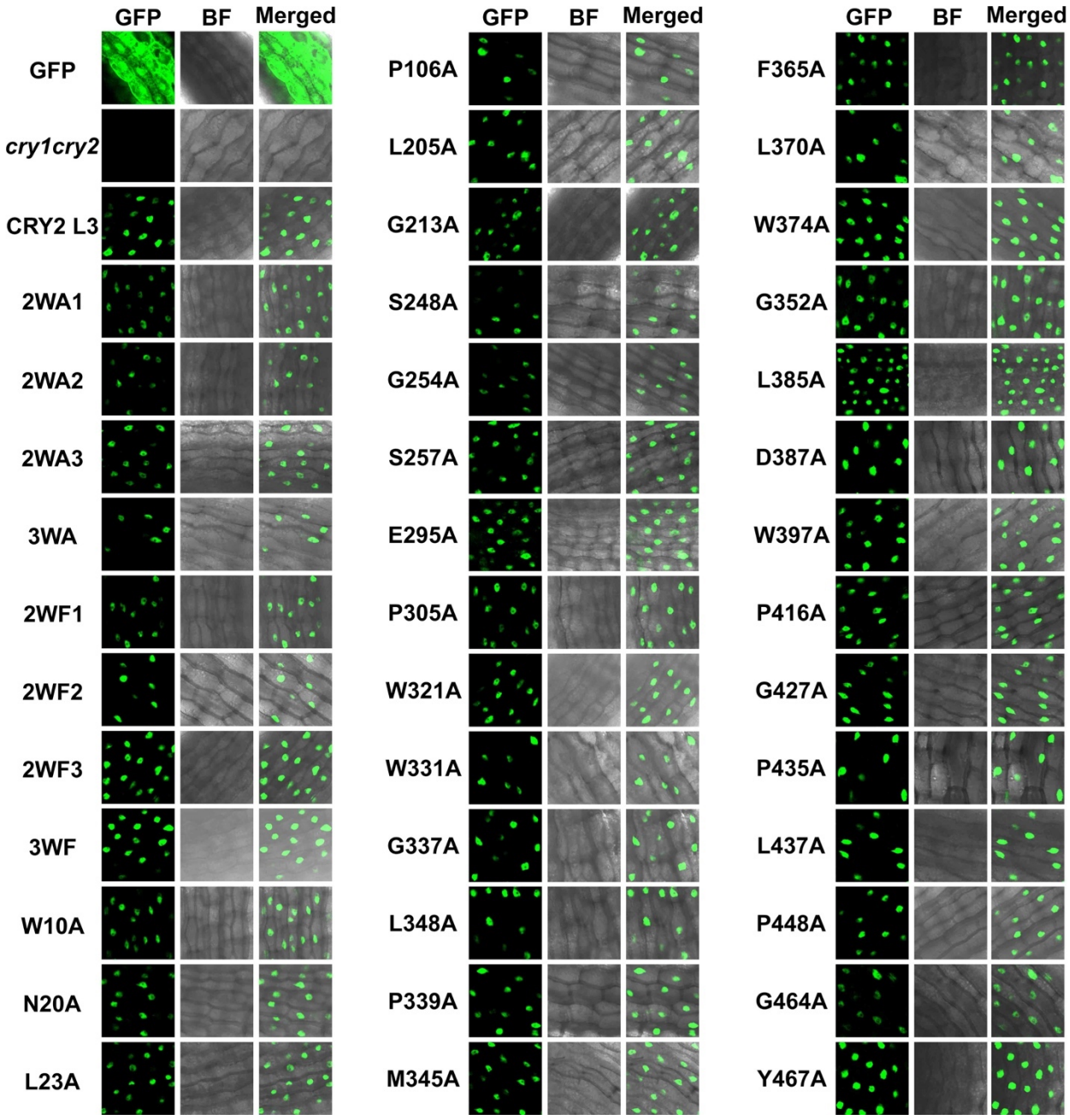

Fig. S2-4: Subcellular localization of the GFP-CRY2 mutants. Seedlings were grown on MS medium in the long-day (16 h day/ 8 h night) period of white light. Roots of two-day-old seedlings were directed analyzed by a Zeiss LSM700 confocal fluorescence microscope, and processed by the Fiji (NIH) software.

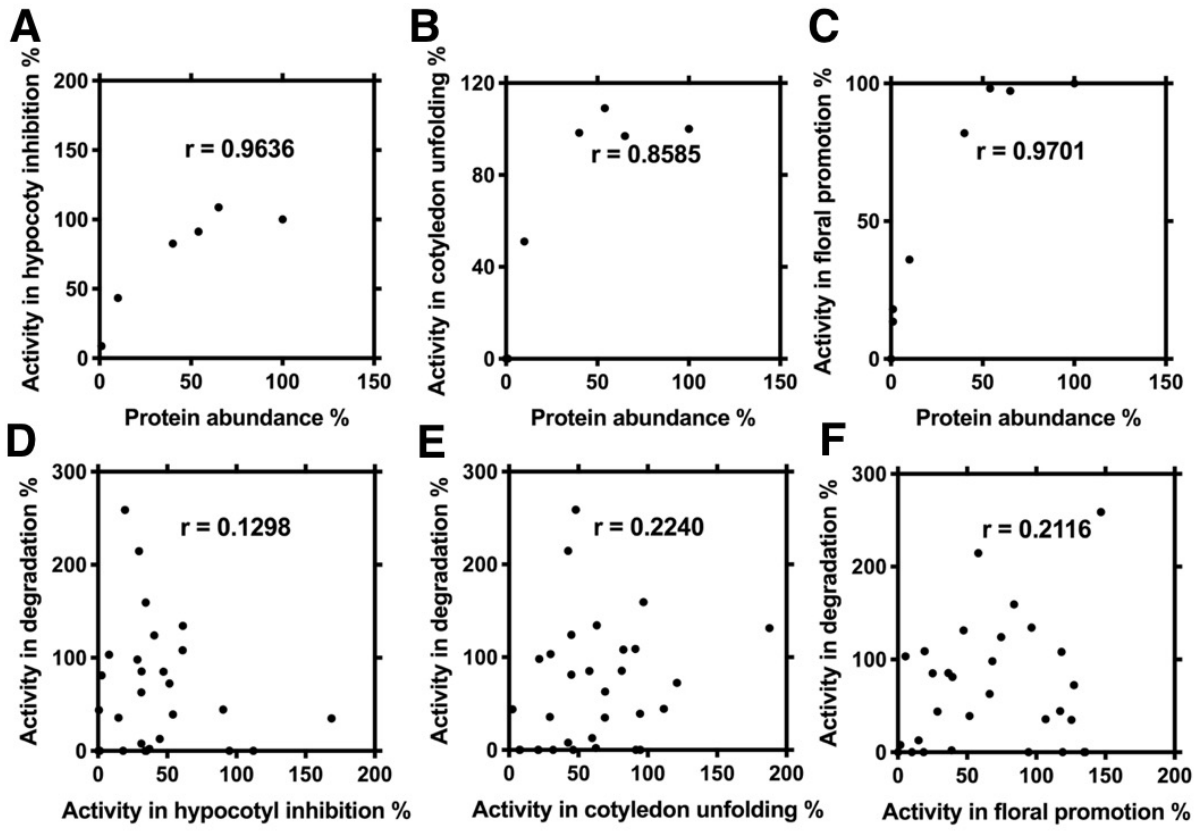

Fig. S2-5: Spearman's correlations.

A-C) Correlations between protein abundance and the relative specific-activities of the blue-light dependent inhibition of hypocotyl elongation (A), cotyledon unfolding (B) and floral promotion (C) for wild-type GFP-CRY2 fusion proteins.

D-F) Correlations between the blue-light dependent proteolysis activities (calculation detailed in SI Appendix, Materials and Methods) and the relative specific-activities of the blue-light dependent inhibition of hypocotyl elongation (D), cotyledon unfolding (E) and floral promotion (F) for transgenic plants of mutants of universally conserved GFP-CRY2 fusion proteins.

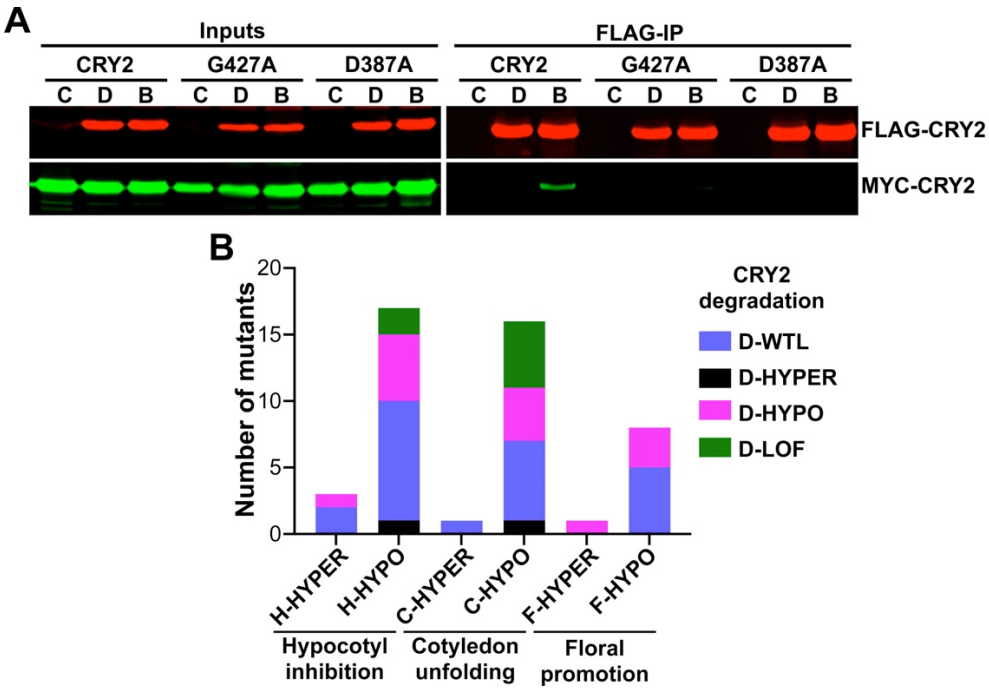

Fig. S2-6: Blue light-dependent homo-dimerization activity of CRY2^{G427A} and blue light-dependent degradation activity of CRY2 mutants examined.

A) HEK293T cells were co-transfected to express the indicated proteins, exposed to blue light (100 μmol m^{-1} s^{-1}) for 120 minutes, and immunoprecipitated by antibody to FLAG (FLAG-IP). The IP signal (FLAG-CRY2) or the co-IP signals (MYC-CRY2) were detected by immunoblots probed with antibodies to FLAG (upper panel) or to MYC (lower panel). The CRY2 null mutant D387A was included as a negative control. C: control (no FLAG-CRY2), D: Dark, B: blue light treated.

B) The blue light-induced proteolytic activities of the CRY2 mutants are classified as hyperactive (D-HYPER), hypoactive (D-HYPO), wild type-like (D-WTL), or loss-of-function (D-LOF), and the corresponding physiological activities of each class are labelled as hypermorphic or hypomorphic for hypocotyl inhibition (H-HYPER, H-HYPO), cotyledon unfolding (C-HYPER, C-HYPO) and floral promotion (F-HYPER, F-HYPO).

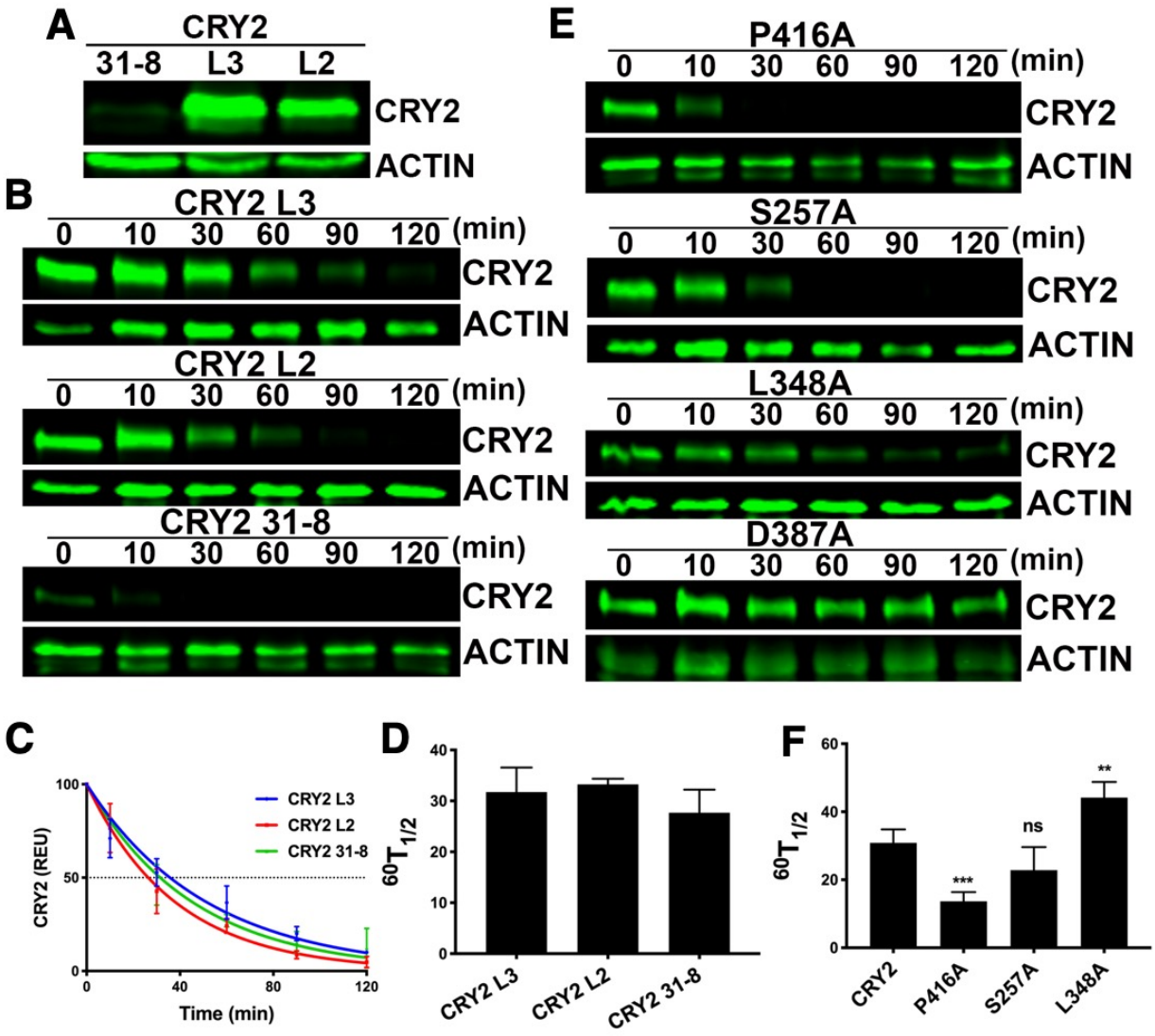

Fig. S2-7: Blue-light dependent proteolysis of CRY2.

A) Immunoblots showing wild-type GFP-CRY2 fusion proteins expressed in transgenic plants. Samples were extracted from seven-day-old seedlings grown in the long-day (16 h day/8 h night) period, fractionated in SDS/PAGE (10%), blotted, probed with anti-CRY2 antibody (CRY2) and anti-ACTIN antibody (ACTIN, control), then probed with fluorophore-conjugated secondary antibodies, and detected by the Odyssey CLx System (LI-COR)

B) Representative immunoblots of samples prepared from six-day-old etiolated seedlings exposed to blue light (60 μmol m^{-2} s^{-1}) for the indicated time in the indicated genotypes. Samples were probed with antibodies to CRY2 and ACTIN (control).

C) Degradation curves of wild-type GFP-CRY2 fusion proteins. Assays were conducted as described in Fig. 2-3H. Bars indicates SD of the mean (n=2).

D) Half-life ($^{60}T_{1/2}$) of proteolysis was calculated from degradation curves of C as described in Fig. 2-3J. Bars indicates SD of the mean (n=2).

E) Representative immunoblots of samples prepared from six-day-old etiolated seedlings exposed to blue light (60 μmol m^{-2} s^{-1}) for the indicated time in the representative mutants of GFP-CRY2 classified as HYPER (P416A), WTL (S257A), HYPO (L348A) and LOF (D387A). Assays were conducted as described in B.

F) Half-life ($^{60}T_{1/2}$) of representative mutants of GFP-CRY2 classified as HYPER (P416A), WTL (S257A), HYPO (L348A) was calculated from degradation curve in Fig. 2-2H as described in Fig. 2-3J. Three asterisks indicate P=0.0003, “ns” indicates P>0.05, double asterisks indicate P=0.0027, compared with half-life of wild-type GFP-CRY2 fusion proteins.

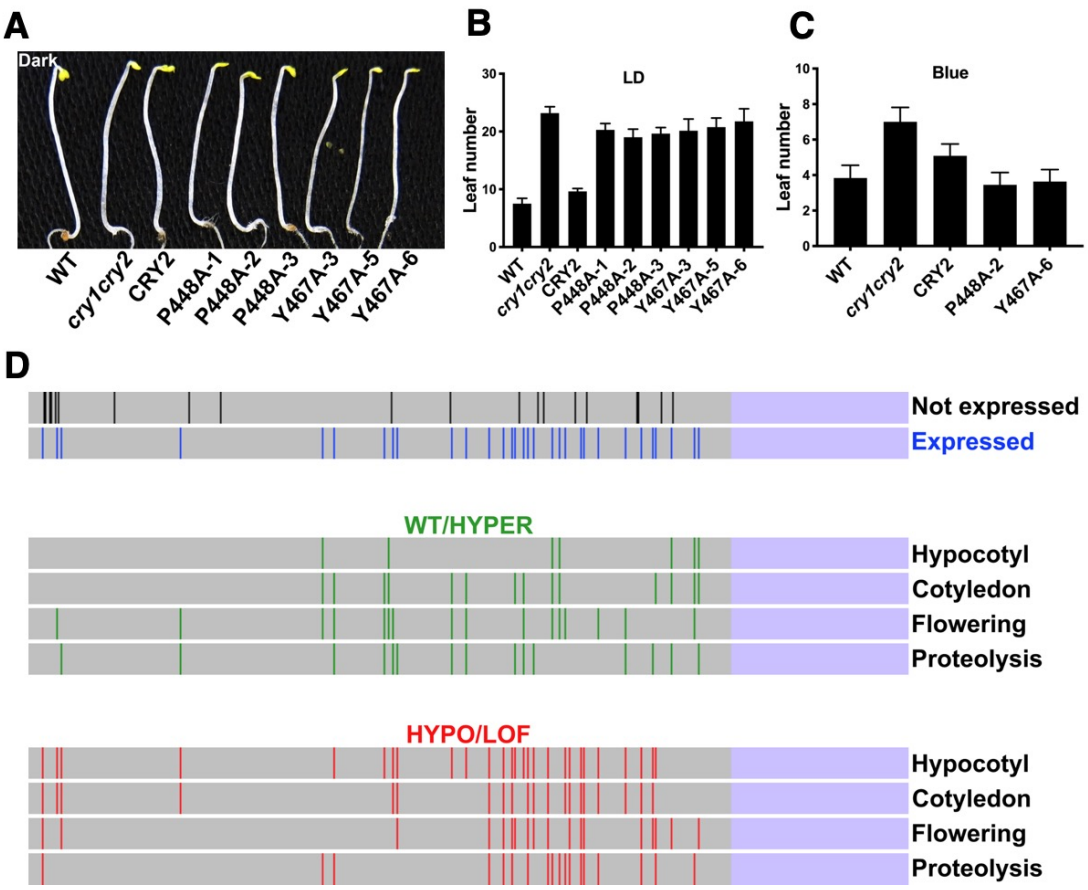

Fig. S2-8: Functional analyses of neighboring UCRs of CRY2.

