Characterization of novel rhodopsins with light-regulated cGMP production or cGMP degradation

Charakterisierung neuartiger Rhodopsine mit Licht-regulierter cGMP-Produktion oder cGMP-Degradation

Doctoral thesis for a doctoral degree
at the Graduate School of Life Sciences,
Julius-Maximilians-Universität Würzburg
Section Integrative Biology

submitted by

Yuehui Tian

From Shanxi, China

Würzburg, 2018

Submitted on:

Members of the *Promotionskomitee*:

Chairperson: Prof. Dr. Martin Müller

Primary Supervisor: Prof. Dr. Georg Nagel

Supervisor (Second): Prof. Dr. Robert Kittel

Supervisor (Third):   Prof. Dr. Thomas Müller

Supervisor (Fourth):

(If applicable)

Date of Public Defence:

Date of Receipt of Certificates:

Contents

Contents

Contents

Abstract

Photoreceptors are widely occurring in almost all kingdoms of life. They mediate the first step in sensing electromagnetic radiation of different wavelength. Absorption spectra are found within the strongest radiation from the sun and absorption usually triggers downstream signaling pathways. Until now, mainly 6 classes of representative photoreceptors are known: five water-soluble proteins, of these three classes of blue light-sensitive proteins including LOV (light-oxygen-voltage), BLUF (blue-light using FAD), and cryptochrome modules with flavin (vitamin B-related) nucleotides as chromophore; while two classes of yellow and red light-sensitive proteins consist of xanthopsin and phytochrome, respectively. Lastly, as uniquely integral membrane proteins, the class of rhodopsins can usually sense over a wide absorption spectrum, ranging from ultra-violet to green and even red light. Rhodopsins can be further divided into two types, i.e., microbial (type I) and animal (type II) rhodopsins. Rhodopsins consist of the protein opsin and the covalently bound chromophore retinal (vitamin A aldehyde). In this thesis, I focus on identification and characterization of novel type I opsins with guanylyl cyclase activity from green algae and a phosphodiesterase opsin from the protist *Salpingoeca rosetta*.

Until 2014, all known type I and II rhodopsins showed a typical structure with seven transmembrane helices (7TM), an extracellular N-terminus and a cytosolic C-terminus. The proven function of the experimentally characterized type I rhodopsins was membrane transport of ions or the coupling to a transducer which enables phototaxis via a signaling chain. A completely new class of type I rhodopsins with enzymatic activity was identified in 2014. A light-activated guanylyl cyclase opsin was discovered in the fungus *Blastocladiella emersonii* which was named Cyclop (Cyclase opsin) by Gao et al. (2015), after heterologous expression and rigorous *in-vitro* characterization. *Be*Cyclop is the first opsin for which an 8 transmembrane helices (8TM) structure was demonstrated by Gao et al. (2015). Earlier (2004), a novel class of enzymatic rhodopsins was predicted to exist in *C. reinhardtii* by expressed sequence tag (EST) and genome data, however, no functional data were provided up to now. The hypothetical rhodopsin included an N-terminal opsin domain, a fused two-component system with histidine

kinase and response regulator domain, and a C-terminal guanylyl cyclase (GC) domain. This suggested that there could be a biochemical signaling cascade, integrating light-induction and ATP-dependent phosphate transfer, and as output the light-sensitive cGMP production.

One of my projects focused on characterizing two such opsins from the green algae *Chlamydomonas reinhardtii* and *Volvox carteri* which we then named 2c-Cyclop (**two**-**c**omponent **Cycl**ase **op**sin), *Cr*2c-Cyclop and *Vc*2c-Cyclop, respectively. My results show that both 2c-Cyclops are light-inhibited GCs. Interestingly, *Cr*2c-Cyclop and *Vc*2c-Cyclop are very sensitive to light and ATP-dependent, whereby the action spectra of *Cr*2c-Cyclop and *Vc*2c-Cyclop peak at ~540 nm and ~560 nm, respectively. More importantly, guanylyl cyclase activity is dependent on continuous phosphate transfer between histidine kinase and response regulator. However, green light can dramatically block phosphoryl group transfer and inhibit cyclase activity. Accordingly, mutation of the retinal-binding lysine in the opsin domain resulted in GC activity and lacking light-inhibition.

A novel rhodopsin phosphodiesterase from the protist *Salpingoeca rosetta* (*Sr*RhoPDE) was discovered in 2017. However, the previous two studies of 2017 claimed a very weak or absent light-regulation. Here I give strong evidence for light-regulation by studying the activity of *Sr*RhoPDE, expressed in *Xenopus laevis* oocytes, *in-vitro* at different cGMP concentrations. Surprisingly, hydrolysis of cGMP shows a ~100-fold higher turnover than that of cAMP. Light can enhance substrate affinity by decreasing the K_m value for cGMP from 80 μM to 13 μM, but increases the maximum turnover only by ~30%. In addition, two key single mutants, *Sr*RhoPDE K296A or K296M, can abolish the light-activation effect by interrupting a covalent bond of Schiff base type to the chromophore retinal. I also demonstrate that *Sr*RhoPDE shows cytosolic N- and C- termini, most likely via an 8-TM structure. In the future, *Sr*RhoPDE can be a potentially useful optogenetic tool for light-regulation of cGMP concentration, possibly after further improvements by genetic engineering.

Zusammenfassung

Photorezeptoren sind in fast allen Lebewesen vorzufinden. Sie vermitteln den ersten Schritt bei der Detektion von elektromagnetischer Strahlung unterschiedlicher Wellenlänge. Ihre Absorptionsspektren finden sich innerhalb des Bereichs der stärksten Sonnenstrahlung (UV bis nahes IR) und die Absorption löst normalerweise nachgelagerte Signalwege aus. Bis jetzt sind hauptsächlich 6 Klassen von Photorezeptoren bekannt: als wasserlösliche Proteine zunächst drei Klassen von Blaulicht-empfindlichen Modulen, die LOV-Domäne (Light/Oxygen/Voltage), die BLUF-Domäne (Blue Light sensing, Using FAD) und Cryptochrom mit Flavinen (Vitamin B-Komplex) als Chromophor, sowie Xanthopsine, die hauptsächlich für gelbes Licht und Phytochrome, die für Rotlicht empfindlich sind. Als integrale Membranproteine kann die Klasse der Rhodopsine jedoch ein breiteres Absorptionsspektrum aufweisen, je nach Rhodopsin empfindlich für UV/blaues, grünes oder sogar rotes Licht. Rhodopsine bestehen aus dem Protein Opsin und dem kovalent gebundenen Chromophor Retinal (Vitamin A-Aldehyd). Sie können weiter in zwei Typen unterteilt werden, mikrobielle (Typ I) und tierische (Typ II) Rhodopsine. In dieser Dissertation konzentriere ich mich auf die Identifizierung und Charakterisierung von neuen Typ I Opsinen mit Guanylylcyclase-Aktivität aus Grünalgen und einem Phosphodiesterase-Opsin aus dem Protisten *Salpingoeca rosetta*.

Bis 2014 wiesen alle bekannten Rhodopsine eine typische Struktur mit sieben Transmembranhelices (7TM), einem extrazellulären N-Terminus und einem zytosolischen C-Terminus auf. Die nachgewiesene Funktion der experimentell charakterisierten Rhodopsine vom Typ I ist der Membrantransport von Ionen oder die Kopplung an einen Transducer, der die Phototaxis über eine Signalkette ermöglicht. Eine völlig neue Klasse von Typ-I Rhodopsinen mit enzymatischer Aktivität wurde 2014 gefunden. Ein lichtaktiviertes Guanylyl-Cyclase-Opsin wurde im Pilz *Blastocladiella emersonii (Be)* entdeckt, das von Gao et al. (2015) nach heterologer Expression und gründlicher *in-vitro*-Charakterisierung Cyclop (Cyclase opsin) genannt wurde. *Be*Cyclop ist das erste Opsin, für das eine Struktur mit 8 Transmembranhelices (8TM) demonstriert wurde (Gao et al., 2015). Bereits früher (2004) wurde durch expressed sequence tag (EST) und Genom-Daten vorhergesagt, dass eine neue Klasse von enzymatischen

Rhodopsinen in *Chlamydomonas reinhardtii* existiert, jedoch gelang bisher keine funktionelle Expression. Eines dieser hypothetischen Rhodopsine umfasste eine N-terminale Opsin-Domäne, ein fusioniertes Zweikomponentensystem mit Histidinkinase- und Regulator-Domäne und eine C-terminale Guanylylcyclase (GC). Dies legte nahe, dass das Protein eine biochemische Signalkaskade inkorporieren könnte, die Licht-Absorption und ATP-abhängigen Phosphattransfer integriert und eine lichtempfindliche cGMP-Produktion bewirkt.

Eines meiner Projekte konzentrierte sich auf die Charakterisierung von zwei solchen Opsinen aus den Grünalgen *C. reinhardtii* und *Volvox carteri*, die wir nun 2c-Cyclop (Zweikomponenten-Cyclase-Opsin), d.h. *Cr*2c-Cyclop und *Vc*2c-Cyclop, nennen. Meine Ergebnisse zeigen, dass beide 2c-Cyclops durch Licht gehemmte GCs sind. Interessanterweise sind *Cr*2c-Cyclop und *Vc*2c-Cyclop sehr empfindlich gegenüber Licht und die cGMP-Produktion ist ATP-abhängig, wobei die Wirkungsspektren von *Cr*2c-Cyclop und *Vc*2c-Cyclop bei ~540 nm bzw. ~560 nm ihren Höhepunkt erreichen. Ich konnte zeigen, dass die Guanylyl-Cyclase-Aktivität von einem kontinuierlichen Phosphat-Transfer zwischen Histidinkinase und Response-Regulator abhängt. Grünes Licht kann jedoch den Phosphoryl-Gruppen-Transfer dramatisch blockieren und die Cyclase-Aktivität inhibieren. Dementsprechend führte die Mutation des Retinal-bindenden Lysins in der Opsindomäne zu einer cGMP Produktion ohne jegliche Lichtinhibierung.