A) Six day-old seedlings expressing wild-type GFP-CRY2 or GFP- CRY2^{P448A} (P448A), CRY2^{Y467A} (Y467A) mutants, and *cry1cry2* mutant or wild-type seedlings grown in dark. **B-C)** Number of rosette leaves at the time of flowering of independent lines of CRY2^{P448A} and CRY2^{Y467A} mutants in the long-day (16 h day/ 8 h night) period (B), or in the continuous blue light (70-80 μmol•m^{-2}•s^{-1}) (C). Bars indicates SD of the mean (n≥8). **D)** Linear representation of mutants of CRY2. Grey and light blue color represents PHR and CCE domain, respectively. Black, blue, green, red, purple and pink lines indicate activities as labeled at the top of each panel.

2.8 Tables

Table 2-1

Residues	Mutant	Protein Abundance	Hypocotyl inhibition		Cotyledon unfolding		Floral initiation	
			Activity	Category	Activity	Category	Activity	Category
W321A + W374A	2WA1	117.43%	30.03%	HYPO	55.67%	HYPO	38.36%	HYPO
W321A + W397A	2WA2	344.10%	30.97%	HYPO	88.43%	HYPO	77.45%	HYPO
W374A + W397A	2WA3	46.09%	39.38%	HYPO	39.37%	HYPO	57.91%	HYPO
W321A + W374A + W397A	3WA	77.53%	51.94%	HYPO	74.10%	HYPO	87.19%	WTL
W321F + W374F	2WF1	26.98%	17.65%	HYPO	34.15%	HYPO	45.10%	HYPO
W321F + W397F	2WF2	106.56%	30.03%	HYPO	59.06%	HYPO	52.62%	HYPO
W374F + W397F	2WF3	38.20%	28.98%	HYPO	44.46%	HYPO	27.84%	HYPO
W321F + W374F + W397F	3WF	17.95%	29.83%	HYPO	49.15%	WTL	21.05%	HYPO

Table 2-1: Summary of the relative protein abundance and relative specific activities of CRY2 trp-triad mutants.

The protein abundance and physiological activities of double and triple mutants of Trp-triad expressed in the *cry1cry2* parental lines were shown. “HYPER”, “WTL”, “HYPO” and “LOF” represent hypermorph (increased activity), “wild-type”-like, hypomorph (reduced activity) and loss-of-function, respectively. Protein abundance and activities were normalized to L5 of

wild-type GFP-CRY2 line (Fig. 2-3, A-C and Table 2-3). Activities were categorized using the standard curves shown in Fig. 2-3 (A-C).

Table 2-2

Name	Group	Family	Species	NCBI	aa
AtCRY2	Plants-Dicots	Brassicaceae	*Arabidopsis thaliana*	NP_171935.1	1-487
AtCRY1	Plants-Dicots	Brassicaceae	*Arabidopsis thaliana*	NP_567341.1	1-501
GmaCRY1a	Plants-Dicots	Fabaceae	*Glycine max*	XP_025980601.1	1-494
GmaCRY1b	Plants-Dicots	Fabaceae	*Glycine max*	NP_001241002.1	1-495
GmaCRY1c	Plants-Dicots	Fabaceae	*Glycine max*	NP_001235205.1	1-494
GmaCRY1d	Plants-Dicots	Fabaceae	*Glycine max*	NP_001240855.1	1-494
GmaCRY2a	Plants-Dicots	Fabaceae	*Glycine max*	KRG92422.1	1-494
GmaCRY2b	Plants-Dicots	Fabaceae	*Glycine max*	XP_006588363.1	1-494
GmaCRY2c	Plants-Dicots	Fabaceae	*Glycine max*	NP_001241551.1	1-495
OsaCRY1a	Plants-Monocots	Poaceae	*Oryza sativa*	BAB70686.1	1-510
OsaCRY1b	Plants-Monocots	Poaceae	*Oryza sativa*	BAB70688.2	1-501
OsaCRY2	Plants-Monocots	Poaceae	*Oryza sativa*	BAC56984.1	1-486
ZmaCRY1a	Plants-Monocots	Poaceae	*Zea mays*	XP_008644161.1	1-506
ZmaCRY1b	Plants-Monocots	Poaceae	*Zea mays*	PWZ40305.1	1-502
ZmaCRY2	Plants-Monocots	Poaceae	*Zea mays*	XP_008677763.1	1-502
CreCRY	Plants-Green algae	Chlamydomonadaceae	*Chlamydomonas reinhardtii*	XP_001698054.1	1-493
DmCRY	Animals-Invertebrate	Drosophilidae	*Drosophila melanogaster*	NP_732407.1	1-514
DreCRY1a	Animals-Invertebrate	Cyprinidae	*Danio rerio*	NP_001070765.1	1-491
DreCRY1b	Animals-Invertebrate	Cyprinidae	*Danio rerio*	NP_571865.4	1-491
DreCRY2a	Animals-Invertebrate	Cyprinidae	*Danio rerio*	CAQ13306.1	1-491
DreCRY2b	Animals-Invertebrate	Cyprinidae	*Danio rerio*	NP_571867.2	1-494
HsCRY1	Animals-Vertebrate	Hominidae	*Homo sapiens*	NP_004066.1	1-491
HsCRY2	Animals-Vertebrate	Hominidae	*Homo sapiens*	NP_066940.3	1-510
MmCRY1	Animals-Vertebrate	Muridae	*Mus musculus*	NP_031797.1	1-491
MmCRY2	Animals-Vertebrate	Muridae	*Mus musculus*	NP_034093.1	1-509

Table 2-2: Eukaryote proteins used for multiple sequence alignment shown in Fig. S2-2. Fifty-seven residues of CRY2 are universally conserved among the cryptochromes family

of the eukaryotes. NCBI accession numbers, and residues forming the PHR domain of each protein used for alignment were shown in “NCBI” and “aa”, respectively.

Table 2-3

CRY2 residues	Mutant	Protein Abundance	Hypocotyl inhibition		Cotyledon unfolding		Floral initiation		Protein degradation		hCRY1 residues	Previously published mutants			
			Activity	Category	Activity	Category	Activity	Category	$^{60}T_{1/2}$	Category		Species	Protein	Mutant	References
Endogenous CRY2		4.159%	45.5%	N/A	99.88%	N/A	100%	N/A	N/A	N/A	N/A		N/A		
GFP-CRY2 L1		10.41%	43.5%	N/A	51.00%	N/A	36%	N/A	N/A	N/A	N/A		N/A		
GFP-CRY2 L2		40.07%	82.6%	N/A	98.31%	N/A	82%	N/A	33	N/A	N/A		N/A		
GFP-CRY2 L3		53.72%	91.3%	N/A	109.1%	N/A	98%	N/A	32	N/A	N/A		N/A		
GFP-CRY2 L4		64.67%	109%	N/A	96.97%	N/A	97%	N/A	N/A	N/A	N/A		N/A		
GFP-CRY2 L5		100.0%	100%	N/A	100.0%	N/A	100%	N/A	N/A	N/A	N/A		N/A		
W10	W10A	64.60%	1.17%	LOF	21.64%	HYPO	10%	LOF	>120	LOF	W8		N/A		
N20	N20A	72.71%	14.8%	HYPO	30.55%	HYPO	105%	WTL	59	HYPO	N18		N/A		
L23	L23A	24.72%	1.83%	LOF	35.70%	HYPO	28%	HYPO	36	WTL	L21		N/A		
P106	P106A	31.77%	48.8%	HYPO	54.51%	HYPO	75%	WTL	25	WTL	P104		N/A		
L205	L205A	10.67%	44.2%	WTL	49.35%	WTL	60%	WTL	119	HYPO	L205		N/A		
G213	G213A	8.971%	13.0%	HYPO	33.78%	WTL	26%	WTL	43	WTL	G213	*Mus musculus*	mCRY2	G230R	[17]
S248	S248A	6.436%	11.4%	HYPO	38.14%	WTL	26%	WTL	22	WTL	S252	*Mus musculus*	mCRY1	S252A/D	[18]
G254	G254A	13.22%	15.1%	HYPO	13.17%	HYPO	35%	WTL	31	WTL	G258	*Pisum sativum*	PsCRY1	G250E	[51]
												Arabidopsis thaliana	CRY2	G254R	[11]
S257	S257A	115.0%	43.6%	HYPO	49.58%	HYPO	79%	HYPO	26	WTL	S261	*Mus musculus*	mCRY1	S261A/D	[18, 52]
E295	E295A	5.963%	6.1%	LOF	17.96%	WTL	43%	WTL	14	HYPER	E294	*Mus musculus*	mCRY2	E312K	[17]
P305	P305A	20.71%	40.8%	HYPO	61.12%	WTL	76%	WTL	29	WTL	P304		N/A		
W321	W321A	414.1%	44.2%	HYPO	74.89%	HYPO	46%	HYPO	113	HYPO	W320	Previously published Trp-triad residues were not listed.			
W331	W331A	26.91%	5.84%	LOF	17.04%	HYPO	0.00%	LOF	>120	LOF	W330	*Arabidopsis thaliana*	CRY1	W334A/F	[13]
												Arabidopsis thaliana	CRY1	G340E	[1, 53]
G337	G337A	28.06%	33.9%	HYPO	49.61%	HYPO	11%	LOF	88	HYPO	G336	*Mus musculus*	mCRY1	G336D	[54]
												Mus musculus	mCRY2	G354D	[17]
P339	P339A	85.98%	32.4%	HYPO	86.56%	WTL	37%	HYPO	35	WTL	P338	*Mus musculus*	mCRY2	P356L	[17]
M345	M345A	22.67%	43.4%	HYPO	93.04%	WTL	86%	WTL	39	WTL	M344		N/A		
L348	L348A	48.37%	0.460%	LOF	2.248%	LOF	25%	HYPO	53	HYPO	K347	*Arabidopsis thaliana*	CRY2	L348F	[14]
G352	G352A	23.09%	5.51%	LOF	23.16%	HYPO	3.6%	LOF	30	WTL	G351		N/A		

F365	F365A	38.09%	76.6%	WTL	101.3%	WTL	97%	WTL	53	HYPO	F364	N/A			
L370	L370A	30.35%	132%	HYPER	70.24%	WTL	95%	HYPER	60	HYPO	L370	N/A			
W374	W374A	69.86%	58.9%	HYPO	47.76%	HYPO	92%	WTL	>120	LOF	W374	Previously published Trp-triad residues were not listed.			
L385	L385A	26.59%	23.3%	HYPO	34.61%	HYPO	1.2%	LOF	98	HYPO	L385	N/A			
												Drosophila melanogaster	dCRY	D410N	[15]
D387	D387A	99.56%	0.910%	LOF	8.592%	LOF	19%	LOF	>120	LOF	D387	*Arabidopsis thaliana*	CRY2	D387A	[39]
												Mus musculus	mCRY1	D387N	[55]
W397	W397A	16.94%	21.2%	HYPO	21.59%	HYPO	78%	WTL	>120	LOF	W397	Previously published Trp-triad residues were not listed.			
P416	P416A	11.88%	14.8%	HYPO	24.39%	HYPO	28%	WTL	17	HYPER	P415	N/A			
G427	G427A	6.763%	6.17%	LOF	3.049%	LOF	0.00%	LOF	>120	LOF	G426	N/A			
P435	P435A	26.13%	34.9%	HYPO	46.81%	HYPO	18%	HYPO	35	WTL	P434	N/A			
L437	L437A	13.33%	28.8%	HYPO	57.44%	WTL	26%	HYPO	57	HYPO	L436	N/A			
P448	P448A	19.19%	130%	HYPER	65.58%	WTL	12%	LOF	29	WTL	P447	N/A			
G464	G464A	85.24%	89.5%	WTL	81.70%	HYPO	92%	WTL	>120	LOF	G463	N/A			
Y467	Y467A	10.13%	145%	HYPER	98.32%	HYPER	20%	HYPO	25	WTL	Y466	N/A			

Table 2-3: Summary of the relative protein abundance and relative specific activities of UCR mutants of CRY2 with stable protein expression.

The protein abundance and activities of wild-type and mutant GFP-CRY2 expressed in the *cry1cry2* parental lines were shown. “HYPER”, “WTL”, “HYPO” and “LOF” represent hypermorph (increased activity), “wild-type”-like, hypomorph (reduced activity) and loss of function, respectively. Protein abundance and activities were normalized to L5 of wild-type GFP-CRY2 line. Activities were categorized using the standard curves shown in Fig. 2-3 (A-C). N/A, not applied.

Table 2-4

CRY2 residues	Mutant	hCRY1 residues	Previously published mutants			
			Species	Protein	Mutant	References
F11	F11A	F9		N/A		
R12	R12A	R10		N/A		
L15	L15A	L13		N/A		
R16	R16A	R14		N/A		
D19	D19A	D17	Arabidopsis	CRY1	D21N	[53]
P21	P21A	P19		N/A		
L60	L60A	L59		N/A		
D112	D112A	D110		N/A		
L134	L134A	L132		N/A		
F253	F253A	F257	mouse	mCRY1	F257A	[16]
R294	R294A	R293		N/A		
D342	D342A	D341		N/A		
H355	H355A	H354		N/A		
R359	R359A	R358	mouse	mCRY1	R358K	[55]
F381	F381A	F381		N/A		
D389	D389A	D389		N/A		
D424	D424A	D423		N/A		
P425	P425A	P424		N/A		
P441	P441A	P440		N/A		
W449	W449A	W448		N/A		

Table 2-4: Summary of UCR mutants of CRY2 without stable expression from the "lack of protein" group. N/A, not applied.

Table 2-5

Arabidopsis	Previously reported cryptochrome mutants					
CRY2 residue	Species	Protein	Mutation	Category	Phenotype	Reference
T7	Mm	mCRY2	S23L	Non-UCR	circadian	[17]
D19	At	CRY1	D21N (cry1-401)	UCR	hypocotyl	[53]
C39	Mm	mCRY1	D38A	Non-UCR	circadian	[16]
P40	Mm	mCRY1	P39G	Non-UCR	circadian	[16]
E42	Mm	mCRY1	F41S	Non-UCR	circadian	[16]
R53	Mm	mCRY1	R51A	UCR	circadian	[16]
S59	At	CRY1	S66N (cry1-388)	Non-UCR	hypocotyl	[56]
S72	Mm	mCRY1	S71A/D	Non-UCR	interaction	[52]
S72	Mm	mCRY1	S71A/D	Non-UCR	circadian	[18]
Y104	Mm	mCRY1	S102A/D	Non-UCR	circadian	[18]
D105	Mm	mCRY1	E103K	Non-UCR	circadian	[16]
D105	Mm	mCRY2	E121K	Non-UCR	circadian	[17]
V107	Mm	mCRY1	F105A	Non-UCR	circadian	[16]
S108	Mm	mCRY1	G106R	Non-UCR	circadian	[17]
S108	Mm	mCRY1	G106R/W	Non-UCR	circadian	[16]
R111	Mm	mCRY1	R109Q	UCR	circadian	[17]
L133	Mm	mCRY1	T131A/D	Non-UCR	circadian	[18]
K156	Mm	mCRY1	T155A/D	Non-UCR	circadian	[18]
L159	Mm	mCRY1	S158A/D	Non-UCR	circadian	[18]
G213	Mm	mCRY2	G230R	UCR	circadian	[17]
W214	Mm	mCRY1	E214K	Non-UCR	circadian	[17]
N216	Mm	mCRY1	E216	Non-UCR	circadian	[17]
A217	At	CRY1	G220D (hy4-6)	Non-UCR	hypocotyl	[57]
A217	At	CRY1	G220D (hy4-6)	Non-UCR	flower	[58]
A217	Mm	mCRY1	A217V	Non-UCR	circadian	[17]
D218	Mm	mCRY1	L218F	Non-UCR	circadian	[17]
F224	Mm	mCRY1	H224E	Non-UCR	interaction	[54]
Y232	Mm	mCRY1	A232T	Non-UCR	circadian	[17]
G241	Mm	mCRY1	S243A/D	Non-UCR	circadian	[18]
S243	Mm	mCRY1	S247A/D	Non-UCR	circadian	[18]
S243	Mm	mCRY1	S247D	Non-UCR	interaction	[54]
S245	Mm	mCRY1	T249A/D	Non-UCR	circadian	[18]
S248	Mm	mCRY1	S252D	UCR	circadian	[18]
Y250	Mm	mCRY1	Y254A/D	Non-UCR	circadian	[18]
F253	Mm	mCRY1	F257A	UCR	circadian	[16]
G254	At	CRY2	G254R	UCR	flower	[11]
E255	Mm	mCRY1	C259Y	Non-UCR	circadian	[17]
S257	Mm	mCRY1	S261A/D	UCR	circadian	[18]
K268	Mm	mCRY1	T270A/D	Non-UCR	circadian	[18]
S278	Mm	mCRY1	S280A/D	Non-UCR	interaction	[52]
S278	Mm	mCRY1	S280A/D	Non-UCR	circadian	[18]
G280	At	CRY1	G283E (hy4-5)	Non-UCR	hypocotyl	[57]
S283	At	CRY1	S286N (cry1-402)	Non-UCR	hypocotyl	[53]
D285	Mm	mCRY1	S285A/D	Non-UCR	circadian	[18]
R289	Mm	mCRY1	G288R	Non-UCR	circadian	[17]
E295	Mm	mCRY2	E312K	UCR	circadian	[17]
Q310	Mm	mCRY1	M309I	Non-UCR	circadian	[17]
G334	At	CRY1	G337D (hy4-4)	UCR	hypocotyl	[1]
G334	Mm	mCRY2	G351D	UCR	circadian	[17]
G337	At	CRY1	G340E (cry1-404)	UCR	hypocotyl	[53]
G337	At	CRY1	G340E (hy4-1)	UCR	hypocotyl	[1]
G337	Mm	mCRY1	G336D	UCR	circadian	[54]
G337	Mm	mCRY2	G354D	UCR	circadian	[17]
P339	Mm	mCRY2	P356L	UCR	circadian	[17]

G344	At	CRY1	G347E (hy4-16)	Non-UCR	hypocotyl	[57]
G344	At	CRY1	G347R (hy4-15)	Non-UCR	hypocotyl	[57]
G344	At	CRY1	G347R (cry1-375)	Non-UCR	hypocotyl	[56]
L348	At	CRY2	L348F	UCR	interaction	[14]
W349	At	CRY2	W349R	Non-UCR	interaction	[14]
T351	Mm	mCRY2	E368K	Non-UCR	circadian	[17]
N356	Mm	mCRY1	H355E	Non-UCR	interaction	[54]
V367	At	CRY2	V367M	Non-UCR	cotyledon	[59]
V367	At	CRY2	V367M	Non-UCR	flower	[60]
G377	At	CRY1	G380R (cry1-344)	UCR	hypocotyl	[56]
W382	Mm	mCRY1	E382A	Non-UCR	circadian	[16]
D393	Mm	mCRY1	N393A/C	Non-UCR	circadian	[18]
Y399	At	CRY2	Y399A/F	Non-UCR	hypocotyl/degradation	[61]
I404	At	CRY1	L407F	Non-UCR	flower/germination	[62]
I404	Mm	mCRY1	S404A/D	Non-UCR	circadian	[18]
P405	Mm	mCRY1	F405A	Non-UCR	interaction	[54]
G420	Mm	mCRY2	G437D (cry1-305)	Non-UCR	circadian	[17]
W433	Mm	mCRY1	Y432A/D	Non-UCR	circadian	[18]
S459	At	CRY1	A462V (cry1-305)	Non-UCR	hypocotyl	[56]
S486	Mm	mCRY1	K485D/E	Non-UCR	interaction	[54]
E490	At	CRY2	E490G	Non-UCR	interaction	[63]
N/A	At	CRY1	E508K (cry1-349)	Non-UCR	hypocotyl	[56]
N/A	At	CRY1	E515K (hy4-19)	Non-UCR	hypocotyl	[57]
N/A	At	CRY1	E531K (hy4-20)	Non-UCR	hypocotyl	[57]
N/A	At	CRY1	R536K (cry1-321)	Non-UCR	hypocotyl	[56]
N/A	At	CRY1	P549L (hy4-9)	Non-UCR	hypocotyl	[57]
N/A	At	CRY1	G380R (cry1-344)	Non-UCR	hypocotyl	[56]
N/A	At	CRY2	K541R	Non-UCR	hypocotyl/degradation	[36]
N/A	At	CRY1	E559L (hy4-22)	Non-UCR	hypocotyl	[57]
N/A	At	CRY1	R581K (hy4-23)	Non-UCR	hypocotyl	[57]
N/A	At	CRY1	611K (hy4-24)	Non-UCR	hypocotyl	[57]
N/A	At	CRY1	E623K (cry1-403)	Non-UCR	hypocotyl	[53]
N/A	Mm	mCRY1	S492A/D	Non-UCR	circadian	[18]
N/A	Mm	mCRY1	S588A/D	Non-UCR	circadian	[18]

Table 2-5: List of previously reported mutants of cryptochromes included in Fig. 2D. Only residues not selected based on level of sequence conservation were included here. Different mutations in the same residue were considered as 2 hits (1) if the mutations were separately examined in two independent studies (e. g. residue S72 of Arabidopsis CRY2), or (2) if a residue was mutated into two different other residues in the same phenotype-based genetic screen in a study (e. g. mutant hy4-15 and hy4-16). In contrast, different mutations of the same residue were considered as 1 hit if the residue was rationally selected to be changed into different residues in the same study

(e. g. Y104 of Arabidopsis CRY2). N/A, not applied because corresponding residues are not available. At, *Arabidopsis thaliana*. Mm, *Mus Musculus*.

Table 2-6

pFGFP-C-CRY2F	tccagctccaggatccATGAAGATGGACAAAAAGAC
pFGFP-C-CRY2R	GAGAAAGCTTGGATCC TCATTTGCAACCATTTTTTCCC
CRY2-W10A-36R	TCTAAACGCAACTATAGTCT
CRY2-W10A-22F	ATAGTTGCGTTTAGAAGAGA
CRY2-F11A-39R	TCTTCTAGCCCAAACTATAG
CRY2-F11A-25F	GTTTGGGCTAGAAGAGACCT
CRY2-R12A-42R	GTCTCTTGCAAACCAAACTA
CRY2-R12A-28F	TGGTTTGCAAGAGACCTAAGGA
CRY2-L15A-51R	AATCCTTGCGTCTCTTCTAA
CRY2-L15A-37F	AGAGACGCAAGGATTGAGGA
CRY2-R16A-54R	CTCAATTGCTAGGTCTCTTC
CRY2-R16A-40F	GACCTAGCAATTGAGGATAA
CRY2-D19A-63R	AGGATTAGCCTCAATCCTTA
CRY2-D19A-49F	ATTGAGGCTAATCCTGCATT
CRY2-N20A-66R	TGCAGGAGCATCCTCAATCC
CRY2-N20A-52F	GAGGATGCTCCTGCATTAGC
CRY2-P21A-69R	TAATGCAGCATTATCCTCAA
CRY2-P21A-55F	GATAATGCTGCATTAGCAGC
CRY2-L23A-75R	TGCTGCTGCTGCAGGATTAT
CRY2-L23A-61F	CCTGCAGCAGCAGCAGCTGC
CRY2-L60A-186R	GTGAGCAGCTGATTGTTTCA
CRY2-L60A-172F	CAATCAGCTGCTCACTTATC
CRY2-P106A-324R	CGAAACAGCATCATAGAGGT
CRY2-P106A-310F	TATGATGCTGTTTCGTTAGT
CRY2-D112A-342R	GGTATGGGCCCGAACTAACG
CRY2-D112A-328F	GTTCGGGCCCATACCGTAAA
CRY2-L134A-408R	TTCATACGCTAGATCTCCAT
CRY2-L134A-394F	GATCTAGCGTATGAACCGTG
CRY2-L205A-621R	AGTTAACGCCGCATTGCTCG
CRY2-L205A-607F	AATGCGGCGTTAACTAGAGC
CRY2-G213A-645R	GCTCCATGCTGGAGACCAAG
CRY2-G213A-613F	TCTCCAGCATGGAGCAATGC
CRY2-S248A-750R	ATACGGAGCAAGTAGTGAAG
CRY2-S248A-736F	CTACTTGCTCCGTATCTCCA
CRY2-F253-765R	TTCCCCGGCATGGAGATACG
CRY2-F253A-751F	CTCCATGCCGGGGAAATAAG
CRY2-G254-768R	TATTTCCGCGAAATGGAGAT
CRY2-G254A-754F	CATTTCGCGGAAATAAGCGT
CRY2-S257A-777R	TCTGACGGCTATTTCCCCGA
CRY2-S257A-763F	GAAATAGCCGTCAGACACGT
CRY2-R294A-888R	ATACTCTGCTAAACCGATTC
CRY2-R294A-874F	GGTTTAGCAGAGTATTCTCGGT

CRY2-E295A-891R	AGAATACGCTCTTAAACCGA
CRY2-E295A-877F	TTAAGAGCGTATTCTCGGTA
CRY2-P305A-921R	AGTAAACGCGAAGTTGAAAC
CRY2-P305A-907F	AACTTCGCGTTTACTCACGA
CRY2-W321A	
CRY2-W331A-999R	TTGTCTCGCGGCCTTGAACT
CRY2-W331A-1004F	AAGGCCGCGAGACAAGGCAG
CRY2-G337A-1017R	CGGATAAGCGGTCCTGCCTT
CRY2-G337A-1003F	AGGACCGCTTATCCGTTGGT
CRY2-P339A-1023R	CACCAACGCATAACCGGTCC
CRY2-P339A-1009F	GGTTATGCGTTGGTGGATGC
CRY2-D342A-1032R	TCCGCCAGCCACCAACGGAT
CRY2-D342A-1018F	TTGGTGGCTGCCGGAATGAGA
CRY2-M345A-1041R	CTCTCTCGCTCCGGCATCCA
CRY2-M345A-1027F	GCCGGAGCGAGAGAGCTTTG
CRY2-L348A-1050R	AGCCCAAGCCTCTCTCATTC
CRY2-L348A-1036F	AGAGAGGCTTGGGCTACCGG
CRY2-G352A-1062R	CATCCATGCGGTAGCCCAAA
CRY2-G352A-1048F	GCTACCGCATGGATGCATAA
CRY2-H355A-1071R	TCTGTTAGCCATCCATCCGGT
CRY2-H355A-1057F	TGGATGGCTAACAGAATAAGA
CRY2-R359A-1083R	AATCACTGCTATTCTGTTAT
CRY2-R359A-1069F	AGAATACGAGTGATTGTTTC
CRY2-F365A-1101R	CACAGCAGCGCTTGAAACAA
CRY2-F365A-1087F	TCAAGCGCTGCTGTGAAGTT
CRY2-L370A-1116R	AAGGAGAGCAAACTTCACAG
CRY2-L370A-1102F	AAGTTTGCTCTCCTTCCATG
CRY2-W374A-1128R	CCATTTGGCTGGAAGGAGAA
CRY2-W374A-1114F	CTTCCAGCCAAATGGGGAAT
CRY2-F381A-1149R	ATCCCAGGCATACTTCATTC
CRY2-F381A-1135F	AAGTATGCCTGGGATACACT
CRY2-L385A-1161R	ATCCAAAGCTGTATCCCAGA
CRY2-L385A-1147F	GATACAGCTTTGGATGCTGA
CRY2-D387A-1167R	ATCAGCAGCCAAAAGTGTAT
CRY2-D387A-1153F	CTTTTGGCTGCTGATTTGGA
CRY2-D389A-1173R	TTCCAAAGCAGCATCCAAAA
CRY2-D389A-1159F	GATGCTGCTTTGGAATGTGA
CRY2-W397A-1197R	ATACTGGGCGCCAAGGATGTC
CRY2-W397A-1183F	CTTGGCGCCCAGTATATCTC
CRY2-P416A-1254R	TAACGCGGCATTGTCCAAGC
CRY2-P416A-1259F	GACAATGCCGCGTTACAAGG
CRY2-D424A-1278R	TTCTGGGGCATATTTGGCGC
CRY2-D424A-1264F	AAATATGCCCCAGAAGGTGA
CRY2-P425A-1281R	ACCTTCTGCGTCATATTTGG

CRY2-P425A-1267F	TATGACGCAGAAGGTGAGTA
CRY2-G427A-1287R	GTACTCAGCTTCTGGGTCAT
CRY2-G427A-1273F	CCAGAAGCTGAGTACATAAG
CRY2-P435A-1311R	AAGCTCGGCAAGCCATTGCC
CRY2-P435A-1297F	TGGCTTGCCGAGCTTGCGAG
CRY2-L437A-1317R	TCTCGCAGCCTCGGGAAGCC
CRY2-L437A-1303F	CCCGAGGCTGCGAGATTGCC
CRY2-P441A-1315R	TTCAGTTGCCAATCTCGCAA
CRY2-P441A-1315F	AGATTGGCAACTGAATGGAT
CRY2-P448A-1350R	GTCCCATGCATGATGGATCC
CRY2-P448A-1336F	CATCATGCATGGGACGCTCC
CRY2-W449A-1353R	AGCGTCCGCTGGATGATGGA
CRY2-W449A-1339F	CATCCAGCGGACGCTCCTTT
CRY2-G464A-1398R	GTTTGTTGCGAGTTCCACAC
CRY2-G464A-1384F	GAACTCGCAACAAACTATGC
CRY2-Y467A-1407R	TTTCGCAGCGTTTGTTCCGA
CRY2-Y467A-1393F	ACAAACGCTGCGAAACCCAT

Table 2-6: Primers used for site-directed mutagenesis.

Chapter 3: Continuous directed evolution of plant blue-light-dependent interactions

3.1 Abstract

Arabidopsis thaliana Cryptochrome2 (CRY2) is a blue-light receptor that mainly regulates plant photomorphogenesis through blue-light-specific interactions with numerous protein partners. Such blue-light-specific interactions have been exploited in optogenetics to manipulated biological events in a timely and precisely manner. CRY2 and its blue-light specific interacting partners have been widely used as optogenetic tools. This study focused on the development of a novel pair of blue-light-dependent interacting proteins: CRY2-BIC1 (Blue-light Inhibitor of Cryptochromes 1). BIC 1 protein is small (15 kDa) and can inhibit the unwanted homooligomerization of CRY2 in optogenetic applications. However, the dynamic range of CRY2-BIC1 blue-light dependent interaction remains to be increased. To evolve CRY2 to increase its interaction affinity with BIC1 in the blue light, I applied a novel continuous directed evolution method, PACE (Phage Assisted Continuous Evolution), which coupled in vivo random mutagenesis with selection, enabling rounds of automatic evolution with minimal human efforts. I isolated variants of CRY2 with stronger interactions with BIC1. To facilitate evolution of CRY2 and to increase the dynamic range and signal-to-background ratio of the CRY2-BIC1 photodimerizer, soluble expression and protein-dissociating PACE methods were developed to increase the solubility of CRY2 in *E. coli*

cells and to reduce interacting background in darkness. PACE variants of CRY2 were introduced in *Arabidopsis thaliana* plants to examining how physiological functions of CRY2 were influenced by altered interaction affinity.

3.2 Introduction

Optogenetics is an innovative method to control biological processes. By combining genetic and optical approaches, it allows fast and targeted regulation of biological events in a wide range of biological systems. Light-dependent dimerization is an important branch of optogenetics, using light-induced protein-protein interaction to manipulate biological events in a timely and precisely manner. For example, CRY2-CIB1 dimerizer was used to regulate transcription [1, 2], Cre-recombinase activity [3, 4], phosphoinositide metabolism [5, 6], signaling [7, 8] and other cellular functions [9, 10]. CRY2-CIB1 is one of the most wildly used interaction pairs because of its rapid light response, strong interaction strength and reversibility [4]. However, CRY2 was found to form homooligomers when illuminated with blue light [11], which poses complications for optogenetic applications in two aspects. First, compared with components in natural pathways, light-induced oligomerization of pathway components may pose different artificial conformations, leading to different signaling outcomes [12]. Second, homooligomerization renders it very hard, if not impossible to use photo-dimerizers to exclusively study the effects of heterodimerization of signaling components. For example, both homodimerization of C-RAF (RAF proto-oncogene serine/threonine-protein kinase) and heterodimerization between C-RAF and B-RAF (serine/threonine-protein kinase B-Raf) leads to activation of downstream pathways [13]. However,

heterodimerization of C-RAF-B-RAF cannot be induced without eliciting homodimerization of C-RAF using CRY2-CIB1 heterodimerizer or CRY2 oligomer. Although CRY2 variants with increased homooligomerization were isolated [14, 15], efforts to limit homooligomerization without affecting CRY2-CIB1 heterodimerization only have modest success [15]. BIC1 naturally inhibits CRY2 homooligomerization through blue-light specific interactions with CRY2. In optogenetic applications, problems introduced by CRY2 homooligomerization can be circumvented by using CRY2-BIC1 photodimerizer. However, the interaction dynamic range (the difference in interaction strengths in blue light and dark condition) of CRY2-BIC1 is limited, at least in mammalian cells, compared with that of CRY2-CIB1. To increase the interaction dynamic range of CRY2-BIC1 dimerizer, I used the protein-binding PACE to evolve CRY2 to get mutants with improved dynamic range. CRY2 variants with increased interaction affinity with BIC1 was isolated.

CRY2 has various blue-light specific interacting protein partners, such as CIB1, SPA1 and BIC1. Different interacting proteins may appear in different CRY2-containing complexome and different complexome may be formed at the same time. The PACE-evolved CRY2 variants have changed interaction affinity with BIC1. To investigate how different complexome reach an equilibrium in *Arabidopsis* cells, I introduced the evolved CRY2 variants into the *Arabidopsis thaliana cry1cry2* CRY-deficient background and examined the physiological functions of the CRY2 variant proteins.

3.3 Results

3.3.1 Establishment of the protein-binding PACE

PACE has been successfully used to evolve T7 RNAP with different promoter specificity [16–19], protease with drug resistance [20] or changed substrate specificity [21], DNA binding affinity [22], orthogonal aminoacyl-tRNA synthetases [23], Cas9 with changed PAM compatibility [24], soluble expression of proteins [25], single-chain variable fragments (scFv) [25], improved base editor [26], split T7 RNAP [27–29] and esterase [30]. PACE for evolution of stronger PPI affinity has also been well-characterized and used to evolve *Bacillus thuringiensis* (Bt) toxin to overcome resistance of insects [31]. The protein-binding PACE exploits bacterial two-hybrid system (B2H) [32] to link PPI affinity with transcription of gIII of M13 phage. Since the set-up of the PACE system is complex, it is important to make sure I can conduct successful evolution experiments using PACE. I used model of high-affinity interaction (HA4 monobody and SH2 domain of ABK1 kinase), the same as the one previously used to develop and optimize the protein-binding PACE [31]. HA4 interacts with SH2 with high affinity (K_D = 7 nM), while HA4^{Y87A} dramatically reduces their interaction by 100- to 1,000-fold [33]. I evolved HA4^{Y87A} on SH2 using the protein-binding PACE and observed if the HA4-SH2 interaction could be restored (Fig. 3-1A). First, HA4^{Y87A} was subjected to 3 days of continuous diversification by supporting M13 phage propagation regardless of interaction affinity (genetic drift) to generate a diversified library (Fig. 3-1B and C). Then we subjected the diversified evolving proteins to both mutagenesis and selection pressure (Fig. 3-1C). After 3-day of continuous evolution and selection, HA4^{Y87A} was no longer detectable. In 2 parallel lagoons, HA4^{Y87A} variant was either

changed into HA4^{87M} (L1) or HA4^{87W} (L2) that can support phage propagation (Fig. 3-1C).

3.3.2 Evolution of CRY2 to bind BIC1

The interaction between CRY2 and BIC1 was reported in *Arabidopsis thaliana* plants and HEK293T (human embryonic kidney) cells [11]. However, it is not clear if the blue-light dependent CRY2-BIC1 interaction is detectable in *E. coli* cells, using the B2H scheme of PACE. Therefore, I checked if the transcription of gIII and the following luciferase genes could be activated by CRY2-BIC1 interaction. The interaction between SH2 and HA4^{Y87A} is very weak and served as negative control (Fig. 3-2A). The interaction between BIC1 and CRY2 was significantly stronger in the blue light than in darkness (Fig. 3-2A), suggesting detectable blue-light specific interaction of CRY2-BIC1 in *E. coli* cells. However, the BIC1-CRY2 interaction in the blue light was much weaker than the SH2-HA4 interaction (Fig. 3-2A), indicating that there was plenty of room for improvement. Therefore, it was suitable to evolve CRY2-BIC1 interaction using the protein-binding PACE.

Apart from the CRY2-BIC1 interaction, I also used the same B2H scheme to explore homodimerization of *E. coli* PHL. *E. coli* PHL was reported to homodimerize in a large scale Y2H (yeast two hybrid) screen [34]. However, such interaction was not verified in vivo. I examined homodimerization of PHL in the B2H system. Surprisingly, PHL did not show signs of homodimerization (Fig. 3-2B). Instead, tethering PHL with DNA-binding domain (DBD) suppressed transcription (Fig. 3-2B), the mechanism of which was unclear but was not further studied in this work.

CRY2 was evolved using the protein-binding PACE to increase its blue-light dependent interaction with BIC1 in 3 independent lagoons (Fig. 3-3A). The evolution experiment was conducted in 2 sections: mutagenesis only, followed by mutagenesis together with selection. Three independent lagoons produced completely different results (Fig. 3-3A), manifested randomness of evolution. In L2, phage titer quickly dropped to non-detectable concentration (washed-out), indicating a failure in evolving better performing CRY2, possibly because of lack of beneficial mutations after the genetic drift section. Phage populations were not washed-out in L1 and L3, so I used PCR (polymerase chain reaction) to test the evolved sequences of individual phage plaques in L1 and L3 (Fig. 3-3B). The evolving sequences in L1 were still full-length *CRY2*. However, the length of the evolving sequences in L3 was equivalent to that of gIII, indicating a recombination event that swapped full-length *CRY2* for gIII had occurred. Sanger sequencing of the individual plaque SP DNA confirmed the recombination event, which led to the dramatically increased phage titer in L3. Bulk Sanger sequencing of SP in L1 revealed 2 accumulated mutations of CRY2 (Fig. 3-3C-E). The detection limit of Sanger sequencing is ca. 15%, that is, only mutations enriched to at least 15% of the population are detectable by Sanger sequencing. The CRY2 mutations arose and accumulated to at least 15% of phage population in L1 within 24 ($CRY2^{H102R}$) or 48 ($CRY2^{H408P}$) hours of selection and persisted in L1 until I stopped the selection (Fig. 3-3C). The enrichment and accumulation of mutations indicated that the evolved variants possibly conveyed strong interaction with BIC1.

Sanger sequencing of individual phage plaques (rather than sequencing in bulk) revealed multiple not-enriched mutations apart from the 2 enriched mutations. I

constructed mutagenic spectra of the PACE system using MP6 (Table 3-1) on *CRY2* gene (Fig. 3-4). Consistent with previous mutagenic spectra, little C:G → G:C and A:T → A:T transversion was detected. However, the mutagenic spectra showed close to equal occurrence of transition and transversion events, while previous mutagenic spectra of MP6 on *rpoB* and *lacZ* genes showed markedly more transition over transversion events [35], indicating that the mutagenic spectra of a mutagenic system might demonstrate subtle shifts when targeted to different DNA. Alternatively, my mutagenic spectra might suffer from small sampling size (n=82) and was not representative.