Eine neuartige Rhodopsin-Phosphodiesterase aus dem Protisten *Salpingoeca rosetta* (*Sr*RhoPDE) wurde im Jahr 2017 entdeckt. Die vorangegangenen zwei Studien von 2017 zeigten jedoch eine sehr schwache oder fehlende Lichtregulation der PDE-Aktivität. Hier konnte ich eine starke Lichtregulation beweisen, indem ich die Aktivität von *Sr*RhoPDE, exprimiert in *Xenopus laevis* Oozyten, *in-vitro* bei verschiedenen cGMP-Konzentrationen untersuchte. Überraschenderweise zeigt die Hydrolyse von cGMP einen etwa 100-fach höheren Umsatz als der von cAMP. Licht kann die Substrataffinität durch Verringern des K_m-Werts für cGMP von 80 µM auf 13 µM erhöhen, erhöht jedoch den maximalen Umsatz nur um ~30%. Darüber hinaus können zwei Einzelmutanten, *Sr*RhoPDE K296A oder K296M, den Lichtaktivierungseffekt aufheben, indem sie eine kovalente Bindung vom Schiff-Base-Typ an das Chromophor Retinal unterbrechen. Ich zeige auch, dass *Sr*RhoPDE zytosolische N- und C-

Termini aufweist, höchstwahrscheinlich über eine 8-TM-Struktur. In Zukunft könnte *Sr*RhoPDE ein potentiell nützliches optogenetisches Werkzeug für die Lichtregulation der cGMP-Konzentration sein, möglicherweise nach weiteren Verbesserungen durch Gentechnik.

1. Introduction

Natural photoreceptors as a source of optogenetic tools are widely discovered from three kingdoms of life. In the chromophore-binding pockets of sensory modules, light illumination can firstly trigger chemical changes of chromophores, involving covalent bond formation (LOV[1]), electronic transfer (Cryptochrome[2]) and isomerization change (Rhodopsins[3, 4]). Conformational changes of sensor domains will further regulate output enzyme catalytic domains, interact with its partners to form dimers, or open ion channels etc.[5]. Type-I and type-II rhodopsins from microbial organisms and animals respectively show distinct signaling pathways under light illumination[6]. In type-I rhodopsins, single-domain opsins usually play roles as light-gated channels or pumps, while enzyme opsins have distinct activities induced by light-regulated conformational changes of sensor modules[3, 4, 7, 8]. cGMP, as one of the second messengers, plays roles in many physiological processes[9]. Identifying and characterizing more cGMP-related opsins is essential to expanding optogenetic tools for further application.

1.1 Type-I and type-II rhodopsins

Rhodopsins have been divided into two types[6]. Type-I rhodopsins, as brief or direct light sensors, come from microbial organisms (e.g. prokaryotes, algae, and fungi). This microbial opsin family members play roles in phototactic response, photosynthesis development, and obtainment of energy source, and so on[10]. Type-II vertebrate rhodopsins belong to G protein-coupled receptors (GPCR), mainly regulating visual systems and nonvisual light-stimulated signaling pathways (e.g. circadian clock) [11-13].

The classical type-I and type-II opsin families share similar transmembrane topologies which consist of seven transmembrane helices, an extracellular N-terminal segment and an intracellular C-terminal tail[12]. The two opsin families have poorly conserved sequences[14, 15], but sequence similarity within one family can be higher (25%-80%)[16]. Moreover, the two families utilize different initial isomers of retinal. In the last transmembrane helix, an interior protonated retinylidene Schiff base links chromophore to the ε-amino group of a conserved lysine residue. Type-I opsins utilize all-*trans*-retinal (ATR) as the chromophore, which can be isomerized to 13-*cis*-retinal after light illumination. While type-II opsins use 11-*cis*-retinal

which can be isomerized to ATR upon photon absorption[17].

Natural photoreceptors especially type-I opsins can be applied as building blocks to engineer optogenetic tools[18]. Obviously, compared with animal rhodopsins utilizing a signaling cascade of several components, microbial opsins have many advantages, for example, one gene can integrate both light sensor module and effector domain, some essential cofactors like retinal and biliverdin exist in vertebrate tissues, and visible-spectrum light can be applied biologically with well-tolerated wavelengths and intensities[19]. The representative photoreceptors are depicted in **Figure 1.1**, comprising their individual absorption spectrum, chromophores, photon sensor modules and schematic of domains. It is clearly shown that these photoreceptors can sense visible-spectrum light from UV-light to Near-infrared (NIR) light[18].
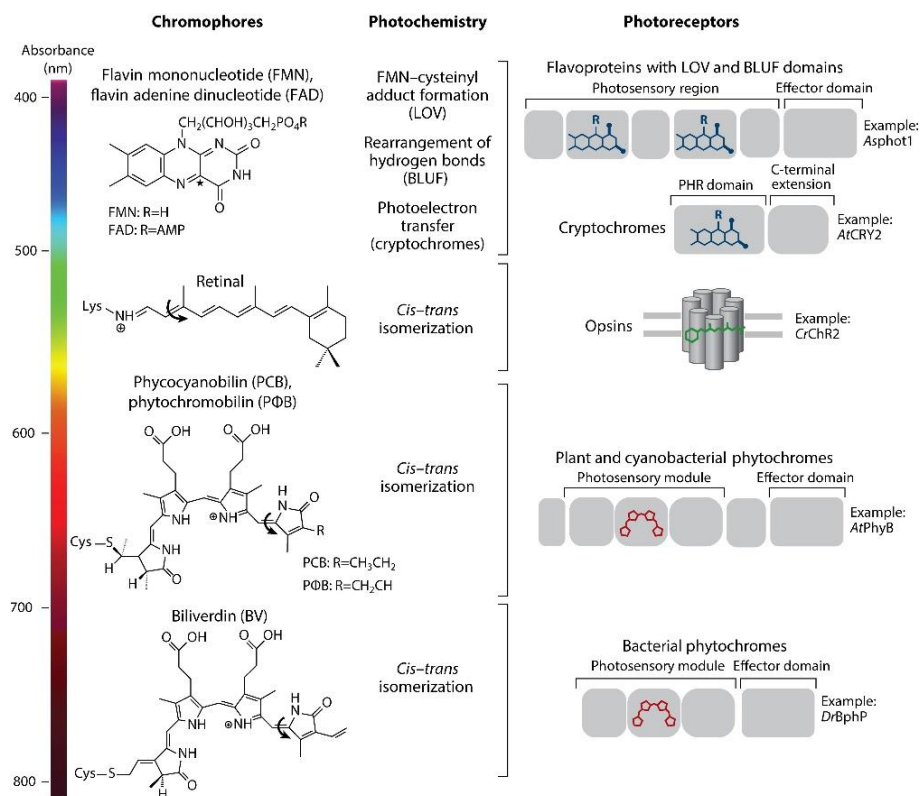
Figure 1.1 Natural photoreceptors and their chromophores for optical tools engineering.

(Left) The color-bar represents absorbance of natural photoreceptors embedded with chemical chromophores. Chromophores are described in their inactive (dark) state. An asterisk ✱ indicates covalent bond formation between the C4a atom in FMN and a conservative cysteine in LOV domains after light illumination. The arrow in some chromophores means isomerization changes in the double bond upon activation.

(Right) Schematic domain structure in each photoreceptor. Apart from single-domain Channelrhodopsin-2, other photoreceptors comprise two or more photosensory modules and effector domains. Taken from [18].

Type-II opsins show more sophisticated signaling cascades and slower kinetics than type-I opsins. Based on sequence homology, type-II opsins belong to GPCRs superfamily A. A typical example is rhodopsins existing in the disc membrane of rod and cone cells of retinas in mammalian eyes, implying the similar molecular mechanism. The signaling cascade initiates upon photons absorption, and then opsins can bind to G protein by catalyzing GDP to GTP. This makes GTP-bound Gα protein dissociate with Gβγ subunits and keep in an active state. Activated Gα will then bind to PDE6 (cyclic nucleotide phosphodiesterase), and activate it to decrease the concentration of cGMP, thus closing CNG (cyclic nucleotide-gated) channels. This allows to hyperpolarize the membrane potential of photoreceptor cells and release neurotransmitters to different cells in downstream, finally forming neuronal signaling in the brain[11, 20, 21]. Some key proteins are involved in rod phototransduction, as shown in **Figure 1.2**. Although light stimulates G protein signaling cascades and causes CNG channel closure, guanylyl cyclase activating proteins (GCAP) can activate guanylate cyclase (GC) at a low concentration of Ca^{2+}. This tends to restore cGMP levels and re-open CNG channels[21].
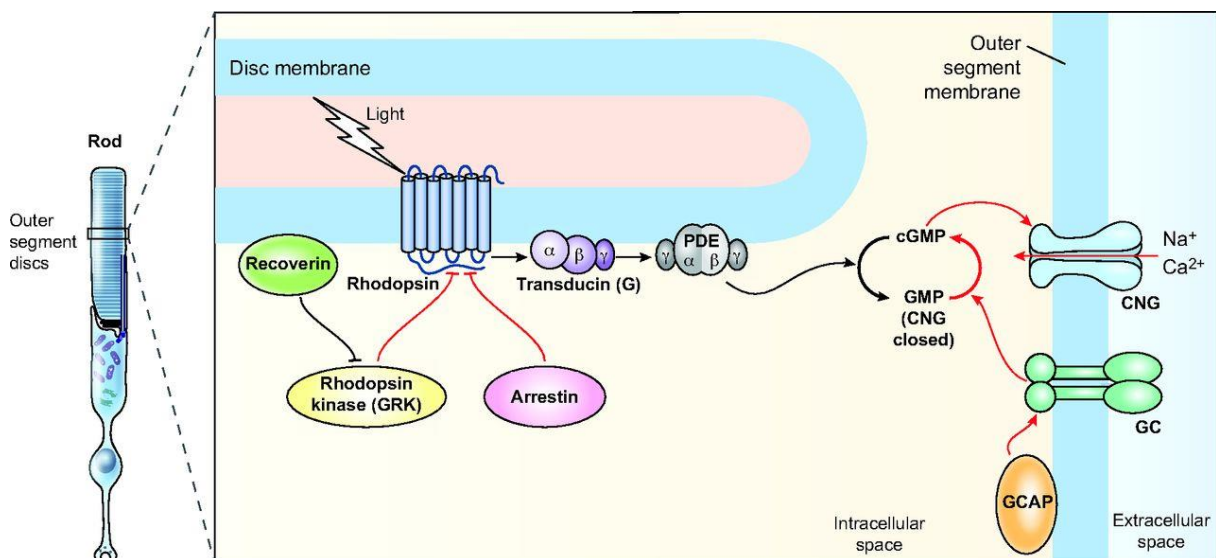
Figure 1.2 Schematic view of key proteins involved in phototransduction.