3.3.3 Characterization of evolved CRY2 variants

HEK293T cells have been used as an exogenous system to explore *Arabidopsis* CRY2-related PPI, avoiding the interference of endogenous plant proteins [11]. I cloned the PACE variants of CRY2 into mammalian expression vectors and co-expressed the proteins with BIC1 or other proteins in HEK293T cells for examining PPI using co-immunoprecipitation (co-IP). Compared with wild-type (wt) CRY2, $CRY2^{H102R}$ and $CRY2^{H408P}$ showed stronger interaction with BIC1, regardless of the presence of light (Fig. 3-5A). The CRY2 variants lost the blue-light response and interacted with BIC1 constitutively. Interestingly, CRY2 homooligomerization, interaction with SPA1 and CIB1 in the blue light were all increased for the PACE variants of CRY2 (Fig. 3-5 B-E). Since the affinity of CRY and any tested interaction partner was increased by the PACE variants, the increased interaction with BIC1 was most likely not through a BIC1-specific mechanism. Instead, some general characteristics of CRY2 might be changed. I tested

the thermal stability of CRY2 variants in HEK293T cells by monitoring decreased luminescence after addition of cycloheximide (CHX) that blocks translation (Fig. 3-5F and G). The rate of degradation and half-life of the PACE variants of CRY2 was similar to that of wild-type CRY2. The increase of overall interaction affinity was not a result of increased thermal stability. The detailed mechanism remained to be explored.

The loss of blue-light responses in CRY2-BIC1 interaction, and the comprehensively enhanced interaction between CRY2 and other CRY2-interacting proteins, might push CRY2 complexomes to a new balance. To explore the effect of change in the balance of CRY2 complexomes, I introduced 4MYC-CRY2^{H102R} (4M-CRY2^{H102R}) and 4MYC-CRY2^{H408P} (4M-CRY2^{H408P}) fusion proteins into *cry1cry2* double mutant *Arabidopsis* plants to examine their ability to rescue the mutant phenotype. The fusion of 4 tandem MYC epitopes to CRY2 was previously reported not to influence physiological functions of CRY2 [11]. Physiological activities are usually determined by two factors: protein abundance and intrinsic specific activity of the protein. To estimate if specific activity is affected by a mutation, protein abundance needs to be accounted. As in Chapter 2, I created a standard curve concerning protein abundance and relative specific activity (calculated from quantified physiological activity), using multiple transgenic lines of 4MYC-wild-type CRY2 (4M-CRY2) in *cry1cry2* plants (Fig. 3-6). Six independent transgenic lines of 4M-CRY2 with different protein abundance were used to construct the hypocotyl inhibition standard curve in blue light (1 μmol m^{-2} s^{-1}). Nine or 8 independent transgenic lines of 4M-CRY2^{H102R} or 4M-CRY2^{H408P} were plotted based on their protein abundance and relative specific activity in blue-light induced hypocotyl inhibition. All 4M-CRY2^{H102R} plants and the majority (6/8) of 4M-CRY2^{H408P} were

located within the upper and lower 95% of prediction bands of the standard curve, suggesting the specific activities of $CRY2^{H102R}$ and $CRY2^{H408P}$ in inhibition of hypocotyl growth under 1 μmol m^{-2} s^{-1} of blue light were not significantly changed compared to wild-type CRY2.

3.3.4 Development of the soluble expression PACE

The solubility of CRYs is limited in *E. coli* cells [36], which may impede the engineering of CRY2 by PACE. Before subjecting CRY2 with further rounds of protein-binding PACE, I planned to develop a soluble expression PACE system to increase the solubility of CRY2 in *E. coli* cells (Fig. 3-7A and B). I sandwiched gene-of-interest (GOI) between transcriptional activation domain (AD) and DBD. Expression of gIII depends on the accessibility of AD and DBD. Misfolded protein-of-interest (POI) would drag AD and DBD into inclusion bodies, render them inaccessible to turn on transcription of gIII, while soluble expression of POI would facilitate transcription of gIII (Fig. 3-7B). I designed 4 versions of construct by manipulating protein fusion linkers and orientation of AD and DBD domains relative to GOI (Fig. 3-7A). Version 4 offered the highest dynamic range when plugged in soluble and insoluble proteins (Fig. 3-7C). Using this version of soluble expression PACE constructs, I first determined if the evolved PACE variants of CRY2 have increased solubility in *E. coli* cells (Fig. 3-7D). $CRY2^{H102R}$ and $CRY2^{H408P}$, and wildt-type CRY2 showed comparable solubility, demonstrating the enrichment of the mutants was not the result of increased solubility.

Solubility of CRYs can also be increased by truncation of CRYs from C-terminal CCE domain [36], without significantly damaging interaction affinity or dynamic range

[4]. Using the soluble expression PACE constructs, CRY2-534 (residue 1–534), CRY2-502 (residue 1–502) and CRY2-489 (residue 1–489) showed increased solubility compared with full-length CRY2 by 1.5 to 2 folds (Fig. 3-7D). I chose CRY2-502 for further rounds of protein-binding PACE because it showed greater increase in solubility than CRY2-534 but retained more residues than CRY2-489.

In previous PACE experiments, I discovered that long time of genetic drift frequently led to truncation of SP (Fig. S3-1). Without selection pressure, the presence of gene-of-interest was not enforced by the need to produce pIII. Instead, shorter SP was favored because it enabled faster phage genome replication. Therefore, in the following PACE experiment, I removed the mutagenesis-only section, and conducted mutagenesis and selection at the same time (Fig. 3-8A). I set up 4 parallel lagoons, and supplied CRY2-502 in L1 and L2, $CRY2^{H102R}$-502 or $CRY2^{H408P}$-502 in L3 or L4. Phage population persisted in all 4 lagoons for at least 111 hours of selection. Sanger sequencing of bulk phage population revealed 4 more enriched substitutions in nucleotide, 1 resulting in change in amino acid residue ($CRY2^{M167I}$, Fig. 3-8 B), while the other 3 were synonymous substitutions ($CRY2^{G49G}$, $CRY2^{K68K}$, $CRY2^{V455V}$, Fig. 3-8 C-E). The sequences resulting from the 3 synonymous substitutions displayed higher codon usage in E. coli, which could improve the efficiency of protein synthesis (Fig. 3-8 F). All residues corresponding to the 3 non-synonymous substitutions I identified ($CRY2^{H102R}$, $CRY2^{H408P}$, $CRY2^{M167I}$) were mapped at the surface of CRY2 protein (simulated structure, Fig. 3-8 G), suggesting their direct involvement in regulating PPI of CRY2.

3.3.5 Development of the protein-dissociating PACE

The CRY2-BIC1 interaction in darkness was elevated for PACE variants of CRY2, thus decreasing the dynamic range of CRY2-BIC1 blue-light dependent interaction. I designed a protein-dissociating PACE system to evolve and select for disrupted interactions (Fig. 3-9 A and S3-2 A). Activity-independent expression of wild-type gIII is under the control of a phage-shock-promoter (P_{psp}::gIII), which expresses after a phage particle infects the *E. coli* cells [37]. PPI controls the expression of a dominant-negative version of gIII on AP (AP-PPI::gIII-neg) [38], so that strong PPI does not support phage propagation whereas disrupted PPI supports phage propagation.

I first placed the P_{psp}::gIII cassette on a separate complementary plasmid (CP-P_{psp}::gIII, Fig. S3-2 A) and examined if the design specifically support phage propagation in the absence of PPI. Surprisingly, phage replication was not favored regardless of the presence of PPI (Fig. S3-2 B). I hypothesized that the P_{psp}::gIII cassette on CP ($CP_{psp\text{-}gIII}$) was not able to support phage propagation by itself, when it was supposed to be able to do so [39]. Consistent with the hypothesis, the ability of $CP_{psp\text{-}gIII}$ to support phage propagation was more than 10,000-fold less effective in supporting phage propagation than the P_{psp}::gIII cassette placed on AP ($AP_{psp\text{-}gIII}$) routinely used to replicate gIII-deficient phages (Fig. S3-2 C). Since both loss of expression and overexpression of gIII impedes phage infection [40], I next compared the protein abundance of pIII expressed from $CP_{psp\text{-}gIII}$ and $AP_{psp\text{-}gIII}$. The abundance of pIII was estimated by luminescence signals emitted by bacterial luciferases, which was translated in tandem with pIII (Table 3-1). $CP_{psp\text{-}gIII}$ showed 7- to 9-fold more expression of pIII than $AP_{psp\text{-}gIII}$ (Fig. S3-2 D). The origin of replication for CP and AP is ColE1 and

pSC101, with copy number to be 15 [41] and 4-6 [42], respectively. The higher copy number of CP-P_{psp}::gIII cassette probably caused the excessive pIII proteins, which impeded phage infection into *E. coli* cells. To tackle the problem, I placed the P_{psp}::gIII cassette from CP onto the AP-PPI::gIII-neg plasmid (AP-PPI::gIII-neg-P_{psp}::gIII) to achieve stoichiometric expression of gIII-neg and gIII (Fig. 3-9 A). The new AP-PPI::gIII-neg-P_{psp}::gIII plasmid was capable of supporting phage propagation only in the absence of PPI as designed (Fig. 3-9 B).

3.4 Discussion

The directed evolution of CRY2 produced variants with increased interaction affinity with BIC1 in the blue light. However, the increase of interaction affinity was promiscuous, as in previous studies [16], leading to elevated level of background interaction in the dark. A counter-selecting protein-dissociating PACE was developed to remove the background. Further directed evolution efforts could alternate between protein-binding PACE in the blue light and protein-dissociating PACE in darkness to increase dynamic range of the blue-light dependent interactions.

The production of CRY2 in E. coli is another factor that needs to be improved. The 3 synonymous mutations that increased codon usage in *E. coli* suggested that protein synthesis of CRY2 was a limiting factor, and comprehensive codon optimization of CRY2 is needed for efficient production of CRY2. The solubility of CRY2 also need to be further improved for effective evolution experiments. Alternatively, eukaryotic directed evolution systems can be adopted to circumvent the problem.

The PACE variants did not affect the physiological function of CRY2 in inhibition of hypocotyl elongation under blue light 1 μmol m^{-2} s^{-1}. However, their effects on other physiological functions have not be thoroughly studied. Other physiological functions, such as fluence rate responses, dark inactivation, and respond time upon blue light activation, and floral initiation, need to be examined.

3.5 Materials and methods

3.5.1 Bacterial strains

Phage packaging, plaque assays were performed using *E. coli* S2208 cells. PACE experiments, luciferase assays were performed in *E. coli* S2060 cells.

The genotype of S2060 is*: F' proA+B+ Δ(lacIZY) zzf::Tn10 lacI^{Q1} P$_{N25}$-tetR luxCDE P$_{psp}$(AR2) lacZ luxR P$_{lux}$ groESL / endA1 recA1 galE15 galK16 nupG rpsL ΔlacIZYA araD139 Δ(ara,leu)7697 mcrA Δ(mrr-hsdRMS-mcrBC) proBA::pir116 araE201 ΔrpoZ Δflu ΔcsgABCDEFG ΔpgaC λ$^-$* [22].

The genotype of S2208 is: *F' proA+B+ Δ(lacIZY) zzf::Tn10 lacI^{Q1} P$_{N25}$-tetR luxCDE P$_{psp}$(AR2) lacZ luxR P$_{lux}$ groESL / endA1 recA1 galE15 galK16 nupG rpsL ΔlacIZYA araD139 Δ(ara,leu)7697 mcrA Δ(mrr-hsdRMS-mcrBC) proBA::pir116 araE201 ΔrpoZ Δflu ΔcsgABCDEFG ΔpgaC λ$^-$ pJC175e* [22].

3.5.2 General cloning

All high-fidelity PCR reactions for cloning were performed with Phusion® High-Fidelity DNA Polymerase (NEB, M0530S). All vectors were constructed by In-Fusion HD Cloning Kit (Clontech, 639650). All DNA cloning was performed with Stellar

Competent Cells (Clontech, 636763). Information of vectors used in this study is detailed in Table 3-1.

3.5.3 PACE

The devices and parts used for PACE experiments are listed in Table 3-2, if not listed in this section. All PACE experiments were conducted in an incubator (with cooling ability to remove the heat produced by blue light illumination) kept at 37˚C.

To prepare for the media use in PACE experiments, one day before the experiment, Nalgene™ Polypropylene Heavy-Duty Vacuum Carboys (Thermo Fisher Scientific, 2226-0020) containing Davis rich medium [18] (DRM, US Biological, CS050H-001, CS050H-003) supplemented with 50 μg/ml carbenicillin and 40 μg/ml chloramphenicol were pre-warmed in the incubator overnight.

To prepare host culture, at least 2 days before PACE experiments, desired AP and MP were co-transformed into S2060 cells and selected on LB agar (1.2%) with 50 μg/ml carbenicillin, 40 μg/ml chloramphenicol and 100 mM glucose (to suppress the expression of mutagenic proteins from P_{BAD} promoter on MP). One day before PACE experiment, 3 colonies were picked into 5 ml DRM with 50 μg/ml carbenicillin and 40 μg/ml chloramphenicol and mixed well by pipetting. Four 50-fold serial dilution was made in the same DRM, and all tubes were cultured at 37˚C with shaking overnight. The two dilution with OD_{600} closest to 1.0 were mixed, and 4 ml of mixture was used to seed 80 ml DRM in chemostats. One 80-ml chemostat can support 3 parallel 15-ml lagoons (Fisher, 03393D), while usually 2 80-ml chemostats are needed to support 4 parallel 15-ml lagoons. The chemostats were cultured by stirring with magnetic stir bars

at 250 rpm at 37˚C for 2 hours, then fresh media began to be introduced into chemostats at flow rate of 100 ml/h. Host cultures were flowed to lagoons that were not seeded with phages at flow rate of 15 ml/h. Meanwhile, 10% (w/v) arabinose (GoldBio, A300-100) was delivered to lagoons by a syringe pump (New Era Pump Systems, Inc., NE-1600) at flow rate of 0.7-1 ml/h throughout all PACE experiments. This step was to pre-induce the MP of the *E. coli* cells.

To prepare phage samples, phages were propagated on S2208 cultures in 2xYT media (1.6% peptone; 1% yeast extract; 0.5% NaCl) with 50 μg/ml carbenicillin overnight and then plated to isolate single plaques. Three single plaques were inoculated in 2xYT media with 50 μg/ml carbenicillin for 8 hours, centrifuged at top speed, filtered with 0.22 μm syringe filter (Genesee Scientific, 25-243), and plaques were counted by plaque assays to determine the titer of the phage samples. The phage samples were temporarily stored at 4˚C.

To begin the evolution of SP, lagoons with volumes fixed at 15-17 ml were seeded with filtered phage (to guarantee the removal of S2208 cells) to reach final titer of 10^3 pfu/ml (if genetic drift was performed) or 10^8 pfu/ml (if genetic drift was not performed). The lagoons were constantly diluted with host cell cultures at flow rates indicated on figures. In all cases, volumes of chemostats and lagoons were kept by setting needles at the location of the desired volume.

3.5.4 Plaque assays

S2208 cells were grown at 37˚C with shaking overnight in LB media (Fisher, 244610) containing 50 μg/ml carbenicillin, re-inoculated in 1:1,000 dilution to fresh LB

media with 50 μg/ml carbenicillin (LabScientific, 1525), grown at 37˚C with shaking for more than 5 hours. Phage samples were centrifuged at top speed for 5 minutes and filtered with 0.22 μm syringe filters. The filtered samples were serially diluted in 10- or 100-fold (depending on expected titer) increments in the following steps: 1) add 100 μl of S2208 cells into each of 4 tubes; 2) add 1 or 10 μl filtered phage samples into the first tube and mix well by vertexing; 3) remove 1 or 10 μl of mixture from the first tube and move into the second tube and mix well with the S2208 cells by vertexing; 4) repeat 3) until phage samples are diluted in the fourth tube. Then 900 μl of LB agar (0.6-0.7%) pre-warmed at 55-65˚C was added into each tube and mix well by pipetting, and plate onto 4-sector LB agar (1.2%) plates (VitaScientific, 753031). The plates were incubated at 37˚C overnight before plaques were counted.

3.5.5 Luciferase assays

CPs were co-transformed with corresponding APs into S2060 and selected on LB agar (1.2%) with 50 μg/ml carbenicillin and 50 μg/ml spectinomycin (Sigma, S9007). Three colonies were mixed and cultured overnight at 37˚C with shaking, then diluted in 700 μl of DRM with 50 μg/ml carbenicillin, 50 μg/ml spectinomycin and 2 μM arabinose for 5 hours. The culture was dispensed into 96-well white wall, clear bottom plates (Fisher scientific, 07-000-098) and OD_{600} and luminescence was measured by Tecan Infinite M1000 Pro. The plates were shaken for 5 seconds with 1 mm diameter, and then luminescence was measured for 5 seconds for each well.

3.5.6 Phage propagation and enrichment assay

Three colonies of *E. coli* cells containing corresponding APs were grown in LB with 50 μg/ml carbenicillin at 37˚C with shaking overnight. The cultures were diluted by 1,000-fold and grown in 2xYT media with 50 μg/ml carbenicillin for 5 hours, before being infected by filtered phage samples to reach final titer of 10^5 pfu/ml. The mixture of phages and *E. coli* cells were grown at 37˚C with shaking overnight, and then centrifuged and filtered for plaque assays to count phage titer.

3.5.7 PCR analysis of phage plaques

Phage samples were taken from lagoons using the needles for pumping out waste cultures, filtered, and plated to check phage titer and isolate single plaques. PCR was performed using DreamTaq DNA polymerase (Thermo Fisher Scientific, FEREP0703). Primers were: gVIII-197F (CGATCCCGCAAAAGCGGCCT) and gIII-1208R (AAAAGAAACGCAAAGACACC). To provide templates for PCR, phage plaques were briefly touched with pipette tips, which were then dipped into PCR reactions.

3.5.8 Sanger sequencing of CRY2 variants

Overnight culture of S2208 cells were re-inoculated into 2xYT with 50 μg/ml carbenicillin by 1:1,000 dilution, cultured at 37˚C with shaking for 5 hours, then mixed with 10 μl of filtered PACE phage samples, cultured with shaking at 37˚C for ≥ 5 hours. Phage circular dsDNA was isolated using GeneJET Plasmid Miniprep Kit (Thermo Fisher Scientific, K0503), PCR amplified with Phusion® High-Fidelity DNA Polymerase using the following primers: gVIII-80F (CGATCCCGCAAAAGCGGCCT) and gVI-20R

(ATACCCAAAAGAACTGGCAT). The PCR products were purified with GeneJET Gel Extraction Kit (Thermo Fisher Scientific, K0692), then sequenced with primer gVIII-197F (CGATCCCGCAAAAGCGGCCT). The variants were identified using Mutation Surveyor (SoftGenetics).

3.5.9 HEK293T cell culture, transfection, protein expression, co-immunoprecipitation, luciferase assay for protein thermal stability, and immunoblot assays

HEK293T cells were routinely cultured in Dulecco's modified Eagle's medium (DMEM, Fisher, MT10017CM) supplemented with 10% (v/v) FBS (fetal bovine serum, Gemini Bio-products, 900-108), 100 IU penicillin and 100 mg/L streptomycin (PS antibiotics, Corning, 30-002-CI), in humidified 5% (v/v) CO_2 incubator (Thermo Fisher Scientific, MIDI 40, Model 3403) at 37˚C. For protein expression assays, HEK293T cells were seeded at a density of $3x10^5$ cells per well of a 6-well plate (Fisher, 130184) in 2 ml DMEM with FBS and PS antibiotics and grown for 16-20 hours. Cells were transfected using polyethylenimine (PEI, Polysciences Inc., 23966-2). At least thirty minutes before transfection, media was changed into fresh media pre-warmed at 37˚C. For detection of CRY2-BIC1 interaction, each well was transfected with 400 ng of pQCMV-F-CRY2 and 200 ng of pQCMV-EGFP-BIC1; for CRY2-CRY2 interaction, 300 ng of pQCMV-F-CRY2 and pCMV-M-CRY2; for SPA1-CRY2 interaction, 300 ng of pCMV-M-CRY2 and 900 ng of pQCMV-F-SPA1; for CIB1-CRY2 interaction, 300 ng of pCMV-M-CRY2 and 900 ng of pQCMV-EGFP-CIB1. For each well of transfection, PEI (1 μg/μl, pH 7.2) was added into 200 μl of DMEM with PI antibiotics but without FBS,

mixed with plasmids in a ratio of 4:1 (µl PEI:µg plasmid), vortexed for a few seconds, briefly centrifuged, and left at room temperature for 15-30 minutes. The mixture was then applied to cells directly. After 36-48 hours after transfection, cells were kept in darkness or exposed to blue light of 100 µmol m^{-2} s^{-1} for 2 hours at 22˚C before removing media by aspiration, being washed by PBS (phosphate-buffered saline, 137 mM NaCl, 2.7 mM KCl, 10 mM Na_2HCO_3, 1.8 mM K_2HCO_3) solution and harvested by 700 µl of 1% Brij buffer (1% Brij-35, 50 mM Tris-HCl pH 8.0, 150 mM NaCl, 1mM EDTA) supplemented with 1x Protease inhibitor cocktail (Sigma, 1187358001). The cells were kept on ice in darkness for 30 minutes for complete lysis, centrifuged at 14,000 g for 10 minutes at 4˚C. Fifty µl of supernatant was removed into new tubes, mixed with 1/3 volume of 4x SDS (sodium dodecyl sulfate–polyacrylamide, Calbiochem, 4280) sample buffer (40% v/v glycerol; 0.04% w/v bromphenol blue; 8% w/v SDS; 0.4 M Tris-HCl pH 6.8; 400 mM DL-dithiothreitol), boiled at 100˚C for 5 minutes, and saved as inputs. The remaining supernatant was removed to new tubes, and EZview™ Red ANTI-FLAG® M2 Affinity Gel (Sigma, F2426) or GFP-Trap Agarose (chromotek, gta-10) affinity agarose beads were added according to manufactures instruction. The supernatant and beads were incubated at 4˚C in darkness with gentle rocking for 2 hours. After incubation, the beads were washed with 1% Brij buffer for three times and proteins were eluted from beads by adding 80 µl of 1x SDS sample buffer (diluted from 4x SDS sample buffer), boiled at 100˚C for 5 minutes, and then analyzed by horseradish peroxidase (HRP) or fluorescent immunoblot.

For luciferase assays examining rate of protein degradation, HEK293T cells were similarly cultured as above, and 400 ng of pQFLUC2-CRY2 was transfected. Thirty-six

to 48 hours after transfection, media was removed by aspiration. Cells in each well were washed with 1 ml PBS, then harvested with 2 ml DMEM w/o phenol red (Thermo Fisher Scientific, 21063029) supplemented with 10% FBS and PS antibiotics. Cells in each well were transferred to two 1.7 ml tubes, 700 µl in each tube, one supplemented cycloheximide (CHX, Sigma, C-7698, dissolved in dimethyl sulfoxide, DMSO) to reach final concentration of 100 µg/ml, the other the same volume of DMSO, both supplemented luciferin (GoldBio, LUCK) to reach final concentration of 1 mM, then mixed well each tube by pipetting. The cells in each tube were dispensed into 3 wells of 96-well tissue culture plates (Fisher, 07200566), 200 µl per well. Luminescence was measured 10 seconds every 10 minutes, for 60 times (10 hours) using luminometer (Turner BioSystems, Modulus microplate).

For immunoblots, protein extract supernatant from HEK293T cells were made as described above. Plant materials were dipped into liquid N_2 and homogenized by TissueLyser (QIAGEN). The resulting plant tissue powders were added 0.8x volume of powder of protein extraction buffer (120 mM Tris-HCl pH 6.8; 100 mM EDTA pH 8.0; 4% w/v SDS; 10% v/v 2-Mercaptoethanol; 5% Glycerol and 0.01% Bromophonol Blue), boiled for 8 minutes, and then centrifuged with table top centrifuges at top speed for 10 minutes. Protein extract supernatant from HEK293T cells or plant samples was separated by home-made 10% SDS-PAGE (sodium dodecyl sulfate–polyacrylamide gel electrophoresis) gels, transferred to Pure Nitrocellulose Blotting Membranes (BioTrace NT, Pall Life Science, 66593) using wet electroblotting system (Bio-Red, Hercules). Ponceau S Red solution (0.1% w/v Ponceau S; 5% v/v acetic acid) was used to stain transferred membranes to gauge transferring efficiency. For immunoblot signals

captured by the Odyssey® CLx Infrared Imaging System (LI-COR), membranes were blocked with 0.5% casein in PBS solution, blotted with primary antibodies in 0.5% casein in PBST (PBS supplemented with 0.3% of Tween-20) solution, and then blotted with fluorescent secondary antibodies in 0.5% casein in PBST solution. Images captured by Odyssey® System were processed with the free Image Studio Lite software (LI-COR) and organized with Adobe Photoshop CC 2017. For immunoblot signals captured by exposure to X-ray film, membranes were blocked with 5% skimmed milk (Lab Scientific, M-0841) in PBST solution, blotted with primary antibodies in PBST solution, then blotted with secondary antibodies in PBST solution. The membranes were then incubated in the home made ECL solution (Solution A: 100 mM Tris-HCl pH8.5; 0.2 mM coumaric acid; Solution B: 100 mM Tris-HCL pH 8.5; 1.25 mM luminol; Right before use, mix 3 ml Solution A with 3 ml solution B and add 2 µl 30% H_2O_2) and exposed to X-ray films. Primary antibodies used in this study were: rabbit-anti-CRY2 (1:3,000, home-made) [43], rabbit-anti-FLAG (1:3,000, Sigma, F7425), mouse-anti-MYC (1:3,000, Millipore, 05-724), rabbit-anti-GFP (Santa Cruz Biotechnology, sc-8334). Secondary antibodies were: goat-anti-rabbit IgG conjugated to Alexa Fluro® 790 (1:15,000, Thermo Fisher Scientific, A11369), goat-anti-mouse IgG conjugated to Alexa Fluro® 790 (1: 15,000, Thermo Fisher Scientific, A11357), donkey-anti-rabbit IgG conjugated to Alexa Fluro® 680 (Thermo Fisher Scientific, A10043), donkey-anti-mouse IgG conjugated to Alexa Fluro® 680 (Thermo Fisher Scientific, A10038), donkey-anti-rabbit IgG conjugated to HRP (1:10,000, Fisher, 45000682) and sheep-anti-mouse IgG conjugated to HRP (1:10,000, Fisher, 45000692).

To quantify protein abundance of CRY2, fluorescent signals captured by Odyssey® CLx System (LI-COR) were quantified by an internal method of Image Studio Lite software (LI-COR). The resulting signals were used to calculate protein abundance by the following formula:

$$\text{Protein abundance} = \frac{\text{CRY2}_{\text{mt}}/\text{ACTIN}_{\text{mt}} - \text{CRY2}_{cry1cry2}/\text{ACTIN}_{cry1cry2}}{\text{CRY2}_{\text{CRY2max}}/\text{ACTIN}_{\text{CRY2max}} - \text{CRY2}_{cry1cry2}/\text{ACTIN}_{cry1cry2}}$$

where CRY2max is the right most wild-type CRY2 transgenic plant on the standard curve (Fig. 3-6).

3.5.10 Plant materials, growth conditions and analysis of hypocotyl inhibition

All *Arabidopsis* plant lines used in this study were in Columbia (Col) background. The wild-type plants used in this study were Col-4. The *cry1cry2* lines were acquired by crossing *cry1-304* [44] and *cry2-1* [43]. The Ti plasmid pDT1h was modified from pDT1 [11] by replacing BlpR (phosphinothricin acetyltransferase) with HygR (hygromycin B phosphotransferase). CRY2 variants were cloned into the BamHI site of pDTh1. The resulting constructs were P_{ACTIN2}:: 4xMYC-CRY2::T_{CaMV} and were introduced into the *cry1cry2* plants by standard floral dip method [45]. The transgenic T1, T2 and T3 populations were screed on MS (Murashige and Skoog) agar (MS salt, Caisson, MSP01-50LT, 0.7% plant agar, Fisher Scientific, A20300-100.0) and maintained on compound soil.

For hypocotyl inhibition assays, seeds were sterilized and sown onto fresh-made MS agar plates, subjected to 4˚C cold treatment in darkness for 4 days, exposed to white light at room temperature for 24 hours, and then put into blue light (1 μmol m^{-2} s^{-1}) conditions at room temperature for 7 days. The resulting seedlings were sandwiched

between two plastic sheets (one transparent, the other black), scanned and measured by Fiji (NIH). Relative specific activity was calculated using the following formula:

$$Relative\ specific\ activity = 1 - \frac{hypocotyl\ length_{mt} - length_{wt}}{length_{cry1cry2} - length_{wt}}$$

Light-emitting diode (LED) was used to obtain monochromatic blue light (peak 450 nm; half-bandwidth of 20 mm).

3.5.11 Structure simulation

The CRY2 structure was simulated using SWISS-MODEL [46] from CRY2 full-length protein sequence based on crystal structure of CRY1 (PDB: 1U3C) [36].

3.6 References

1. Hughes, R.M., Bolger, S., Tapadia, H., and Tucker, C.L. (2012). Light-mediated control of DNA transcription in yeast. Methods 58, 385–391.
2. Konermann, S., Brigham, M.D., Trevino, A.E., Hsu, P.D., Heidenreich, M., Cong, L., Platt, R.J., Scott, D.A., Church, G.M., and Zhang, F. (2013). Optical control of mammalian endogenous transcription and epigenetic states. Nature 500, 472–476.
3. Boulina, M., Samarajeewa, H., Baker, J.D., Kim, M.D., and Chiba, A. (2013). Live imaging of multicolor-labeled cells in Drosophila. Development 140, 1605–1613.
4. Taslimi, A., Zoltowski, B., Miranda, J.G., Pathak, G.P., Hughes, R.M., and Tucker, C.L. (2016). Optimized second-generation CRY2–CIB dimerizers and photoactivatable Cre recombinase. Nat. Chem. Biol., 1–8.
5. Idevall-Hagren, O., Dickson, E.J., Hille, B., Toomre, D.K., and De Camilli, P. (2012). Optogenetic control of phosphoinositide metabolism. Proc. Natl. Acad. Sci. U. S. A. 109, E2316–E2323.
6. Giordano, F., Saheki, Y., Idevall-Hagren, O., Colombo, S.F., Pirruccello, M., Milosevic, I., Gracheva, E.O., Bagriantsev, S.N., Borgese, N., and De Camilli, P. (2013). XPI(4,5)P2-Dependent and Ca2+-Regulated ER-PM interactions mediated by the extended synaptotagmins. Cell 153, 1494–1509.
7. Aoki, K., Kumagai, Y., Sakurai, A., Komatsu, N., Fujita, Y., Shionyu, C., and Matsuda, M. (2013). Stochastic ERK activation induced by noise and cell-to-cell propagation regulates cell density-dependent proliferation. Mol. Cell 52, 529–540.
8. O'Neill, P.R., and Gautam, N. (2014). Subcellular optogenetic inhibition of G

proteins generates signaling gradients and cell migration. Mol. Biol. Cell 25, 2305–2314.

9. Maiuri, P., Rupprecht, J.F., Wieser, S., Ruprecht, V., Bénichou, O., Carpi, N., Coppey, M., De Beco, S., Gov, N., Heisenberg, C.P., Lage Crespo, C., Lautenschlaeger, F., Le Berre, M., Lennon-Dumenil, A.M., Raab, M., Thiam, H.R., Piel, M., Sixt, M., and Voituriez, R. (2015). Actin flows mediate a universal coupling between cell speed and cell persistence. Cell 161, 374–386.

10. Duan, L., Che, D., Zhang, K., Ong, Q., Guo, S., and Cui, B. (2015). Optogenetic control of molecular motors and organelle distributions in cells. Chem. Biol. 22, 671–682.

11. Wang, Q., Zuo, Z., Wang, X., Gu, L., Yoshizumi, T., Yang, Z., Yang, L., Liu, Q., Liu, W., Han, Y.-J., Kim, J.-I., Liu, B., Wohlschlegel, J.A., Matsui, M., Oka, Y., and Lin, C. (2016). Photoactivation and inactivation of Arabidopsis cryptochrome 2. Science 354, 343–347.

12. Zhang, K., and Cui, B. (2015). Optogenetic control of intracellular signaling pathways. Trends Biotechnol. 33, 92–100.

13. Wend, S., Wagner, H.J., Muller, K., Zurbriggen, M.D., Weber, W., and Radziwill, G. (2014). Optogenetic control of protein kinase activity in mammalian cells. ACS Synth. Biol. 3, 280–285.

14. Taslimi, A., Vrana, J.D., Chen, D., Borinskaya, S., Mayer, B.J., Kennedy, M.J., and Tucker, C.L. (2014). An optimized optogenetic clustering tool for probing protein interaction and function. Nat. Commun. 5, 4925.

15. Duan, L., Hope, J., Ong, Q., Lou, H.Y., Kim, N., McCarthy, C., Acero, V., Lin,

M.Z., and Cui, B. (2017). Understanding CRY2 interactions for optical control of intracellular signaling. Nat. Commun. 8.

16. Esvelt, K.M., Carlson, J.C., and Liu, D.R. (2011). A system for the continuous directed evolution of biomolecules. Nature 472, 499–503.

17. Dickinson, B.C., Leconte, A.M., Allen, B., Esvelt, K.M., and Liu, D.R. (2013). Experimental interrogation of the path dependence and stochasticity of protein evolution using phage-assisted continuous evolution. Proc. Natl. Acad. Sci. U. S. A. 110, 9007–9012.

18. Carlson, J.C., Badran, A.H., Guggiana-Nilo, D.A., and Liu, D.R. (2014). Negative selection and stringency modulation in phage-assisted continuous evolution. Nat. Chem. Biol. 10, 216–22.

19. Pu, J., Chronis, I., Ahn, D., and Dickinson, B.C. (2015). A Panel of Protease-Responsive RNA Polymerases Respond to Biochemical Signals by Production of Defined RNA Outputs in Live Cells. J. Am. Chem. Soc. 137, 15996–15999.

20. Dickinson, B.C., Packer, M.S., Badran, A.H., and Liu, D.R. (2014). A system for the continuous directed evolution of proteases rapidly reveals drug-resistance mutations. Nat. Commun. 5, 5352.

21. Packer, M.S., Rees, H.A., and Liu, D.R. (2017). Phage-assisted continuous evolution of proteases with altered substrate specificity. Nat. Commun. 8, 956.

22. Hubbard, B.P., Badran, A.H., Zuris, J.A., Guilinger, J.P., Davis, K.M., Chen, L., Tsai, S.Q., Sander, J.D., Joung, J.K., and Liu, D.R. (2015). Continuous directed evolution of DNA-binding proteins to improve TALEN specificity. Nat. Methods 12, 939–942.

23. Bryson, D.I., Fan, C., Guo, L.T., Miller, C., Söll, D., and Liu, D.R. (2017). Continuous directed evolution of aminoacyl-tRNA synthetases. Nat. Chem. Biol. 13, 1253–1260.

24. Hu, J.H., Miller, S.M., Geurts, M.H., Tang, W., Chen, L., Sun, N., Zeina, C.M., Gao, X., Rees, H.A., Lin, Z., and Liu, D.R. (2018). Evolved Cas9 variants with broad PAM compatibility and high DNA specificity. Nature 556, 57–63.

25. Wang, T., Badran, A.H., Huang, T.P., and Liu, D.R. (2018). Continuous directed evolution of proteins with improved soluble expression. Nat. Chem. Biol. 14, 972–980.

26. Thuronyi, B.W., Koblan, L.W., Levy, J.M., Yeh, W.-H., Zheng, C., Newby, G.A., Wilson, C., Bhaumik, M., Shubina-Oleinik, O., Holt, J.R., and Liu, D.R. (2019). Continuous evolution of base editors with expanded target compatibility and improved activity. Nat. Biotechnol. 37, 1070–1079.

27. Pu, J., Zinkus-Boltz, J., and Dickinson, B.C. (2017). Evolution of a split RNA polymerase as a versatile biosensor platform. Nat. Chem. Biol. 13, 432–438.

28. Pu, J., Kentala, K., and Dickinson, B.C. (2018). Multidimensional Control of Cas9 by Evolved RNA Polymerase-Based Biosensors. ACS Chem. Biol. 13, 431–437.

29. Pu, J., Disare, M., and Dickinson, B.C. (2019). Evolution of C-Terminal Modification Tolerance in Full-Length and Split T7 RNA Polymerase Biosensors. ChemBioChem.

30. Jones, K.A., Kentala, K., Beck, M.W., An, W., Lippert, A.R., Lewis, J.C., and Dickinson, B.C. (2019). Development of a Split Esterase for Protein–Protein Interaction-Dependent Small-Molecule Activation. ACS Cent. Sci.

31. Badran, A.H., Guzov, V.M., Huai, Q., Kemp, M.M., Vishwanath, P., Kain, W., Nance, A.M., Evdokimov, A., Moshiri, F., Turner, K.H., Wang, P., Malvar, T., and Liu, D.R. (2016). Continuous evolution of Bacillus thuringiensis toxins overcomes insect resistance. Nature 533, 58–63.

32. Dove, S.L., and Hochschild, A. (1998). Conversion of the ω subunit of Escherichia coli RNA polymerase into a transcriptional activator or an activation target. Genes Dev. 12, 745–754.

33. Wojcik, J., Hantschel, O., Grebien, F., Kaupe, I., Bennett, K.L., Barkinge, J., Jones, R.B., Koide, A., Superti-Furga, G., and Koide, S. (2010). A potent and highly specific FN3 monobody inhibitor of the Abl SH2 domain. Nat. Struct. Mol. Biol. 17, 519–527.

34. Rajagopala, S. V., Sikorski, P., Kumar, A., Mosca, R., Vlasblom, J., Arnold, R., Franca-Koh, J., Pakala, S.B., Phanse, S., Ceol, A., Häuser, R., Siszler, G., Wuchty, S., Emili, A., Babu, M., Aloy, P., Pieper, R., and Uetz, P. (2014). The binary protein-protein interaction landscape of escherichia coli. Nat. Biotechnol. 32, 285–290.

35. Badran, A.H., and Liu, D.R. (2015). Development of potent in vivo mutagenesis plasmids with broad mutational spectra. Nat. Commun. 6, 8425.

36. Brautigam, C.A., Smith, B.S., Ma, Z., Palnitkar, M., Tomchick, D.R., Machius, M., and Deisenhofer, J. (2004). Structure of the photolyase-like domain of cryptochrome 1 from Arabidopsis thaliana. Proc. Natl. Acad. Sci. 101, 12142–12147.

37. Brissette, J.L., Weiner, L., Ripmaster, T.L., and Model, P. (1991). Characterization

and sequence of the Escherichia coli stress-induced psp operon. J. Mol. Biol. 220, 35–48.

38. Bennett, N.J., and Rakonjac, J. (2006). Unlocking of the filamentous bacteriophage virion during infection is mediated by the C domain of pIII. J. Mol. Biol. 356, 266–273.

39. Rakonjac, J., Jovanovic, G., and Model, P. (1997). Filamentous phage infection-mediated gene expression: Construction and propagation of the gill deletion mutant helper phage R408d3. Gene 198, 99–103.