After initiation of phototransduction by light illumination, phosphorylation of rhodopsin kinase [G-protein receptor kinase (GRK)] can facilitate arrestin to bind and inactivate rhodopsin, thus terminating phototransduction. Recoverin can inhibit GRK and control the lifetime of activated rhodopsin in dark and high intracellular Ca^{2+} levels. The concentration of Ca^{2+} for negative feedback is essential to maintain phototransduction. Taken from [21].

1.2 Light-gated channels or pumps as Type-I opsins

Type-I rhodopsins enable the generation of optogenetic tools because of their simple and efficient structures and functions. They mainly exist in archaea, prokaryotes, and lower eukaryotes (like algae and fungi), living in hot springs and salt lakes (for archaea), oceans, lakes, rivers, soil and leaf surfaces of plants[22]. During over one billion years evolution, microbial rhodopsins as single protein scaffolds have developed and a wide diversity of biochemical functions have emerged, playing roles in regulating many physiological processes.

As single domain microbial rhodopsins, they can directly convert light energy to electrochemical potentials due to properties of channels or pumps. In the early 1970s, the first microbial rhodopsin named bacteriorhodopsin (BR) was discovered in "purple membrane" of *Halobacterium salinarum*, and then it was proved as a light-gated proton pump[23-25]. Ten years later, Halorhodopsin (generally named HR) from *H. salinarum* was determined as an inwardly directed chloride pump stimulated by light[26]. Moreover, *Np*HR from *Natronobacterium pharaonic* was presented with similar properties as a light-activated chloride pump[27]. BR and HR come from members of the domain Archaea, which can preserve their functions in extreme environments including a high concentration of NaCl and extremes of pH for growth. This will build up and maintain ion gradients across cell membranes [28].

Photoreceptors above in archaea belong to light-gated proton or chloride pumps. Remarkably, in the year of 2002, Georg Nagel et al. firstly demonstrated that Channelrhodopsin-1 (ChR1) from the green alga *Chlamydomonas reinhardtii* was a light-gated proton channel after heterologous expression in *Xenopus* oocytes[3]. One year later, Nagel et al. firstly determined that Channelrhodopsin-2 (ChR2) from the same green alga was a non-selective cation channel by blue light illumination. This was then suggested to depolarize different types of cells[4]. After application in different neural cells, Channelrhodopsin-2 and its mutants (e.g. ChR2-XXL, D156C mutant with extra high expression and long open state[29]) were determined as powerful optogenetic tools to excite cells by light-induced membrane depolarization. For example, ChR2-derived optogenetic tools were successfully activating many cell types, including *C. elegans* muscle cells, *Drosophila* motor neurons as well as mouse hippocampal neural cells[29-31]. Optogenetics comes from the development of those single-domain tools to control excitation

or silence of neural cells by light, thus dramatically improving researches in the neuroscientific field. Three typical signal-domain rhodopsins are depicted in **Figure 1.3** [32]. In addition, a red-shifted *Cn*ChR1 (*Chlamydomonas noctigama*) named Chrimson was discovered by *de novo* sequencing, with absorption spectrum peaking at 590 nm[33]. This could be further mutated and applied in neural cells of deep tissues by red light stimulation. On the contrary, other light-gated anion channels like *Gt*ACRs were identified from cryptophyte *Guillardia theta* and they can be further utilized to hyperpolarize cell membrane or silence neurons[34].

Until now, Channelrhodopsin-2 and its variants are widely applied in many neuroscientific fields. More importantly, it is likely to utilize ChR2 in the restoration of vision because of the ubiquitous expression of ChR2 in retinal ganglion cells[35]. Integration of both ChR2 and chloride pump Halorhodopsin could also proceed with next investigations of neural diseases like Parkinson's disease or auditory disorders[36, 37].
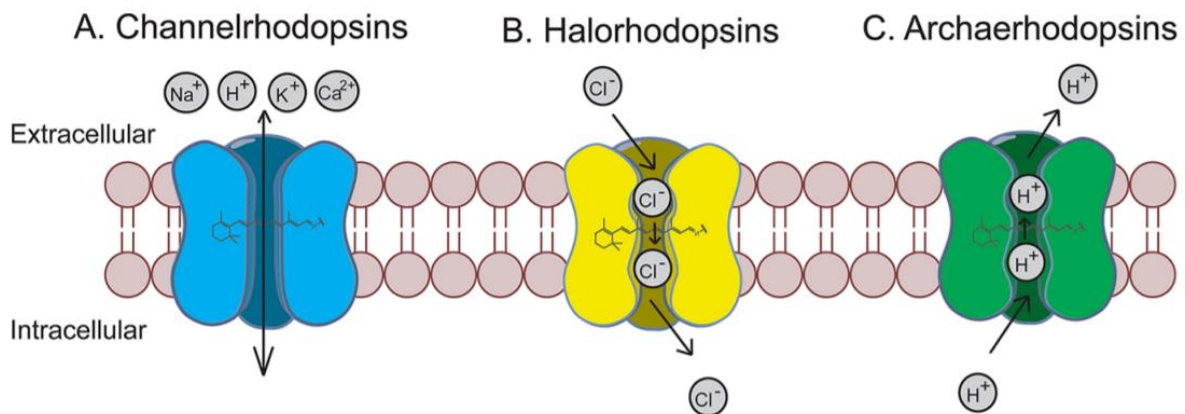
Figure 1.3 Single-domain optogenetic tools.

Light-stimulation can isomerize all-*trans*-retinal to 13-*cis*-retinal, opening the channels or pumps.

A, Channelrhodopsins can passively transport cation ions including Na^+, K^+, H^+, Ca^{2+} etc. following electrochemical gradients, thus depolarizing neurons;

B, Halorhodopsins can hyperpolarize neurons by pumping Cl^- into the cell;

C, Archaerhodopsins can hyperpolarize neurons by pumping H^+ out of the cell. Taken from [32].

1.3 cGMP or cAMP related microbial photoreceptors

Other microbial rhodopsins, existing in bacteria, algae and fungi, also integrate two or more domains. For instance, light-induced adenylyl and guanylyl cyclases comprise light sensor modules (BLUF or Rhodopsin) and output enzymatic domains, which can manipulate levels of cyclic adenosine 3'-5'-monophosphate (cAMP) or cyclic guanosine 3'-5'-monophosphate (cGMP) by light illumination.

1.3.1 Cyclic nucleotides (cAMP and cGMP) and signal transduction

Regulatory systems relating cyclic nucleotides have been awarded no less than five Nobel prizes in Physiology or Medicine field[9]. In animals, first messengers such as light, nitric oxide, hormones can regulate second messenger cyclic nucleotides (cAMP and cGMP) levels. This will further have effects on vision, muscle contraction, cardiovascular, memory and many other functions[38-41].

Although eukaryotes and bacteria have different essential signal-transduction pathways, they still share certain signaling components, such as similar class III adenylate/guanylate cyclase family based on PF00211 domain in Pfam database[42]. A few residues near substrate binding pocket could be able to determine the selectivity of cyclases to nucleotide substrates[43]. In eukaryotes, cNMP (e.g. cAMP or cGMP) is more likely to participate in several regulatory processes, activating other components like PKG or PKA (protein kinase G or A), cNMP-dependent phosphodiesterase and CNG channels (**Figure 1.4**)[44]. In prokaryotes, cGMP is involved in binding cyst regulatory elements of *Rhodospirillum centenum*, thus inducing cyst development[45]. Interestingly, cGMP can also interact with effectors which further regulate cyclic di-GMP signaling. This tends to regulate virulence and biofilm formation in the plant pathogen *Xanthomonas campestris*[46, 47].
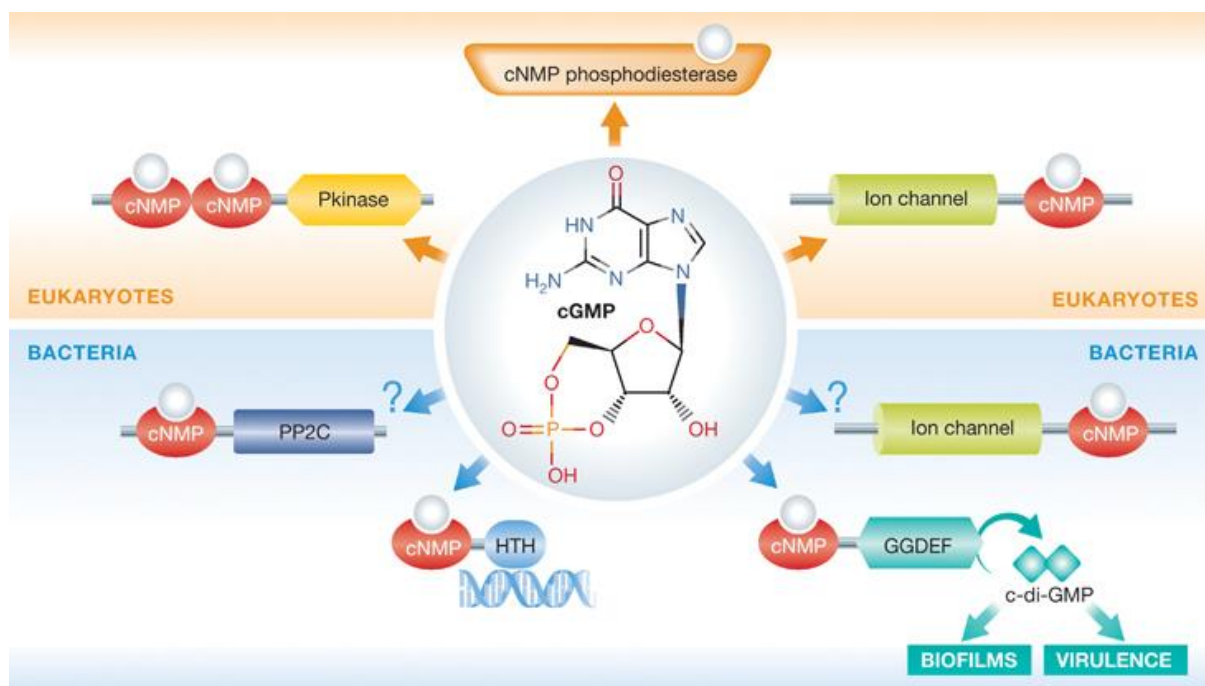
Figure 1.4 Key cGMP receptors in eukaryotes and prokaryotes.