40. Boeke, J.D., Model, P., and Zinder, N.D. (1982). Effects of bacteriophage f1 gene III protein on the host cell membrane. MGG Mol. Gen. Genet. 186, 185–192.

41. Twigg, A.J., and Sherratt, D. (1980). Trans-complementable copy-number mutants of plasmid ColE1. Nature 283, 216–218.

42. Hashimoto-Gotoh, T., Franklin, F.C.H., Nordheim, A., and Timmis, K.N. (1981). Specific-purpose plasmid cloning vectors I. Low copy number, temperature-sensitive, mobilization-defective pSC101-derived containment vectors. Gene 16, 227–235.

43. Guo, H., Yang, H., Mockler, T.C., and Lin, C. (1998). Regulation of flowering time by Arabidopsis photoreceptors. Science 279, 1360–1363.

44. Mockler, T.C., Guo, H., Yang, H., Duong, H., and Lin, C. (1999). Antagonistic actions of Arabidopsis cryptochromes and phytochrome B in the regulation of floral induction. Development 126, 2073–2082.

45. Clough, S.J. (2005). Floral dip: Agrobacterium-mediated germ line transformation. Transgenic Plants, 091–102.

46. Waterhouse, A., Bertoni, M., Bienert, S., Studer, G., Tauriello, G., Gumienny, R., Heer, F.T., de Beer, T.A.P., Rempfer, C., Bordoli, L., Lepore, R., and Schwede, T. (2018). SWISS-MODEL: homology modelling of protein structures and complexes. Nucleic Acids Res. 46, W296–W303.

3.7 Figures

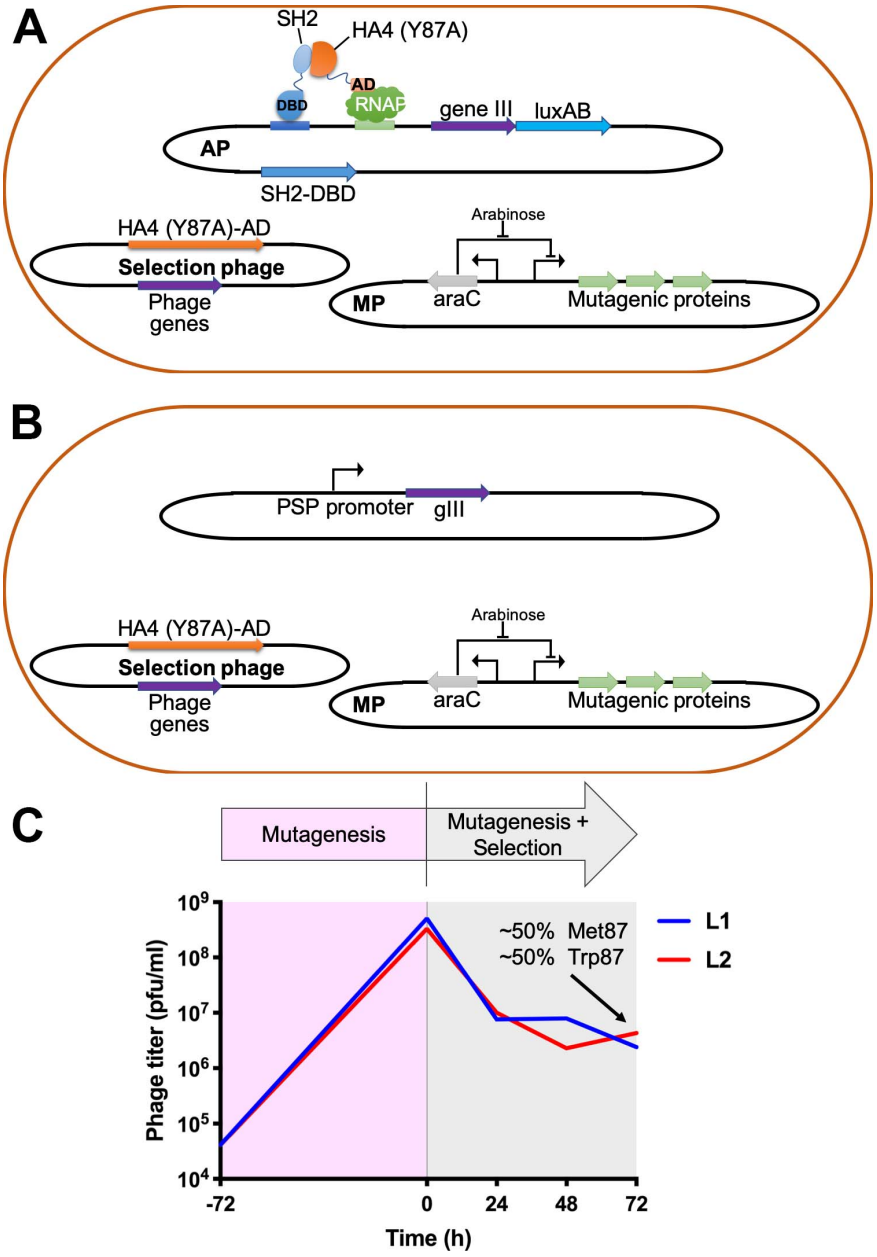

Fig. 3-1: The establishment of the protein-binding PACE.

A) Establishment of the PACE for evolution of stronger protein-protein interactions. The strong interaction between SH2 domain of ABK1 kinase and HA4 monobody is disrupted by the HA4^{Y87A} mutant. To validate the protein-binding PACE, HA4^{Y87A} was evolved to restore the SH2-HA4 interaction.

B) Genetic drift of the $HA4^{Y87A}$ mutant library. Accessary plasmids (AP) containing the PPI selection reporter was replaced by plasmids expressing gIII under the control of a phage shock promoter (psp promoter), which expresses after M13 phage infects cells, therefore enabling selection phage (SP) replication without selection.

C) Continuous evolution of $HA4^{Y87A}$ variants that bind SH2. PACE was executed in 2 segments. In the first segment (-72 h – 0), mutagenesis was applied without selection to generate a diversified library. In the second segment (0 – 72 h), both mutagenesis and selection were applied. PACE was conducted in 2 parallel lagoons: L1 and L2. After 72 h of selection, phage was sampled from both L1 and L2 and sequences. All $HA4^{Y87A}$ was changed into $HA4^{87W}$ or $HA4^{87M}$.

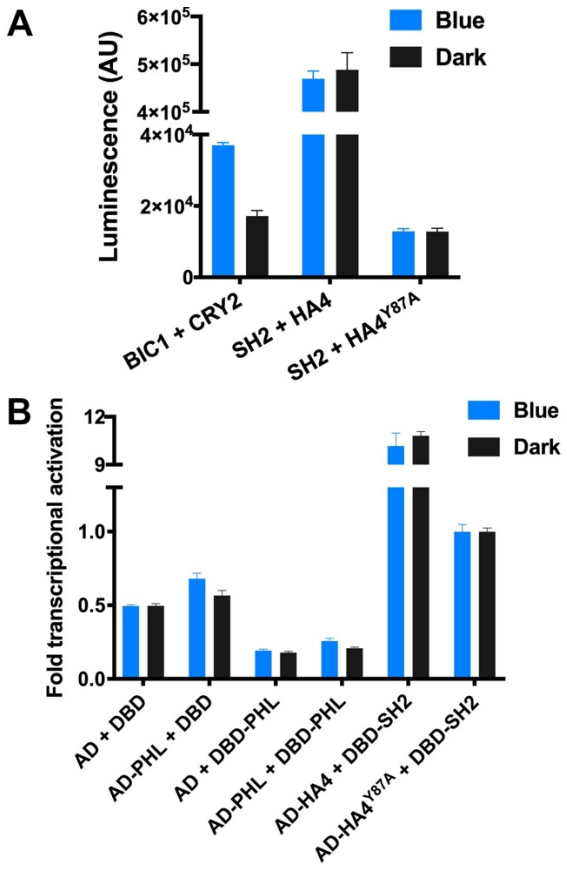

Fig. 3-2: Validation and characterization of CRY/PHL family protein-protein interaction using the bacterial two-hybrid scheme.

A) Blue-light specific interaction between CRY2 and BIC1 in *E. coli* cells. BIC1 and SH2 proteins were fused with DNA binding domain, while CRY2, HA4 and HA4^{Y87A} proteins were fused with transcriptional activation domain. The SH2-HA4 interaction served as positive control, while the SH2-HA4^{Y87A} interaction served as negative control. See Table 3-1 and Material and Methods for detailed information.

B) *E. coli* PHL suppressed transcription in bacterial two hybrid system. AD: transcriptional activation domain; DBD: DNA-binding domain.

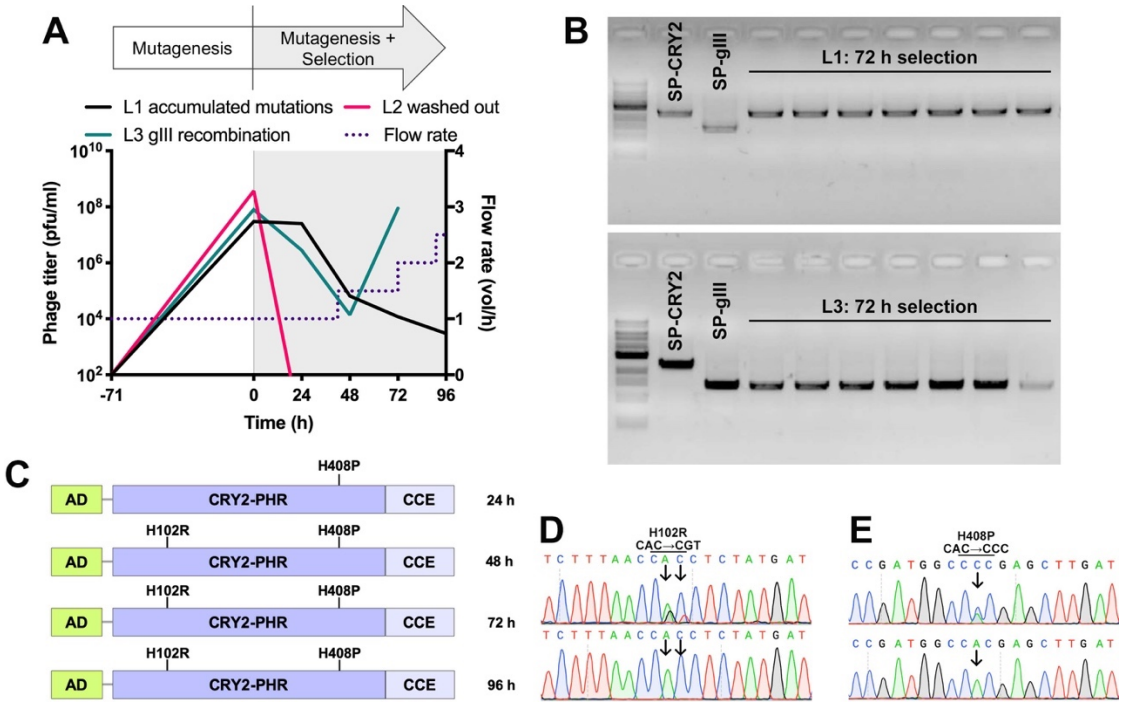

Fig. 3-3: Continuous evolution of CRY2 variants that bind BIC1 in the blue light.

A) Continuous evolution of CRY2 in two segments as in Fig. 3-1C. PACE was conducted in 3 parallel lagoons: L1, L2 and L3, and the result of evolution in each lagoon was labeled. Phage was sampled every 24 hours starting from selection at time 0. The flow rate of *E. coli* cells in lagoon was labeled by dotted line, using lagoon volume per hour as unit.

B) Characterization and identification of individual phage plaques in L1 and L3 at 72 hours of selection. Phage sampled at 72-hour of selection was plated on *E. coli* cells. The length of the gene-of-interest DNA insertion of individual plaques was analyzed by PCR using the following primers: gVIII-197F, GAAATTCACCTCGAAAGCAA and gVI-20R, ATACCCAAAAGAACTGGCAT. Selection phage with CRY2 insertion (SP-CRY2) or wild type M13 phage (SP-gIII) served as control.

C) The positions of the enriched mutations of CRY2 selection phage in L1, sampled at different selection time point as indicated. Length of each domain was in proportion to the actual length of amino acid residues. AD: transcriptional activation domain; CRY2-PHR: the photolyase homologous region of CRY2; CCE: CRY-C terminal extension. **D-E)** Sanger sequencing chromatographs showing enriched mutations of $CRY2^{H102R}$ and $CRY2^{H408P}$.

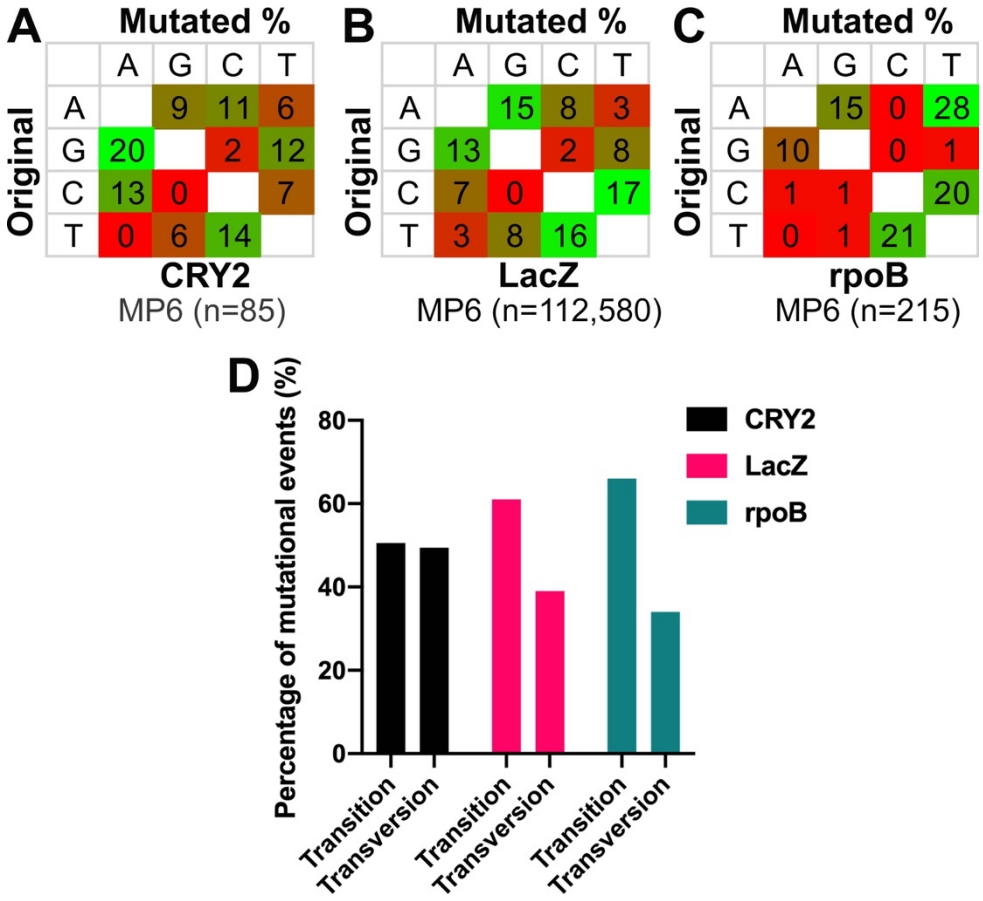

Fig. 3-4: Comparison of the mutagenic spectra of CRY2 with previous reports.

A) The mutagenic spectra of CRY2 evolved using MP6 (see Table 3-1 for details) in protein-binding PACE. Individual plaques were PCR amplified and sequenced. The number of unique mutational events were recorded.

B-C) Previously published mutagenic spectra of LacZ (B) and ropZ (C), modified from [35].

D) Comparison of percentage of mutagenic events of CRY2, LacZ and ropZ.

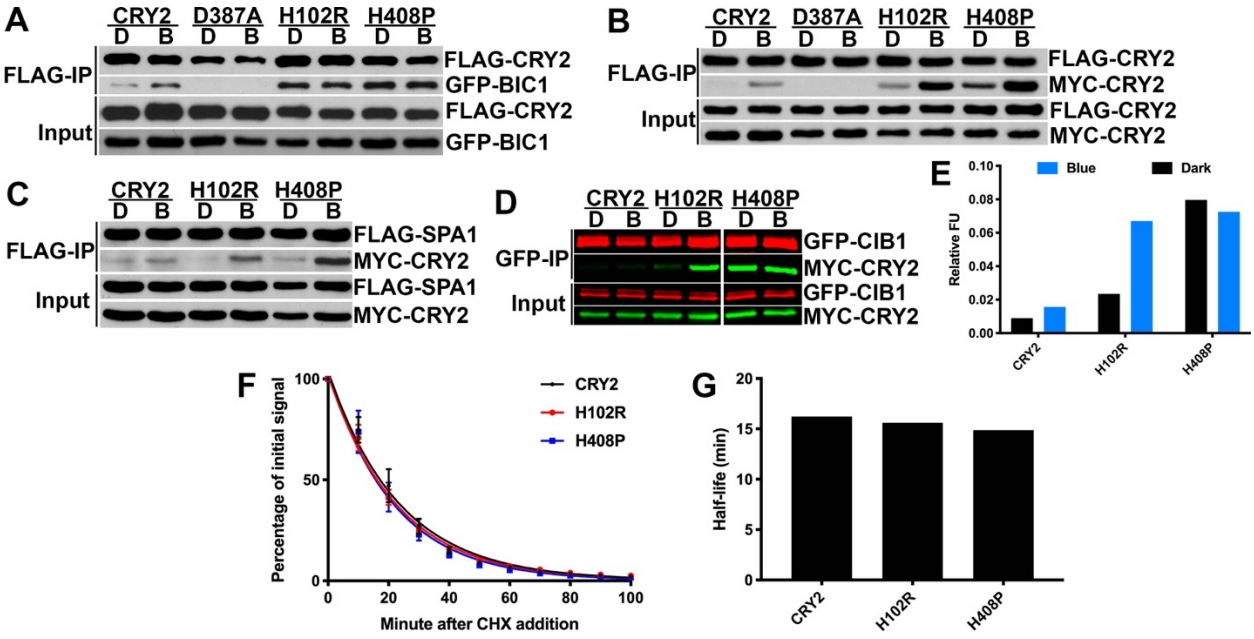

Fig. 3-5: Universally increased protein interaction of the CRY2 PACE variants.

A-D) Co-immunoprecipitation (co-IP) of CRY2 PACE variants with BIC1 (A), CRY2 (B), SPA1 (C) and CIB1 (D). Proteins fused with the indicated epitopes were expressed in HEK293T cells, cultured in darkness or illuminated with blue light (100 μmol m^{-2} s^{-1}) for 2 hours, and immunoprecipitated with FLAG or GFP-Trap beads (see Materials and Methods for details). A-C was visualized using secondary antibodies coupled with HRP, while D was visualized using secondary antibodies coupled with fluorescent molecules.

E) Quantification of co-IP signals of D. Relative fluorescent unit (FU) was calculated by dividing co-IP signals (MYC-CRY2 in GFP-IP) by IP signals (GFP-CIB1 in GFP-IP).

F) Degradation of CRY2 and its PACE variants. CRY2-Luciferase proteins were expressed in HEK293T cells. Cycloheximide (CHX) that inhibits translation and the substrate of firefly luciferase, luciferin was added shortly after the cells were lysed, and luminescence was measured (n=3, see Materials and Methods for details).

G) Half-life of degradation calculated from curves in F.