The cGMP-specific cNMP-binding domains based on PF00027 domain analysis in Pfam database. In bacteria, two potential cGMP targets comprise a likely cNMP-gated ion channel BBta_5447 from *Bradyrhizobium sp.* (UniProt: A5EML8_BRASB, PF00520) and a protein serine phosphatase Cagg_2419 from *Chloroflexus aggregans* (UniProt: B8G3B9_CHLAD, comprise PP2C or SpoIIE domain, PF07228). Taken from [44].

Walseth and Goldberg suggested that cGMP or cAMP subcellular levels turned over rapidly due to robust phosphodiesterase activities in many tissues[48, 49]. The rapid turnover was required to uncover physiological consequences from then on. To study these rapid processes about regulation of cGMP or cAMP levels in spatiotemporal subcellular compartments, new and better methods are needed[9]. Therefore, light-induced cyclases or phosphodiesterases could be applied as updated optogenetic tools to regulate cGMP or cAMP levels when explaining related physiological processes.

1.3.2 Enzyme catalytic activities are induced by N-terminal light sensor domains

Considering important roles of second messengers cGMP or cAMP, identification and characterization of light-regulated adenylate or guanylate cyclase will enable us to manipulate cyclic nucleotides levels in *in vivo* experiments at high temporal-spatial resolution.

In 2002, the first photoactivated adenylyl cyclase (PAC) was identified from flagellate *Euglena gracilis*, which mediates photophobic response under blue light illumination[50]. This *Eu*PAC

consists of two subunits named *Eu*PACα and *Eu*PACβ with two BLUF sensor domains and two adenylyl cyclase domains in each subunit, and they can accumulate cAMP after blue light illumination. The PACs were applied to activate protein kinase A to regulate human cystic fibrosis transmembrane conductance regulator (CFTR, as Cl⁻ channel) and to induce to open CNG channels when expressing in *Xenopus* oocytes and HEK293 cells. Meanwhile, it was confirmed to change behaviors by light illumination after transgenic PAC expression in *Drosophila melanogaster*[51]. Another PAC from *Beggiatoa* (BlaC or bPAC) was then identified with a smaller size (350 amino acids) than *Eu*PAC, including only one BLUF domain and adenylyl cyclase domain[52]. With lower dark activity, bPAC tends to be a superior optogenetic tool to manipulate cAMP levels in neural cells or live animals[53].

In 2014, the first enzyme microbial rhodopsin was identified from *Blastocladiella emersonii*, briefly named *Be*GC1, comprising rhodopsin domain in the N-terminal and guanylyl cyclase domain in the C-terminal. It was suggested that *Be*GC1 was likely to play roles in zoospore phototaxis with cGMP accumulation after green light illumination[54]. This opsin was then renamed as BeCyclOp (**Cycl**ase **Op**sin) and applied to manipulate cGMP levels in muscle cells of live *C. elegans* by light illumination. After co-expression BeCyclOp with cGMP-gated heteromeric CNG channel in live *C. elegans*, cGMP can be produced to open CNG cation channel after green light exposure, thus depolarizing muscle cells and causing body contraction[7]. In addition, this light-activated opsin (RhGC) was then confirmed to activate CNG-A2 channel by light illumination after co-expression in Chinese hamster ovary (CHO) cells and rat hippocampal neurons[55]. The following study about RhoGC (=*Be*GC1, BeCyclOp or RhGC) determined that light-induced dimerization of guanylyl cyclase domain increased substrate affinity and enzymatic activity[56]. Later on, the full-length RhoGC protein was expressed and purified in HEK293 cells with absorption spectrum peaking at 527 nm. It was also shown that the enzymatic activity was completely destroyed with a truncated N-terminal domain (17-to 20-kDa deletion)[57]. Interestingly, BeCyclOp was determined with 8-transmembrane helices (8-TM) in rhodopsin domain in two different methods[7, 57]. Overall, BeCyclOp was proven to be a powerful optogenetic tool to rapidly regulate intracellular cGMP levels in cells or animals with green light stimulation.
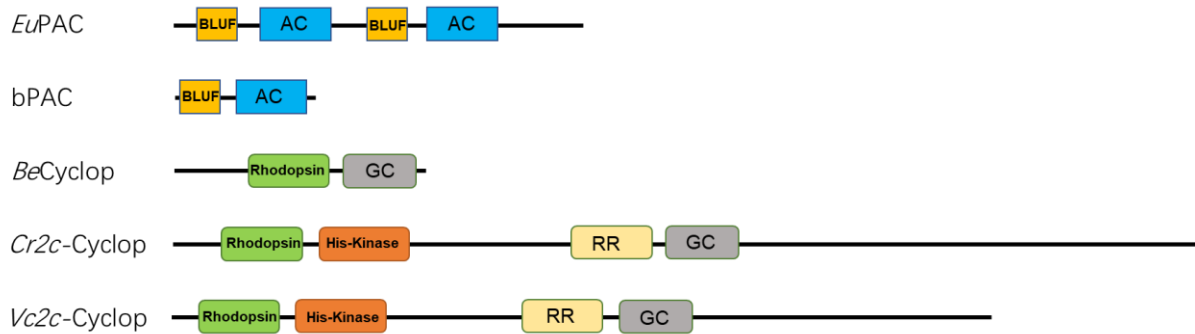
Figure 1.5 Schematic models of light-regulated adenylate or guanylate cyclases.

Colored boxes indicate different domains in each photoreceptor. AC (blue box) means adenylyl cyclase domain, GC (gray box) indicates guanylyl cyclase domain, RR (light yellow box) means response regulator. *Eu*PAC (*Euglena gracilis*), NCBI accession number: AB031225; bPAC (*Beggiatoa*), NCBI accession number: GU461307; *Be*Cyclop (*Blastocladiella emersonii*), NCBI accession number: AIC07007; *Cr*2c-Cyclop (*Chlamydomonas reinhardtii*), JGI accession number: Cre11.g467678; *Vc*2c-Cyclop (*Volvox carteri*), JGI accession number: Vocar20005159m.g.

As sensory photoreceptors, light-regulated enzymes still enrich in microalgae, especially in freshwater algae[58]. In the genome of *Chlamydomonas reinhardtii*, many different photoreceptors were previously named Chlamyopsins1-7 (Cop1 to Cop7)[59]. Cop1 and Cop2, as retinal binding proteins with high homology, were localized in the eyespot organelle of *Chlamydomonas* during growth and cell division. They were proposed to be apoprotein of the phototaxis receptors[60]. But RNAi experiment proved that no photophobic responses and phototaxis were induced by Cop1 or Cop2 protein[61]. Therefore, their functions are still uncertain nowadays. Cop3 and Cop4 showed some homology to microbial opsins, and they were tested as light-gated cation channels (named Channelrhodopsin-1 and 2 separately)[3, 4].

Apart from photoactivated adenylate cyclases (PACs) and light-activated guanylate cyclase opsins (e.g. BeCyclOp), a novel class of enzyme opsins was predicted to exist in *C. reinhardtii* by ESTs and genome database[62]. Several homologous opsins (e.g. COP5, COP6, COP7) are identified in different loci on *C. reinhardtii* chromosomes[63]. This large 2-component cyclase opsins (briefly named *Cr*2c-Cyclops by us) mainly encodes four different domains, including rhodopsin domain in the N-terminal and Guanylyl cyclase (GC) or Adenylyl cyclase (AC) domain in the C-terminal, fused with the two components histidine-kinase (HK) domain and response regulator (RR) domain in the middle (**Figure 1.5**). Although this class of opsin was

predicted existing in green algae for over 12 years, the light-regulated mechanism of 2-component cyclase opsins is still unclear. Until now, only the rhodopsin domain (named "histidine kinase rhodopsin" or "HKR1") from one of *Cr*2c-Cyclops (named "Chalmyopsin-5" or "Cop5") was heterologously expressed, purified and determined with a very slow photocycle between absorption in UV and blue light[64-67].

One part of this thesis focused on the characterization of *Cr*2c-Cyclop (previously named COP6) from *C. reinhardtii* and a homologous protein *Vc*2c-Cyclop from *Volvox carteri*. It was determined that both opsins function as light-inhibited guanylyl cyclases. Action spectra of *Cr*2c-Cyclop and *Vc*2c-Cyclop peak in the green light range (541 nm and 556 nm separately). Interestingly, 2c-Cyclops are very sensitive to light and ATP-dependent.

1.3.3 Two-component systems in prokaryotes and eukaryotes

As members of a novel enzyme opsin family, *Cr*2c-Cyclop and *Vc*2c-Cyclop are distinct from other cyclase opsins due to two additional components between rhodopsin and cyclase domain, suggesting that the two components may play roles in light-induced signal transmission from rhodopsin to guanylyl cyclase.

Two-component systems (TCS) consist of histidine kinases (HK) and response regulators (RR), which can transfer phosphoryl group from the key histidine residue (located in H-box motif of His-kinase domain) to aspartic acid (located in response regulator). The processes of phosphotransfer are widely distributed in most bacteria, and a reduced number of TCS still exist in archaea or several eukaryotes (such as algae, fungi, and plants). However, no TCS in animals or protists was reported. TCS are able to sense and adapt to many environmental stimuli by inducing different signaling responses and establishing regulatory networks. This improves cellular survival, development, and virulence in microorganisms[68]. Interestingly, in evolutionary views, TCS existing in archaea and eukaryotes could be explained as consequences of horizontal gene transfer events from bacteria with numerous two-component systems[69-72].