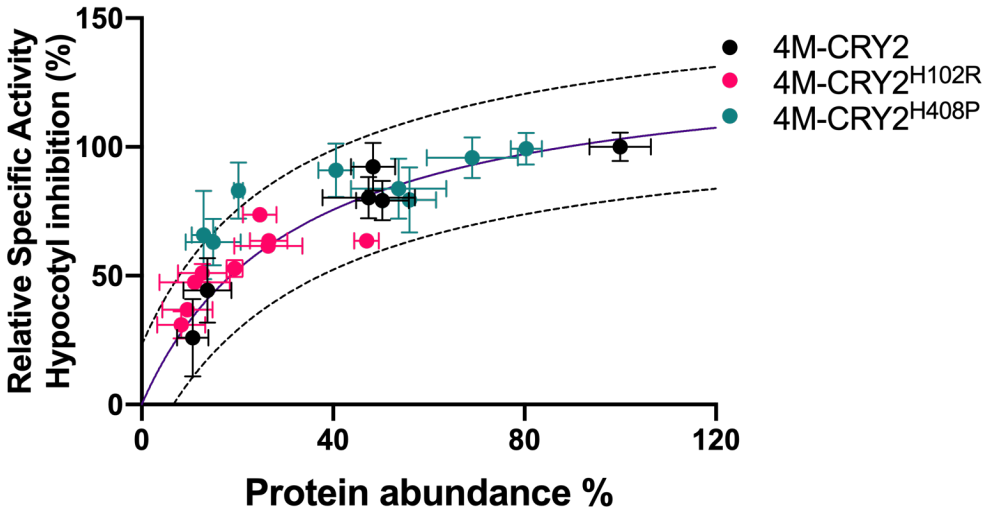

Fig. 3-6: Analysis of the blue-light induced hypocotyl inhibition activities of PACE variants of CRY2.

Standard curve concerning protein abundance and relative specific activity in hypocotyl inhibition was constructed using 6 independent lines of *cry1cry2* transgenic plants expressing 4xMYC fused with wt-CRY2 at varying protein abundance. Dashed lines represent 95% of prediction bands of the standard curve. Regions between the 95% prediction bands cover the likely location of 95% of additional data points. Individual data points of PACE variants of CRY2 represent independent transgenic *cry1cry2* plants expressing each variant. Calculation of protein abundance and relative specific activities was detailed in Materials and Methods. Both horizontal (n≥3) and vertical (n≥20) error bars represent standard deviation.

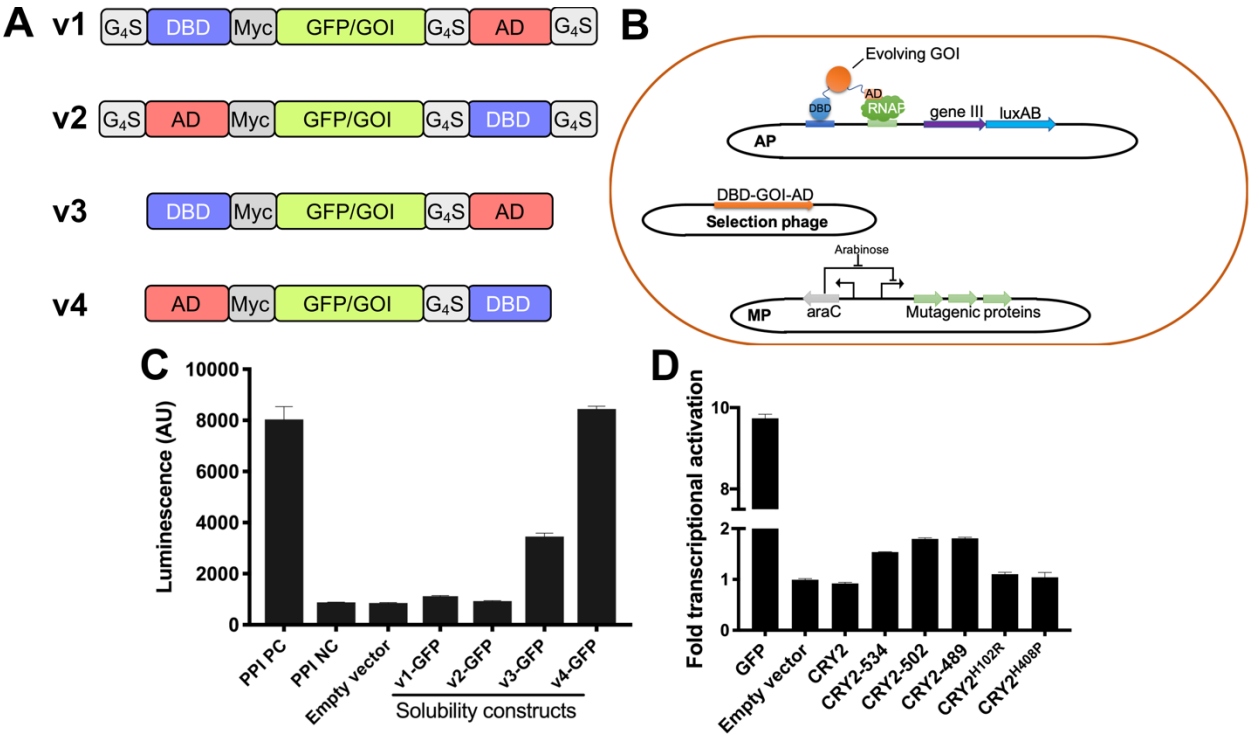

Fig. 3-7: Design of PACE constructs for evolving soluble expression.

A) Scheme of 4 versions of PACE constructs. G4S: GGGGS flexible protein linker; AD: transcriptional activation domain; Myc: epitope; GFP/GOI: green fluorescent proteins (example of soluble proteins) or gene of interest; DBD: DNA binding domain.

B) Scheme of soluble expression PACE. Gene of interest (GOI) was fused with AD and DBD. When GOI is soluble expressed, transcription of gIII will be initiated from AP. Otherwise, pIII cannot be produced.

C) Transcriptional activation of luciferase (Luminescence arbitrary units, AU) by the soluble expression constructs with GFP. PPI PC: protein-protein interaction PACE constructs, positive control using SH2-HA4 interaction; PPI NC: PPI PACE constructs, negative control using SH2-HA4^{Y87A} interaction.

D) Transcriptional activation by the soluble expression constructs with CRY2 variants. CRY2-534, 502 and 489 are truncation of CRY2 from C terminus, remaining 534, 502 and 489 amino acid residues at N terminus.

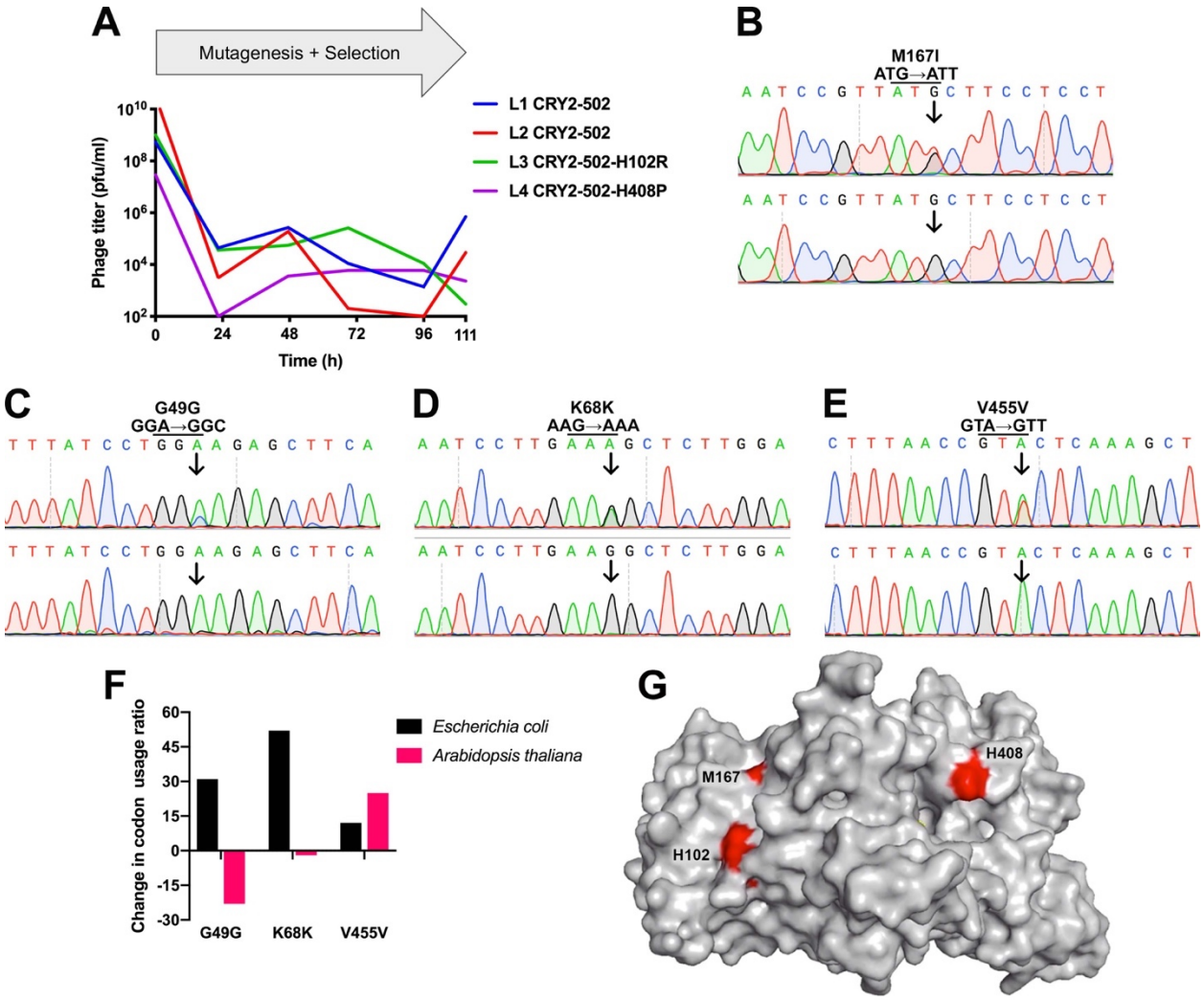

Fig. 3-8: Improved continuous evolution of CRY2 variants that bind BIC1 in the blue light.

A) Continuous evolution of CRY2 in one segment. PACE was conducted in 4 parallel lagoons: L1, L2, L3, and L4. The selection phage used at the beginning of evolution in each lagoon was as labeled. Phage was sampled every 24 hours.

B-E) Sanger sequencing chromatographs showing enriched non-synonymous mutation of $CRY2^{M167I}$ (B) and synonymous mutations $CRY2^{G49G}$ (C), $CRY2^{K68K}$ (D) and $CRY2^{V455V}$ (E).

F) Change in codon usage for the 3 enriched synonymous mutations. The ratio represents the ratio of codon usage before and after the synonymous mutations. **G)** Structure model of the PHR domain of *Arabidopsis* CRY2. Red marked the 3 residues underwent enriched non-synonymous substitution during PACE. FAD (yellow), UCRs (red), the N-terminus, the C-terminus (arrows), and 5 UCRs discussed in the text are indicated.

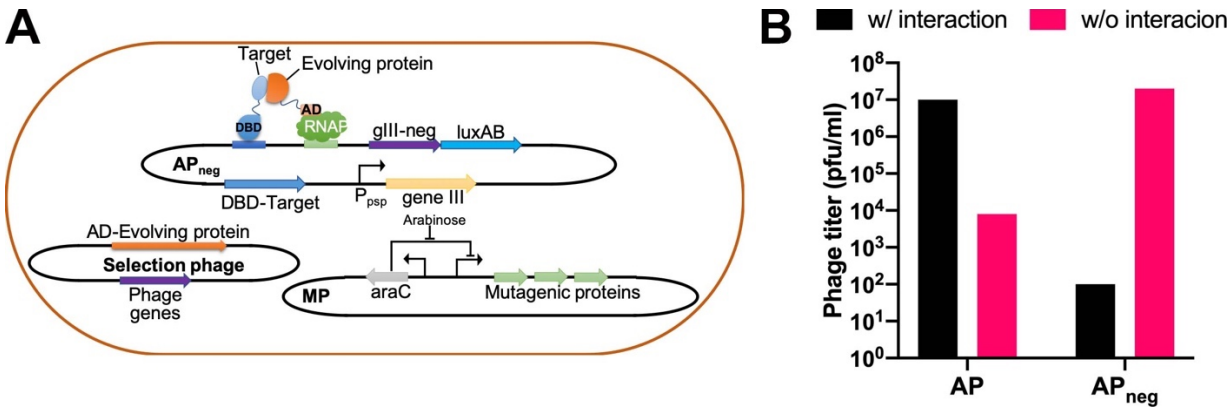

Fig. 3-9: Development of protein-dissociating PACE.

A) Strong interaction initiates transcription of a dominant negative version of gIII (gIII-neg) from AP_{neg}, suppressing the wild-type gIII expressed from phage-infection induced phage shock promoter (P_{psp}). Weak interaction cannot initiate transcription of gIII-neg from AP_{neg}. Meanwhile, gIII can be expressed from AP_{neg}, supporting phage propagation.

B) Overnight phage propagation on *E. coli* cells containing corresponding AP (pAB076i3 for AP, expressing DBD-SH2 and gIII; AP3neg2-SH2 for AP_{neg}, expressing DBD-SH2, gIII, and $P_{psp\text{-}gIII}$, see Table 3-1 for plasmid details, below the same) without constant dilution as in PACE. W/ interaction: strong SH2-HA4 interaction (SP096, expressing AD-HA4); w/o interaction: weak SH2-$HA4^{Y87A}$ interaction (SP097, expressing AD-$HA4^{Y87A}$).

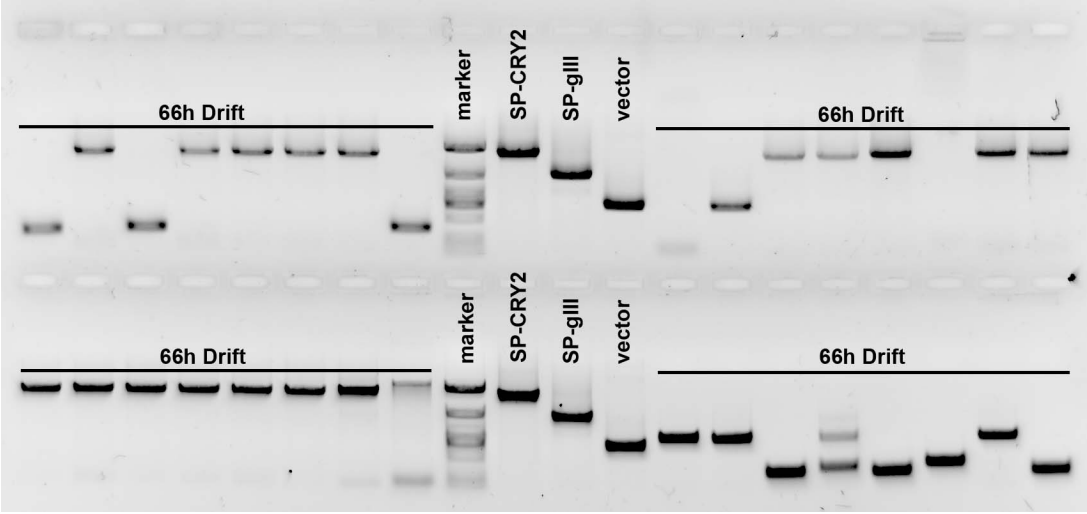

Fig. S3-1: Characterization of individual phage plaques during genetic drift.

Phage sampled at 62-hour of genetic drift was plated on *E. coli* cells. The length of the gene-of-interest DNA insertion of individual plaques was analyzed by PCR using the following primers: gVIII-197F, GAAATTCACCTCGAAAGCAA and gIII-1208R, AAAAGAAACGCAAAGACACC. Selection phage with CRY2 insertion (SP-CRY2), wild type M13 phage (SP-gIII) or empty vector without gIII and any GOI served as control.

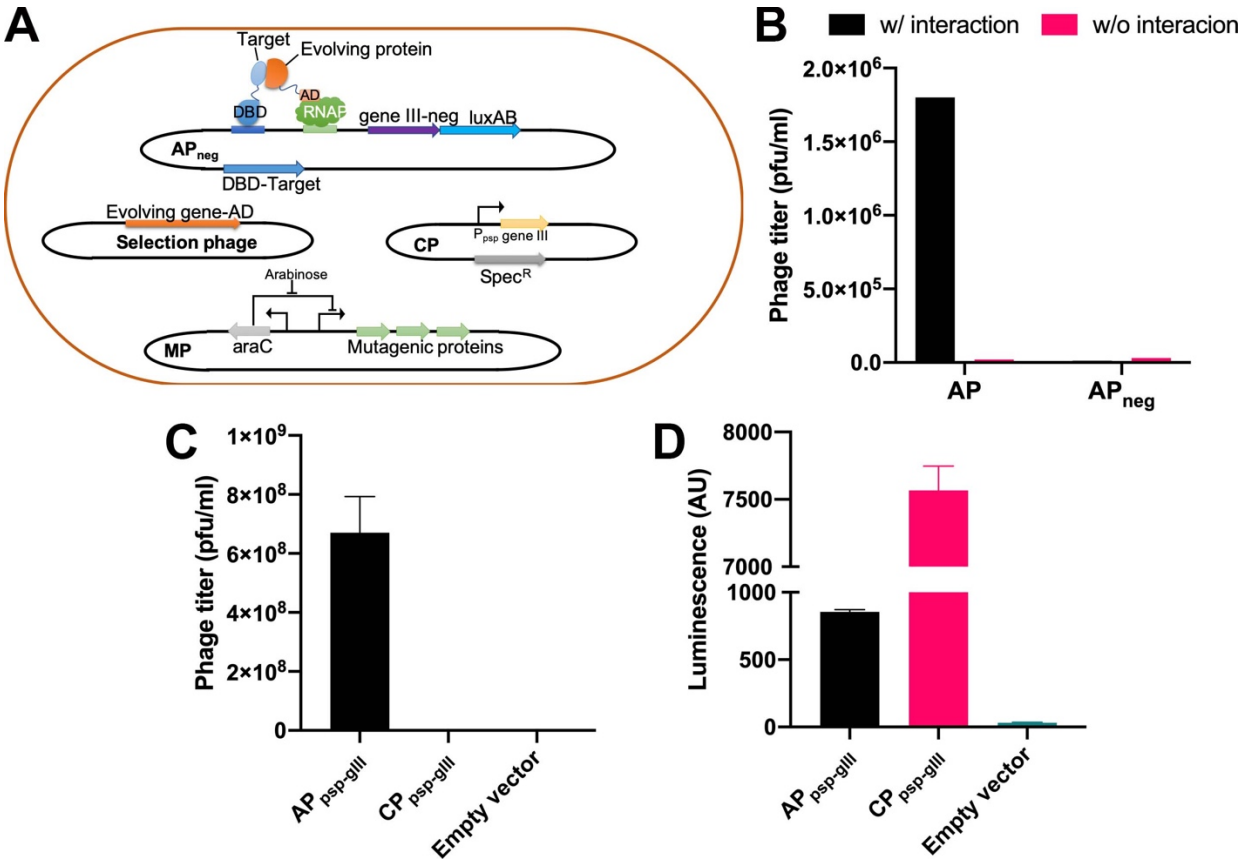

Fig. S3-2: The first design of protein-dissociating PACE.

A) Strong interaction initiates transcription of a dominant negative version of gIII (gIII-neg) from AP_{neg}, suppressing the wild-type gIII expressed from phage-infection induced phage shock promoter (P_{psp}) on CP (complementary plasmids). Weak interaction cannot initiate transcription of gIII-neg from AP_{neg}. Meanwhile, gIII can be expressed from CP.