TCS were well studied in bacteria, for example, many two-component (HK-RR) pairs were characterized in *Escherichia coli*, playing roles in chemotaxis, compounds or ions transport,

genome transcription etc.[73]. One typical HK-RR pair is EnvZ/OmpR in *E. coli*, involving in the osmoregulation of porins. A small inner membrane protein (MzrA) was determined to modulate the EnvZ/OmpR system[74]. Considering the signal transduction mechanism of TCS, another representative HK853 and its cognate RR468 (from *Thermotoga maritima*) were determined by crystal structures. The histidine kinase domain can be further divided into two-helix dimerization and histidine-containing phosphotransfer (DHp) domain and C-terminal catalytic and ATP-binding (CA) domain (**Figure 1.6**). HK853-RR468 complex included one HK853 dimer and two RR468 monomers[75]. Interestingly, this study also demonstrated that phosphorylated RR468 (P~RR468) can be dephosphorylated quickly with HK853 wild type, but P~Asp of P~RR468 was highly stable without HK853. Further researches demonstrated that the H-box motif of histidine kinase DHp domain played dual roles in autophosphorylation of histidine residue as well as dephosphorylation reactions. The phosphatase activity was normally induced by the fourth amino acid Threonine/Asparagine after histidine residue (e.g. HK853 with H260 as autophosphorylation site, T264 as key phosphatase site)[76, 77].
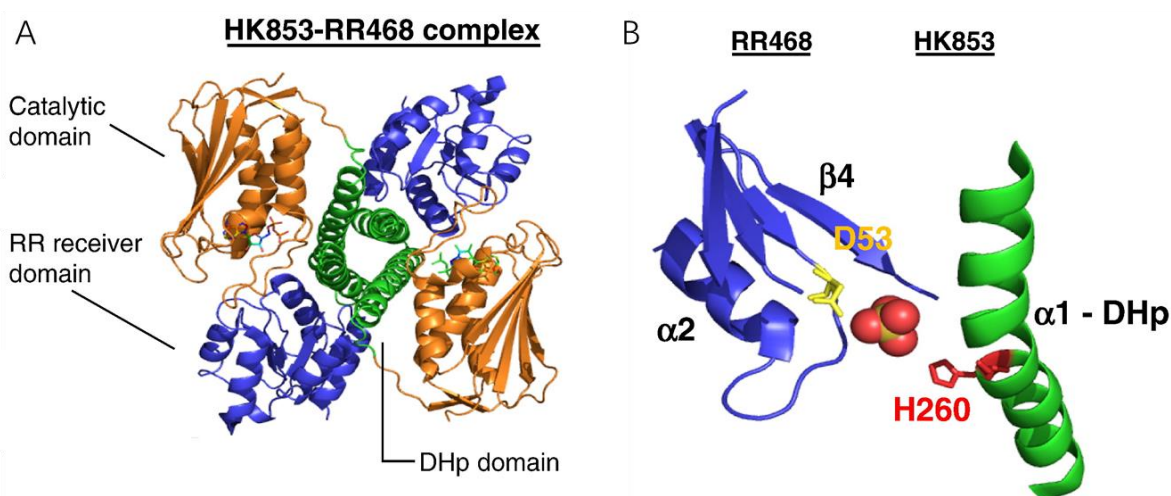
Figure 1.6 Structure complex and phosphotransfer of the HK853-RR468 pair.

A, The complex formation of HK853-RR468 pair is depicted. DHp domain, kinase catalytic domain and response regulator are depicted in green, orange and blue color separately.

B, Close view of the interaction between DHp domain and RR domain. Phosphotransfer sites are located in close distance between H260 (Red) and D53 (yellow) after complex formation. PDB ID: 3DGE was used. Modified from [68].

Based on the alignment between two components (HK-RR) of 2c-Cyclops and above two typical HK-RR pairs (HK853/RR468, EnvZ/OmpR), see **Figure 3.15**, two components shared some conserved sequences especially with H-box and G-box motifs. Therefore, further mutagenesis of key residues was applied in my study to address how phosphoryl transfer has effects on C-terminal guanylyl cyclase activity under light and dark treatments. It is suggested that light illumination tends to interrupt phosphoryl transfer between two components, similar effects as *Cr*2c-Cyclop H352F, G533A or D1092T mutants (block phosphorylation).

1.3.4 Coiled-coil linkers in light-regulated enzymes

In sensory photoreceptors, light simulates chromophore chemical changes to cause conformational alterations of photosensors, which further transduce signals to effector domains. The mechanism about how to transmit signals between sensors and effectors is still unclear.

Coiled-coils (CC), as essential connectors, play roles in diverse functions, such as signal transduction, membrane fusion, and motors. Generally, coiled-coil motifs share similar characters with repeated seven-residue, but one or more residues insertions are still accommodating different signaling states[78]. Signaling transmission via coiled-coils has been studied in recent years, which could further explain not only the mechanisms of light-induced enzymes but also the available generation of chimeric photoreceptors.

Andreas Möglich et al. demonstrated that coiled-coil was applied to design light-regulated histidine kinase successfully. They reprogrammed a new signaling cascade with an exogenous light-oxygen-voltage (LOV) photosensor domain replacing a natural chemosensor domain to control output histidine kinase activity. A blue light sensor LOV came from *Bacillus subtilis* YtvA (YtvA), while histidine kinase domain was taken from *Bradyrhizobium japonicum* (FixL) without N-terminal heme-binding PAS sensor domain[79, 80]. J-α linkers in both natural proteins were modified in certain points as an integrated α-helical coiled-coil linker. Finally, they determined that one synthetic light-regulated kinase (named YF1) activity was reduced by ~1000-fold *in vitro* after light-illumination. Moreover, YF1 can control gene expression with light stimulation *in vivo*[81]. The crystal structure of engineered YF1 was determined that the α-helical coiled-coil linker between LOV photosensor and effector module plays key roles in

modulating signal transduction by light-stimulated α-helix rotation within microseconds (**Figure 1.7**). This rearranges kinase domain by influencing interface between the DHp and CA domains, thus simulating working models as many other natural photoreceptors[82, 83].
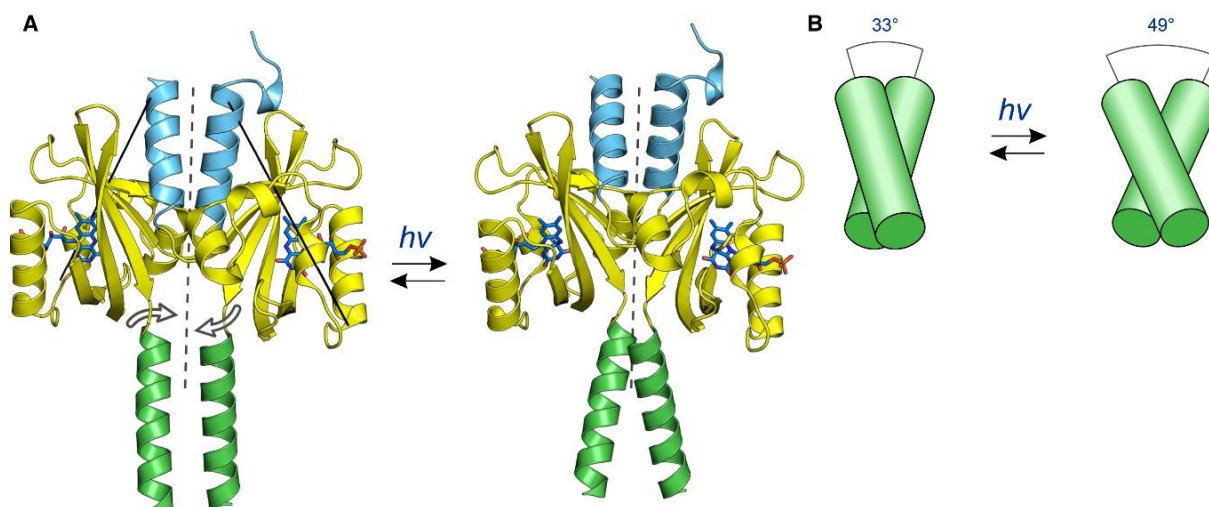
Figure 1.7 Structure-Based Model for Signal Transduction.

A, In dark stage, photosensor LOV domain of YF1 was determined by x-ray crystallography (left) which was compared to a homology model in light stage (right) about crystallized PpSB1-LOV protein from *Pseudomonas putida* [84]. Upon light illumination, rearrangements of the quaternary structure will cause a supertwist of C-terminal coiled-coil (arrows). Both structural models were aligned based on LOV core domains (25 aa-126 aa). The dashed line means the dimer axis, and the solid line means screw-rotation of LOV core domains.

B, Jα helices around the dimer axis could change the crossing angle from 33° to 49°, and this tends to make a coiled coil with a left-handed supertwist. Taken from [82].

On the other hand, to demonstrate signal transduction roles of coiled-coils, two natural photoactivated adenylate cyclases (PACs) have been crystallized, one from *Oscillatoria acuminata* (OaPAC) and the other from *Beggiatoa* sp. (bPAC). It was suggested that light stimulation makes the central coiled-coil transmit signals from photosensor domain to effector domain, which is mainly induced by structural fluctuations between the BLUF core and $\alpha3_{BLUF}$[85, 86]. Later on, a rhodopsin guanylyl cyclase from *Catenaria anguillulae* (CaRhGC) was demonstrated as a light-activated cyclase opsin (L/D activity ratio > 1000). They also assumed that the coiled-coil motif transmits a signal from a light-induced conformational change of rhodopsin to cyclase domain[87].

To obtain light-regulated enzymes, new chimeric photoreceptors could be engineered based on characters of coiled-coil sequences. For example, in N-terminal of the PDE2A catalytic domain from *Homo sapiens* cAMP/cGMP-specific PDE2A (*Hs*PDE2A), a coiled-coil sequence was fused to C-terminal of the two sensor GAF-domains. The photosensor module (including PAS, GAF, PHY domains) from *Pseudomonas aeruginosa* was connected with PDE2A catalytic domain in a central α-helical coiled-coil peptide sequence. This achieved a light-regulated phosphodiesterase (LAPD) which can be activated by ~6 fold under near-infrared light illumination[88]. In the future, it will be more valuable if the large photosensor module can be replaced by a vivid LOV domain, connecting *Hs*PDE2A with J-α linker. Other chimeric enzyme opsins could also be obtained by integrating rhodopsin domain with catalytic domain according to predicted coiled-coil modification.

1.4 cGMP, cAMP hydrolysis by light-regulated phosphodiesterase

Cyclic Nucleotide (cN) Phosphodiesterases (PDE) are widely distributed in the animal kingdom, hydrolyzing 3' phosphate bond of cyclic nucleotides (cGMP or cAMP). Three classes of catalytic domains can be identified in unicellular organisms such as protozoa (in class-I), fungi, slime mold, amoebae (in class-II) and *Dictyostelium discoideum* (in class-III). However, class I catalytic domain mainly exist in metazoan[89]. Until now, PDE enzymes are divided into 11 families in mammals, which can hydrolyze cAMP and/or cGMP depending on their catalytic selectivity[90]. As ubiquitous second messengers, cAMP and cGMP play essential roles in cell division, differentiation, growth, and death. Therefore, PDE enzymes are crucial to manipulate concentrations of both second messengers to maintain normal responses. In N-termini of most PDE family members, different regulatory modules have diverse functions such as ligand binding, oligomerization, and kinase recognition/phosphorylation. Some typical PDE family members are described in **Figure 1.8**, comprising certain regulatory modules and cAMP and/or cGMP specific PDE catalytic domains.

To identify or engineer light-regulated PDE is another important project nowadays because many inherited diseases (including disruption of endocrine cells, smooth muscle cells, hippocampus, cardiac myocytes, immune responses) are closely related with dysfunctional PDE enzymes[89]. Such optogenetic tools could manipulate concentration of cAMP and/or

cGMP by light to uncover physiological impacts and biochemical mechanisms in certain different tissues or subcellular compartments.
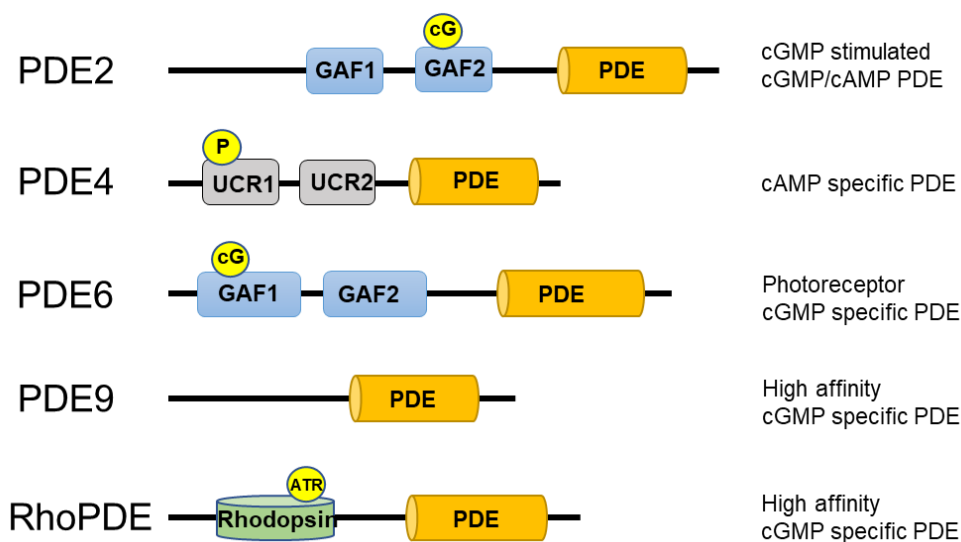
Figure 1.8 Models of some typical PDE family members.

All PDE catalytic domains locate in the C-terminal of proteins, specifically hydrolyzing cAMP and/or cGMP. Regulatory modules include GAF, UCR (upstream conserved regulatory), and Rhodopsin, which can be stimulated to regulate PDE activity by cGMP binding (encircled cG), phosphorylation (encircled P) and all-*trans*-retinal (ATR) isomerization (encircled ATR), respectively.

The first engineered light-regulated PDE consisted of a photosensor from *Pseudomonas aeruginosa* and the PDE2A catalytic domain from *Homo sapiens*, which was determined as a light-activated phosphodiesterase (LAPD)[88]. In addition, the first natural microbial phosphodiesterase from the genome of the protist *Salpingoeca rosetta* (named Rh-PDE[91] or RhoPDE[92]) was discovered with N-terminal rhodopsin domain and C-terminal PDE domain. Rh-PDE was expressed in HEK293 cells and a slightly light-activated PDE effect was observed. Illumination increased PDE hydrolytic activity for cGMP and cAMP as substrates by ~1.4 fold and ~1.6 fold separately. It was suggested that the slightly light-activated PDE showed ~10 fold more active with cGMP than with cAMP. The full-length of Rh-PDE was purified and the maximal absorption spectrum located at 492 nm[91]. However, Lamarche et al. did not observe the light-regulated activity of RhoPDE. The crystallized PDE domain formed as a dimer, similar to human PDE9. Interestingly, they determined that there is an even number of transmembrane helices (TMs) in rhodopsin domain by immunofluorescence microscopy experiments[92]. This *Sr*RhoPDE was further determined by us that it is a clearly light-activated phosphodiesterase

and shows light-enhanced substrate affinity. Light can cause a strong decrease of K_m value for cGMP from 80 to 13 μM, but only increase the maximum turnover by ~30%. We also confirmed that there is an even number of TMs (probably 8-TM based on alignment with classical 7-TM opsins) in the rhodopsin domain by BiFC (Bimolecular fluorescence complementation) experiment[8]. We anticipate that *Sr*RhoPDE could be further improved by genetic engineering as optogenetic tools for light-regulated cGMP manipulation.

1.5 Light-induced modules in association, dissociation or gene expression

Optogenetic tools start with light-gated channels or pumps, which have been widely applied and transformed in neuroscience. In recent years, the second generation of optogenetic tools continue with light-regulated guanylyl or adenylyl cyclases or phosphodiesterase, and this could be used to manipulate second messengers (cAMP and/or cGMP) levels *in vivo* and uncover regulatory signaling cascades behind. Interestingly, the third generation of optogenetic tools, as a non-neuronal toolbox, were identified and engineered to obtain light-induced dimerization, oligomerization as well as gene expression in subcellular levels. Applications of optogenetic tools possess many advantages, such as available fast kinetics of light-induced reversible systems, endogenous chromophores (e.g. flavin, retinal, biliverdin) existing in most mammalian cells, and precise manipulation of intracellular signals in space and time[93].

In the blue spectral range, a rapidly blue-light-induced dimerization in living cells was determined based on cryptochrome 2 and CIB1 from *Arabidopsis thaliana*. This pair was then optimized with light-induced minimal interaction domains of CRY2-CIB1 by yeast two-hybrid assay. This optimized system was then applied in protein translocation of CRY2 induced by light, light-regulated transcription and DNA recombination[94]. Another typical blue light absorbed sensor is EL222 model system (222 amino acids) from *Erythrobacter litoralis*, including Light-oxygen-voltage (LOV) domain and HTH (helix-turn-helix) DNA binding domain identified by *in vivo* ChIP-Seq (chromatin immunoprecipitation and high-throughput sequencing)[95]. This EL222 system was then modified as a light-induced transcription tool in many eukaryotic cells as well as zebrafish embryos. It was reported that the system has rapid kinetics with activation (<10 s) and deactivation (<50 s)[96]. Interestingly, this EL222 system can also be applied to improve metabolic engineering in *Saccharomyces cerevisiae* for the

production of valuable products, like isobutanol, by regulating fermentation only with light[97].

On the other hand, in the red spectral range, a bacterial phytochrome DrBphP from *Deinococcus radiodurans* can be engineered as infrared-fluorescent proteins (IFPs), with maximal excitation at 687 nm and emission at 708 nm. IFPs with ubiquitous biliverdin as the chromophore can be applied in *in vivo* imaging because red light is able to penetrate into deep tissues[98]. To control cell signaling in subcellular level, a natural pair of BphP1-PpsR2 optogenetic system was identified from *Rhodopseudomonas palustris*. It was demonstrated that near-infrared (NIR) light (used at 740 nm) can induce their heterodimer formation and relocate one partner to the plasma membrane (**Figure 1.9**) or nucleus. This NIR system can be utilized to regulate cell morphology or induce gene expression in cultured cells or mouse tissues[99].
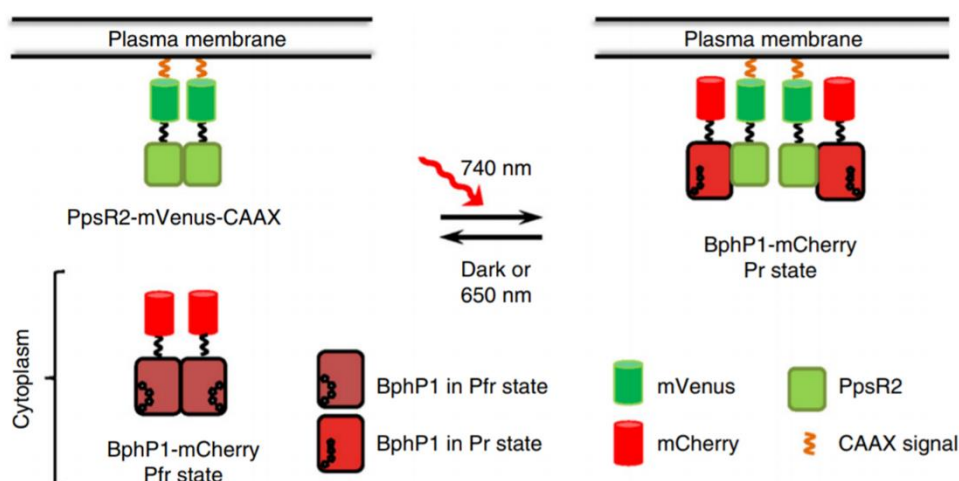
Figure 1.9 A model about recruitment of BphP1 to the plasma membrane by NIR illumination.
PpsR2 was targeted into plasma membrane with CAAX signal peptide, fusion with mVenus tag. The BphP1 was expressed in cytoplasm fusing with mCherry tag. Modified from [99].