B) Overnight phage propagation on *E. coli* cells containing CP (CP-psp-gIII-luxAB), corresponding AP and selection phage without constant dilution as in PACE. W/ interaction: strong SH2-HA4 interaction (SP096, see Table 3-1 for plasmid details, below the same); w/o interaction: weak SH2-$HA4^{Y87A}$ interaction (SP097); AP

(pAB076i3): the regular AP where strong interaction initiates transcription of wild-type gIII; AP_{neg} (AP3neg-SH2): the protein-dissociating PACE where strong interaction initiates transcription of gIII-neg.

C) Overnight phage propagation on *E. coli* cells containing $AP_{psp\text{-}gIII}$ (pJC175e, see Table 3-1 for plasmid details, below the same) or $CP_{psp\text{-}gIII}$ (CP-psp-gIII-luxAB).

D) Luminescence as a proxy of the amount of pIII produced from $AP_{psp\text{-}gIII}$ (pJC175e, see Table 3-1 for plasmid details, below the same) or $CP_{psp\text{-}gIII}$ (CP-psp-gIII-luxAB).

3.8 Tables

Table 3-1

Name	Class	Resistance	Origin of replication	ORF1		ORF2		ORF3		References
				Promoter	Gene/MCS	Promoter	Gene/MCS	Promoter	Gene/MCS	
pAB094a	PACE-CP	Spec	ColE1	P_{BAD}	rpoZ-MYC-HA4	P_C	araC	–	–	[31]
pAB094e	PACE-CP	Spec	ColE1	P_{BAD}	rpoZ-MYC-HA4 (Y87A)	P_C	araC	–	–	[31]
pAB094aHL	PACE-CP	Spec	ColE1	P_{BAD}	rpoZ-MYC-SacI	P_C	araC	–	–	This work
CP-empty	PACE-CP	Spec	ColE1	SpeI	-	–	–	–	–	This work
CP-psp-gIII-luxAB	PACE-CP	Spec	ColE1	P_{psp}	gIII, luxAB	–	–	–	–	This work
CP-CRY2	PACE-CP	Spec	ColE1	P_{BAD}	rpoZ-MYC-CRY2-FL	P_C	araC	–	–	This work
CP-CRY2-534	PACE-CP	Spec	ColE1	P_{BAD}	rpoZ-MYC-CRY2-534	P_C	araC	–	–	This work
CP-CRY2-502	PACE-CP	Spec	ColE1	P_{BAD}	rpoZ-MYC-CRY2-502	P_C	araC	–	–	This work
CP-EcPHL	PACE-CP	Spec	ColE1	P_{BAD}	rpoZ-MYC-EcPHL	P_C	araC	–	–	This work
CPS0	PACE-CP	Spec	ColEI	P_{BAD}	NsiI	–	–	–	–	This work
CPS1	PACE-CP	Spec	ColEI	P_{BAD}	G4S-434cI-MYC-rpoZ-G4S	–	–	–	–	This work
CPS2	PACE-CP	Spec	ColEI	P_{BAD}	G4S-rpoZ-MYC-434cI-G4S	–	–	–	–	This work
CPS3	PACE-CP	Spec	ColEI	P_{BAD}	434cI-MYC-rpoZ	–	–	–	–	This work
CPS4	PACE-CP	Spec	ColEI	P_{BAD}	rpoZ-MYC-434cI	–	–	–	–	This work
CPS1-GFP	PACE-CP	Spec	ColEI	P_{BAD}	G4S-434cI-MYC-rpoZ-G4S-GFP	–	–	–	–	This work
CPS2-GFP	PACE-CP	Spec	ColEI	P_{BAD}	G4S-rpoZ-MYC-GFP-434cI-G4S	–	–	–	–	This work
CPS3-GFP	PACE-CP	Spec	ColEI	P_{BAD}	434cI-MYC-GFP-rpoZ	–	–	–	–	This work
CPS4-GFP	PACE-CP	Spec	ColEI	P_{BAD}	rpoZ-MYC-GFP-434cI	–	–	–	–	This work
CPS4-CRY2	PACE-CP	Spec	ColEI	P_{BAD}	rpoZ-MYC-CRY2-434cI	–	–	–	–	This work

CPS4-CRY2-534	PACE-CP	Spec	ColEI	P_{BAD}	rpoZ-MYC-CRY2-534-434cI	–	–	–	–	This work
CPS4-CRY2-502	PACE-CP	Spec	ColEI	P_{BAD}	rpoZ-MYC-CRY2-502-434cI	–	–	–	–	This work
CPS4-CRY2-489	PACE-CP	Spec	ColEI	P_{BAD}	rpoZ-MYC-CRY2-489-434cI	–	–	–	–	This work
CPS4-CRY2-H102R	PACE-CP	Spec	ColEI	P_{BAD}	rpoZ-MYC-CRY2^{H102R}-434cI	–	–	–	–	This work
CPS4-CRY2-H408P	PACE-CP	Spec	ColEI	P_{BAD}	rpoZ-MYC-CRY2^{H408P}-434cI	–	–	–	–	This work
pJC175e	PACE-AP	Carb	SC101	P_{psp}	gIII, luxAB	–	–	–	–	[18]
pAB076i3	PACE-AP	Carb	SC101	$P_{lacZ\text{-}opt}$ (OR1)	gIII, luxAB	P_{pro1}	434cI-FLAG-SH2$_{ABL1}$	–	–	[31]
pAB107a	PACE-AP	Carb	SC101	$P_{lacZ\text{-}opt}$ (OR1)	gIII, luxAB	P_{pro1}	434cI-FLAG	–	–	[31]
pAB107aHL3 (AP3)	PACE-AP	Carb	SC101	$P_{lacZ\text{-}opt}$ (OR1+2)	gIII, luxAB	P_{pro1}	434cI-FLAG-KpnI	–	–	This work
AP3neg	PACE-AP	Carb	SC101	$P_{lacZ\text{-}opt}$ (OR1+2)	gIII-neg, luxAB	P_{pro1}	434cI-FLAG-KpnI	–	–	This work
AP3neg-SH2	PACE-AP	Carb	SC101	$P_{lacZ\text{-}opt}$ (OR1+2)	gIII-neg, luxAB	P_{pro1}	434cI-FLAG-SH2$_{ABL1}$	–	–	This work
AP3neg2	PACE-AP	Carb	SC101	$P_{lacZ\text{-}opt}$ (OR1+2)	gIII-neg, luxAB	P_{pro1}	434cI-FLAG-KpnI	P_{psp}	gIII, luxAB	This work
AP3neg2-SH2	PACE-AP	Carb	SC101	$P_{lacZ\text{-}opt}$ (OR1+2)	gIII-neg, luxAB	P_{pro1}	434cI-FLAG-SH2$_{ABL1}$	P_{psp}	gIII, luxAB	This work
AP3-BIC1	PACE-AP	Carb	SC101	$P_{lacZ\text{-}opt}$ (OR1+2)	gIII, luxAB	P_{pro1}	434cI-FLAG-BIC1	–	–	This work
AP3-EcPHL	PACE-AP	Carb	SC101	$P_{lacZ\text{-}opt}$ (OR1+2)	gIII, luxAB	P_{pro1}	434cI-FLAG-EcPHL	–	–	This work
APS1	PACE-AP	Carb	SC101	$P_{lacZ\text{-}opt}$ (OR1)	gIII, luxAB	–	–	–	–	This work
M13	PACE-SP	none	M13 f1	P_{gIII}	gIII	–	–	–	–	wild-type M13
SP096	PACE-SP	none	M13 f1	P_{gIII}	rpoZ-MYC-HA4	-	-	–	–	[31]
SP097	PACE-SP	none	M13 f1	P_{gIII}	rpoZ-MYC-HA4 (Y87A)	-	-	–	–	[31]
SP098	PACE-SP	none	M13 f1	P_{gIII}	rpoZ-MYC	-	-	–	–	[31]
SPHL3	PACE-SP	none	M13 f1	P_{gIII}	rpoZ (–A runs)-MYC	–	–	–	–	This work
SPHL3-CRY2	PACE-SP	none	M13 f1	P_{gIII}	rpoZ (–A runs)-MYC-CRY2	–	–	–	–	This work
SPHL3-CRY2-502	PACE-SP	none	M13 f1	P_{gIII}	rpoZ (–A runs)-MYC-CRY2-502	–	–	–	–	This work
SPHL3-CRY2-502-H102R	PACE-SP	none	M13 f1	P_{gIII}	rpoZ (–A runs)-MYC-CRY2^{H102R}-502	–	–	–	–	This work
SPHL3-CRY2-502-H408P	PACE-SP	none	M13 f1	P_{gIII}	rpoZ (–A runs)-MYC-CRY2^{H408P}-502	–	–	–	–	This work

MP6	PACE-MP	Chlor	CloDF13	P_{BAD}	dnaQ926, dam, seqA, emrR, ugi, cda1	P_C	araC	–	–	[35]
pQCMV-F-CRY2	Mammalian	Kan	ColE1	P_{CMV}	FLAG-CRY2	–	–	–	–	[11]
pQCMV-F-CRY2-H102R	Mammalian	Kan	ColE1	P_{CMV}	FLAG-CRY2^{H102R}	–	–	–	–	This work
pQCMV-F-CRY2-H408P	Mammalian	Kan	ColE1	P_{CMV}	FLAG-CRY2^{H408P}	–	–	–	–	This work
pCMV-M-CRY2	Mammalian	Kan/Carb	ColE1	P_{CMV}	MYC-CRY2	–	–	–	–	[11]
pCMV-M-CRY2-H102R	Mammalian	Kan/Carb	ColE1	P_{CMV}	MYC-CRY2^{H102R}	–	–	–	–	This work
pCMV-M-CRY2-H408P	Mammalian	Kan/Carb	ColE1	P_{CMV}	MYC-CRY2^{H408P}	–	–	–	–	This work
pQCMV-EGFP-BIC1	Mammalian	Kan	ColE1	P_{CMV}	EGFP-BIC1	–	–	–	–	[11]
pQCMV-EGFP-CIB1	Mammalian	Kan	ColE1	P_{CMV}	EGFP-CIB1	–	–	–	–	[11]
pQCMV-F-SPA1	Mammalian	Kan	ColE1	P_{CMV}	FLAG-SPA1	–	–	–	–	[11]
pQFLUC2-CRY2	Mammalian	Kan	ColE1	P_{CMV}	FLAG-fLUC-CRY2	–	–	–	–	This work
pQFLUC2-CRY2-H102R	Mammalian	Kan	ColE1	P_{CMV}	FLAG-fLUC-CRY2^{H102R}	–	–	–	–	This work
pQFLUC2-CRY2-H408P	Mammalian	Kan	ColE1	P_{CMV}	FLAG-fLUC-CRY2^{H408P}	–	–	–	–	This work
pDT1h-4M-CRY2	Plant	Kan/Hyg	ColE1	P_{ACTIN2}	4xMYC-CRY2	–	–	–	–	[11]
pDT1h-4M-CRY2-H102R	Plant	Kan/Hyg	ColE1	P_{ACTIN2}	4xMYC-CRY2^{H102R}	–	–	–	–	This work
pDT1h-4M-CRY2-H408P	Plant	Kan/Hyg	ColE1	P_{ACTIN2}	4xMYC-CRY2^{H408P}	–	–	–	–	This work

Table 3-1: Summary of plasmids

Table 3-2

Peristaltic pumps	Cat #	Supplier
Masterflex L/S Digital Drive, 100 RPM, 115/230 VAC	EW-07522-30	Cole-Parmer
L/S 8-channel multichannel pump head for microbore tubing	EW-07534-08	Cole-Parmer
Syringe pumps	**Cat #**	**Supplier**
NE-1600 Six Channel Programmable Syringe Pump	NE-1600	New Era Pump Systems, Inc.
Tubing	**Cat #**	**Supplier**
Microbore 2-stop, silicone (platinum cured); 0.89 mm ID, 6/pk	EW-06421-26	Cole-Parmer
Microbore 2-stop, silicone (platinum cured); 1.42 mm ID, 6/pk	EW-06421-34	Cole-Parmer
Microbore 2-stop, silicone (platinum cured); 2.06 mm ID, 6/pk	EW-06421-42	Cole-Parmer
Microbore 2-stop, silicone (platinum cured); 2.79 mm ID, 6/pk	EW-06421-48	Cole-Parmer
Connections	**Cat #**	**Supplier**
1/16" hose barb adapter, PVDF, 25/pk	WU-45512-00	Cole-Parmer
1/16" with lock ring, hose barb, PVDF, 25/pk	WU-45513-00	Cole-Parmer
3/32" hose barb adapter, PVDF, 25/pk	WU-45512-02	Cole-Parmer

3/32" with lock ring, hose barb, PVDF, 25/pk	WU-45513-02	Cole-Parmer
1/8" hose barb adapter, PVDF, 25/pk	WU-45512-04	Cole-Parmer
1/8" with lock ring, hose barb, PVDF, 25/pk	EW-45513-04	Cole-Parmer
Stirplates and stirbars	**Cat #**	**Supplier**
Cimarec i Poly 15 and Multipoint Stirrers, 15 point	50093538	Thermo Scientific
1/2 x 1/8 Displosable Spinbar, Octagon, bg/100	58947-140	VWR
Chemostats/Lagoon vessels	**Cat #**	**Supplier**
VWR 100 mL w/out Cap, 10/pk	89012-114	VWR
Corning High-Temperature PBT Cap, Red, Open, GL45	11310-688	VWR (Corning)
Corning PTFE-Faced Silicone Septa, 10/pkm 45 mm	1395-45TS	VWR (Corning)
22ml Clear Vial, Open Top, Cap 10/90 Septa	B7990-5	Thermo Scientific
Carboy for media	**Cat #**	**Supplier**
Thermo Sci., Nalgene, 3-port	2162-0831	Thermo Scientific
10L Heavy duty Corboy	2226-0020	Thermo Scientific
Needles/Syringe/Filters	**Cat #**	**Supplier**
Air-Tite Veterinary Needle, Luer, 22-gauge, 3 1/2"	89219-276	VWR (AirTite)

Air-Tite Veterinary Needle, Luer, 18-gauge, 1 1/2"	89219-300	VWR (AirTite)
Cadence Science Blunt End Pipetting Needle, Luer, 14-gauge, 12/pk	20068-690	VWR (Cadnece Sci.)
Cadence Science Blunt End Pipetting Needle, Luer, 16-gauge, 12/pk	20068-686	VWR (Cadnece Sci.)
Soft-Ject® 3-Part Disposable Syringes, Air-Tite, 60 ml	89215-240	VWR
HSW® Norm-Ject® Sterile Luer-Slip Syringes, Air-Tite, 1 ml	53548-001	VWR
Syringe filters PES, 0.22µm, 13 mm diameter	25-243	Genesee

Table 3-2: PACE devices and parts.

Chapter 4 Concluding remarks

Cryptochromes are evolutionarily ancient proteins found in all three kingdoms of life, with diverse functions for different homologs and paralogs of the protein family members [1]. Presumably evolved from DNA photolyases, CRYs and PHLs share similar primary, secondary, tertiary protein structures and the same co-factor FAD. Although the phylogenetic evolution of members in the CRY/PHL family was well studied, the structure-function relationship in evolutionary perspective was not understood. In Chapter 2, I explored the relationship between UCRs of CRY/PHL family and the functions of CRY2, by studying functional consequences of UCR (universally conserved residue) substitution. In Chapter 3, I explored how protein sequences were changed along with changes in specific activities.

UCRs are residues that are invariable in across different members from various evolutionary lineages of a protein family [2]. UCRs are traditionally considered to be universally important for maintaining structure integrity, and therefore affects all aspects of the protein functions, such that mutations at those positions is prevented from accumulating under selection pressures over long time of evolution. However, the notion was anecdotal and was not examined comprehensively. The comprehensive study of UCR needs to be conducted in an evolutionarily ancient protein family, with diverse functions for different protein members, and needs accurate readouts of diverse functions. The CRY/PHL protein family is therefore a suitable system to comprehensively study the importance of UCR. In Chapter 2, I systematically examined the functional consequences of replacement of 51 UCRs of *Arabidopsis thaliana* CRY2.

Four accurate functional readouts of relative-specific-activities were provided by a standard curve-based quantitative method. To my knowledge, this is the first time that the photophysiological or photobiochemical activities of CRY2 mutant proteins are measured using the protein abundance-dependent method or in the form of relative-specific-activity. Consistent with the perceived importance of UCRs, all UCRs are required for stable expression of proteins or at least 1 of the 4 examined phenotypes. However, not all UCRs are universally required for stable protein expression or functions of CRY2, exemplified by the fact that 45% (23/51) of the UCR mutant proteins retained wild-type-like relative-specific-activities in at lease 1 of the tested phenotypes. Although the UCRs define common structure elements of members of the CRY/PHL protein family, the respective functions of the UCRs were probably repurposed in different proteins. For example, some proteins in the CRY/PHL family have a secondary co-factor-binding pocket to bind either pterin or deazaflavin. However, residues in secondary pockets remain conserved even in CRYs without secondary co-factor, which can be explained by repurposing the secondary pockets for novel functions [3]. For example, secondary pockets were adapted for interactions with CLOCK/BMAL1 in mammalian CRYs. Given the above, I hypothesize that the maintenance of the extreme conservation of UCRs does not require all protein functions to depend on UCR. It is possible that a subset of functions is enough to affect fitness to enforce the dependence upon the UCR, thus keeping UCRs invariable during millions of years of evolution. However, functions of proteins are determined by both UCRs and non-UCRs. The features that distinguish the two groups of residues remain elusive.

In Chapter 3, I adopted an opposite method to study the molecular evolution of CRY2 proteins. I used directed evolution to modify the biochemical activities of CRY2 to observe how UCRs and non-UCRs were changed accordingly. I demonstrated that the PACE method was capable of introducing multiple mutations to improve the desired biochemical functions of CRY2. However, the number of acquired mutations was not enough to manifest the difference between UCRs and non-UCRs. I developed 2 new PACE systems to facilitate the evolution of CRY2 for a more comprehensive analysis of features of UCRs and non-UCRs regarding the biochemical activity of CRY2. In addition, the engineer of CRY2-BIC1 interaction sheds light on optogenetics and the balance of CRY2-complexome. Blue-light specific interaction of CRY2-BIC1 can be useful optogenetic tools with advantages of smaller protein size and less homooligomerization. Modification of PPI affinity of CRY2-BIC1 provides a unique opportunity to thoroughly investigate the relationship between biochemical activities and physiological functions of CRY2 proteins.

One of the interesting follow-up questions is to study the co-evolution of different UCRs. In this thesis, effects of UCR substitutions were only interrogated separately, while genetic interactions of residues, such as additive, epistatic, suppressed and synergistic effects, were not taken into consideration. Additionally, further directed evolution efforts to reprogram CRY2-BIC1 interaction surface can probe the characteristics of UCRs and non-UCRs, enrich optogenetic tools, and provide insights into the interacting surface of CRY2-BIC1 and CRY2-complexome balance.

What I find, in general, is that UCRs are not universally required for physiological and biochemical functions of cryptochromes. Clearly, there are many questions and

hypotheses regarding the characteristics of UCRs and non-UCRs to be further explored. This work provides first systematic evidence of functions of UCRs experimentally.

Reference

1. Mei, Q., and Dvornyk, V. (2015). Evolutionary history of the photolyase/cryptochrome superfamily in eukaryotes. PLoS One 10, e0135940.
2. Liu, H., Su, T., He, W., Wang, Q., and Lin, C. (2019). The universally conserved residues are not universally required for stable protein expression or functions of cryptochromes. Mol. Biol. Evol.
3. Rosensweig, C., Reynolds, K.A., Gao, P., Laothamatas, I., Shan, Y., Ranganathan, R., Takahashi, J.S., and Green, C.B. (2018). An evolutionary hotspot defines functional differences between CRYPTOCHROMES. Nat. Commun. 9, 1138.