Later on, the same research group further modified one component PpsR2 in a minimal size named Q-PAS1 which can still interact with its partner BphP1 under near-infrared light. Interestingly, this BphP1-Q-PAS1 NIR system can be applied with a blue-light-activated LOV-domain-based system without spectral crosstalk, thus engineering multicomponent systems to regulate tri-directional subcellular protein targeting (in the plasma membrane, cytoplasm or nucleus) by blue (used at 460 nm) or NIR illumination (used at 740 nm)[100].

1.6 Objectives

Optogenetic tools are explored successfully in single-domain channelrhodopsins or pumps, which are widely applied in neuroscience. Another part is about identification and characterization of sensory photoreceptors, which integrate two or more domains to regulate output enzymatic activities (e.g. light-regulated adenylyl or guanylyl cyclases, light-stimulated phosphodiesterase). Manipulating second messengers (cAMP and/or cGMP) levels by light enables us to uncover related signaling cascades in many organisms.

In my study, I mainly characterized two different microbial enzyme rhodopsins. One was *Cr*2c-Cyclop from *C. reinhardtii*, integrating two components between rhodopsin domain and guanylyl cyclase domain. Until now, it is still unclear about properties of this enzyme rhodopsin. My results confirmed that *Cr*2c-Cyclop is a novel light-inhibited guanylyl cyclase, which is different with light-activated enzymes such as bPAC, BeCyclOp. Therefore, the mechanism of inhibition effects needs to be further uncovered. Based on single site-directed mutagenesis in key residues, phosphorylation transfer between two components plays an important role in activating guanylyl cyclase, while light illumination tends to interrupt phosphorylation processes and dramatically inhibit cyclase activity. Overall, this project uncovered the mechanism of a novel enzyme opsin with additional phosphorylation steps.

The other interesting enzyme opsin, rhodopsin phosphodiesterase *Sr*RhoPDE, was identified from a protist *Salpingoeca rosetta*. Previous two publications presented some conflicting views on whether or not this enzyme opsin possesses light-regulated influence. It is clearly explained by us that *Sr*RhoPDE is a novel light-activated phosphodiesterase, and light can enhance substrate affinity.

Interestingly, light-activated GC (Cyclop), light-inhibited GC (2c-Cyclop), and light-activated PDE (RhoPDE) somehow resemble the more sophisticated light-regulation of cGMP by type-II rhodopsin in the mammalian vision system judging from the functional output. In addition, based on signaling transduction modules, it is also possible to engineer novel photoreceptors as optogenetic tools with similar linker modification in the future.

2. Materials and Methods

In this part, all chemicals come from Sigma (Deisenhofen), Fluka (Neu-Ulm), Applichem and Roth if no special labeling.

2.1 DNA Manipulation in *E. coli*
2.1.1 *E. coli* strains for cloning and overexpression

To generate vectors with targeted DNA fragments, *E. coli* MRF strain was applied to produce enough copy numbers of plasmids efficiently after transformation of pGEMHE vector. Before transformation, *E. coli* strain was treated by a certain concentration of $CaCl_2$ solution to generate competent *E. coli* cells due to enhancing DNA penetration. Then exogenous vectors can be transformed into *E. coli* MRF by heating or electric shock treatment. The transformed *E. coli* MRF was then put in an LB agar plate with anti-biotic ampicillin or kanamycin depending on a resistant marker gene on the vector. A certain amount of plasmid can be extracted from a certain volume of cultured *E. coli* MRF from a positive colony. To overexpress targeted proteins, *E. coli* BL21(DE3) chemical competent cells were transformed by a PET-28b vector with targeted DNA fragments. With 0.5 mM-2 mM IPTG induction, the T7 RNA polymerase can be recruited to the promoter site and the targeted gene was then overexpressed.

2.1.2 DNA cloning

pGEMHE vector with multiple cloning sites (MCS) was applied to ligate targeted DNA fragments. For *Cr2c*-Cyclop cloning methods, RNA was isolated from *Chlamydomonas reinhardtii*, and then the cDNA was generated from RNA by reverse transcription PCR (RT-PCR) in previous experiments from our lab. Finally, the DNA can be amplified from cDNA without introns for further protein expression. The PCR amplification system is shown in **Table 2.1**. The full-length of *Cr2c-Cyclop* fragment (5'-BamHI-*Cr2c-Cyclop*-XhoI-HindIII-3') was ligated in pGEMHE vector with BamHI in 5' cloning site and HindIII in 3' cloning site. Further 5'-XhoI-YFP-HindIII-3' fragment was fused in the 3' cloning site of 5'-BamHI-*Cr2c-Cyclop*-XhoI-HindIII-3' fragment (no stop codon). To generate single mutants of *Cr2c*-Cyclop (including K298A, H352F, G533A, D1092T), the individual mutation site was designed in a

pair of primers with unique restriction sites. After PCR processes, certain fragments can be further ligated into pGEMHE vector with correspondingly unique restriction sites. The **Supplementary Table 1** illustrates the primers used in all experiments.

For *Sr*RhoPDE, the DNA fragment was synthesized by GeneArt Strings DNA fragments (Life Technologies, Thermo Fisher Scientific). The sequence information comes from NCBI gene database (XP_004998010.1), with BamHI restriction in 5' cloning site, HindIII restriction in 3' cloning site. The codon was optimized to *Mus musculus*. Then the RhoPDE DNA fragment was inserted into the PGEMHE vector with 5' BamHI and 3' HindIII digestion sites. The YFP::RhoPDE construct was generated by inserting two fragments into above PGEMHE vector, that is, 5'-BamHI-YFP (no stop codon)-XhoI-3' and 5'-XhoI-RhoPDE-HindIII-3'. In contrast, RhoPDE::YFP was generated using 5'-BamHI-RhoPDE (no stop codon)-XhoI-3' and 5'-XhoI-YFP-HindIII-3'. The BiFC-Opsin construct was made by inserting opsin part of RhoPDE sequence without stop codon to the BiFC (Bimolecular Fluorescence Complementation) vector, with KpnI and XhoI as restriction sites. Single mutants were generated by the Quickchange PCR Method with primer pairs presented in **Supplementary Table 1**.

All the constructs were confirmed by complete DNA sequencing. NheI restriction enzyme was used to linearize all the plasmids. After DNA gel electrophoresis to confirm single linear DNA band, the linearized DNA was then used for *in vitro* transcription to obtain cRNA with the AmpliCap-MaxT7 High Yield Message Maker Kit (Epicentre Biotechnologies).

Table 2.1   Reaction system of PCR amplification

Template	~20 ng (diluted to ~1 μL)
5× DHG buffer	8.0 μL
dNTP (10 mM)	1.6 μL
Forward primer (10 μM)	1.6 μL
Reverse primer (10 μM)	1.6 μL
BSA (100 mg/mL)	0.4 μL
Phusion DNA polymerase (2 U/μL)	0.3 μL
Add ddH$_2$O to final 40 μL	

Steps of PCR amplification show as below:

Step 1:   98 °C   50 s

Step 2:   98 °C   15 s, 66 °C   15 s, 72 °C 15 s/kb; 6 Cycles

Step 3:   98 °C   15 s, 62 °C   15 s, 72 °C 15 s/kb; 6 Cycles

Step 4:   98 °C   15 s, 58 °C   15 s, 72 °C 15 s/kb; 20 Cycles

Step 5:   72 °C 2~5 min, 10 °C ∞

Phusion High-Fidelity DNA Polymerase (F530-L, Life Technologies) was used in the PCR reaction system, and the PCR program was operated in a Piko Thermal Cycler (Thermo Scientific). PCR products were loaded with an aliquot of 5× DNA loading buffer for 1% DNA agarose gel electrophoresis. DNA agarose gel was prepared with 0.005% GelGreen (Biotium, Hayward, USA), and DNA bands were visualized by Imagemaster VDS (Pharmacia, Uppsala, Sweden). PCR products or gel slices can be frozen under -20 °C for future use.

To generate accurate chimeric photoreceptors, overlapping PCR was applied as an efficient strategy to fuse two different fragments in a certain junction. The actual overlapping PCR processes in my experiments were modified based on some cloning strategies[101]. Firstly, the two pairs of primers (labeled as No. 1, 2; 3, 4) for two fragments were designed for the normal PCR process as above, then two PCR products were separately purified with PCR purification kit and eluted with an aliquot of ddH$_2$O. Secondly, a new PCR program was designed with above two PCR products as templates (~100-300 ng for each), and primers were not applied in this mixture reaction system. The annealing step was increased to 1 min at 68 °C, and extension step was adjusted to 2 min at 72 °C based on DNA length. If the fused fragment was detected in a correct DNA size, gel purification would continue. Finally, the two end primers (No. 1 and 4) were used to do normal PCR with eluted targeted fragment as the template (**Figure 2.1**). The single band was obtained at the correct DNA size.
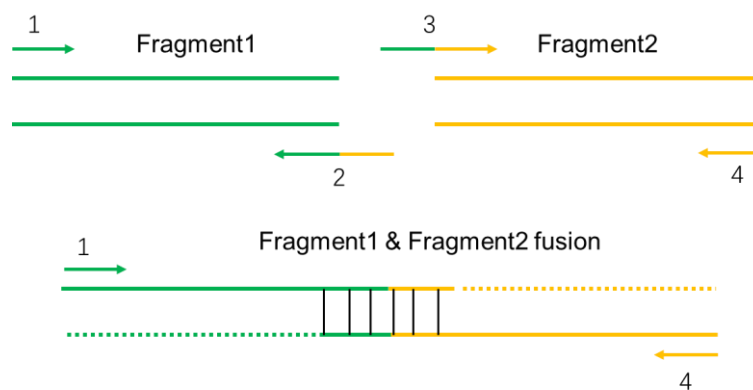
Figure 2.1 Processes of overlapping PCR.

Two different fragments can be amplified by primers 1 and 2, 3 and 4 separately. The final PCR program needs two end primers 1 and 4 to obtain a final fusion construct.

2.1.3 DNA purification

DNA fragments were purified with two methods, including PCR purification and gel purification.

PCR purification would be applied if the targeted DNA fragment was pure as a single bind after gel electrophoresis. The protocol was following QIAquick PCR Purification Kit (Qiagen, Hilden, Germany). ddH$_2$O was used in the final step to elute DNA fragment from QIAquick column. The purified DNA concentration was measured by NanoPhotometer, and then stored at -20 ℃ for future use.

The other way is gel purification. This protocol was following QIA quick Gel Extraction Kit (Qiagen, Hilden, Germany). PCR products or digested vectors were firstly loaded in the DNA agarose gel to do electrophoresis, and the targeted DNA band was cut out from gel under blue light illumination. The gel slice was then dissolved by QG buffer from the kit at 50 ℃ for further processes. ddH$_2$O was used in the final step to elute DNA fragment from QIAquick column. The concentration of extracted DNA was measured by NanoPhotometer, and then stored at -20 ℃ for future use.

2.1.4 DNA digestion and ligation

After extraction of targeted DNA fragments, they were next digested by certain restriction

enzymes (Fermentas, Thermo, Waltham, USA) for further ligation with correspondingly restricted vectors. Aliquots of restricted DNA fragments were then purified by above PCR purification protocol. DNA marker was using Lambda PstI (Gibco/Invitrogen, Carlsbad, USA).

The ligation system contains: a mixture of ~50 ng vector and certain amount (ng) of DNA fragments based on the molar ratio (Vector/DNA fragment: 1/10~1/5), 1× T4 ligation buffer, T4 DNA ligase 0.5 μL (Fermentas, Thermo, Waltham, USA), ddH$_2$O was added to final 10 μL or 20 μL. The mixed aliquot was incubated at ~22 °C for 30~60 min, and then it was transfected in competent *E. coli* cells.

2.1.5 Chemical competent *E. coli* preparation and transformation

Competent *E. coli* was prepared by chemical treatment. The processes were operated as followings: stored competent *E. coli* was taken, put 10 μL in 5 mL LB liquid medium without antibiotic, shaking (~200 rpm) overnight in a constant 37 °C shaker (G25 Incubator, New Brunswick Scientific). 1 mL pre-cultured medium was transferred into 1 L new LB liquid medium without antibiotic for further culture, shaking (~200 rpm) until OD value located in a range of 0.3~0.5. Then 1 L cultured medium was divided into 50 mL sterile falcon tubes and put them on ice to cool down for ~10 min, 4000 rpm centrifuge at 4 °C for 10 min, discarded supernatant. In each 50 mL tube, the pellet was suspended with 10 mL pre-cooling 0.1 M CaCl$_2$ solution, and incubated on ice for ~30 min, 4000 rpm centrifuge at 4 °C for 10 min again, discarded supernatant. The pellet in each tube was resuspended with 2 mL 0.1 M CaCl$_2$ solution (including 15% Glycerol), put on ice for ~5 min, and finally separated in 1.5 mL tubes with 50 μL for each, froze in liquid N$_2$, and then stored at -80 °C.

After ligation of restricted DNA fragment and vector, an aliquot of the volume was added in above ~50 μL competent *E. coli*, kept on ice for ~15 min, and then incubated at 42 °C for 45 s-90 s, put on ice for ~2 min. Certain mixed volume was put in a solid LB plate with an antibiotic depending on the antibiotic marker in vectors (e.g. ampicillin or kanamycin), and then cultured in an incubator at 37 °C overnight.

2.1.6 Screening for positive bacterial colonies

To detect positive colonies, there were two ways in my experimental processes.

One way was that a certain number of colonies were taken out for cracking. The 2× cracking solution contained: 0.2 M NaOH, 0.5% SDS, 20% sucrose. Colonies were marked on the plate and then taken into correspondingly numbered tubes with 8 µL sterile ddH$_2$O for each. After mixture with colonies, 8 µL 2× cracking buffer was then added in each tube, incubated for ~10 min at room temperature, and loaded 5× DNA loading buffer in each tube. The mix was centrifuged at 13,000 rpm for 2 min, and loaded in 1% DNA agarose gel to check the super-coiled plasmids. The correct plasmids with inserted fragment normally showed larger DNA bands than incorrect plasmids without targeted DNA segment insertion. The positive colonies were then taken to culture in ~6 mL LB liquid medium in 15 mL falcon tubes (SARSTEDT, Nümbrecht, Germany) with a required antibiotic, shaking at 37 °C overnight.

In addition, positive colonies can be selected by colony PCR. The processes are similar with normal PCR, but the template is replaced with an individual colony of *E. coli*. The first PCR step is changed to heat at 98 °C for 5 min, due to denaturing DNA double helix more efficiently from transformed *E. coli*. PCR products were loaded in a DNA agarose gel, and colonies with right DNA size were then taken to culture in ~6 mL LB liquid medium (+Ampicillin or kanamycin), shaking at 37 °C overnight.

2.1.7 Plasmids extraction and restriction test

The above-cultured medium from positive colonies was centrifuged at 4,000 rpm for 10 min. The supernatant was discarded, then plasmids were extracted from pellet by using the protocol from Qiaprep spin miniprep kit (Qiagen, Hilden, Germany). Finally, DNA was eluted with ~50 µL sterile ddH$_2$O in each column. Concentrations of plasmids were detected by NanoPhotometer (NanoPhotometer™, Implen, Munich, Germany). dsDNA absorbed at 260 nm, and A260/280 ratio for good quality of DNA was located at around 1.8. The extracted plasmids were detected by certain restriction enzymes and sent to GATC Biotech for sequencing (GATC Biotech, Constance, Germany). Correct plasmids were stored at -20 °C for future use.

2.2 cRNA (complementary RNA) generation and quality evaluation

2.2.1 *In vitro* transcription

After obtaining correct DNA sequences, NheI digestion enzyme was used to linearize plasmids. Linearized plasmids were finally eluted with ~15 μL RNase-free H_2O and 1 μL from each sample was detected by DNA agarose gel electrophoresis. The *in vitro* transcription was done by mixing linearized plasmid in certain aliquots from AmpliCap-MaxT7 High Yield Message Maker Kit (Epicentre Biotechnologies), as depicted in **Table 2.2**. The 10 μL mixture for each sample was then incubated at 37 °C for ~3 h. After incubation, 12 μL Ammonium acetate (CH_3COONH_4) was added to the above reaction mix, and then frozen at -20 °C for at least 1 h. The mixture was centrifuged at 4 °C, ~20,000 × g for 1 h. The pellet was washed in 50 μL 70% ethanol (prepared in DEPC H_2O or RNase-free H_2O) and centrifuged for another 1 h, then the supernatant was discarded and the pellet was dried in a fuming cupboard. Finally, 10-15 μL RNase-free H_2O was added to dissolve the pellet. RNA concentration was detected by NanoPhotometer, and A260/280 ratio for good quality of RNA was located at around 2.0. Then samples were frozen at -20 °C for future use.

Table 2.2   *In vitro* transcription

AmpliCap-Max Cap/NTP PreMix	4.0 μL
DTT (Dithiothreitol) (100 mM)	1.0 μL
10× AmpliCap-Max™T7Transcription Buffer	1.0 μL
AmpliCap-Max™ T7 Enzyme (5 U/μL)	1.0 μL
RNase Inhibitor (20 U/μL)	0.3 μL
Add Linear DNA (~600 ng) to final 10 μL	

2.2.2 Determination of RNA quality

After *in vitro* transcription, the quality of RNA samples was evaluated in two ways. Firstly, the purity of RNA samples was estimated by the ratio of absorbance between nucleic acid and contaminants by using NanoPhotometer. For example, A260/A280 ratios normally locate in a

range of 1.8-2.2, while A260/A230 ratios are normally >2.0.

Secondly, RNA samples were detected in a denaturing gel system due to the intramolecular base pairing formation. RNA gel preparation was applying MOPS buffer system based on the protocol from ThermoFisher Scientific. For example, 100 mL 1% RNA agarose gel included 1 g agarose with 90 mL ddH$_2$O boiled, after cooling down at ~60 °C, 10 mL 10× MOPS was added in the running buffer (0.4 M MOPS pH 7.0, 0.1 M Na-Acetate, 0.01 M EDTA), and the 37% formaldehyde was added ~20 μL as well as 4 μL Gel Red as fluorescent nucleic acid dye. A small aliquot of RNA sample was prepared for RNA gel electrophoresis, consisting of 1 μL RNA, 5 μL 2× Formaldehyde Loading Dye (including 95% formamide, 0.025% SDS, 0.025% bromophenol blue, 0.025% xylene cyanol, 0.025% ethidium bromide, 0.5 mM EDTA), 4 μL RNAase-free water or DEPC water to final 10 μL. To denature RNA secondary structure, RNA samples were incubated at 72 °C for 20 min, then put on ice for ~2 min and loaded in an RNA agarose gel.

2.3 Manipulation of protein expressing in *Xenopus* oocytes

Xenopus oocytes were applied as a convenient system to express exogenously targeted proteins. A large number of membrane proteins like channels or pumps from microorganisms or plants can be efficiently expressed in oocytes and measured by electrophysiology, while other membrane-bound or cytosolic enzymes can also be characterized by extracting membranes or cytosol from oocytes and doing *in vitro* tests.

2.3.1 *Xenopus laevis* operation for oocytes

Unfertilized oocytes were harvested from female *Xenopus laevis*, and each *Xenopus laevis* was operated every ~6 months. During surgical processes, some instruments like scalpels and surgical scissors were sterilized in advance.

The frog was anesthetized in a few liters of ice-cooled water with 1 g/L tricaine methane-sulfonate (MS-222) for 30 min, and then put in a tricaine solution soaked tissue paper on the ice, with abdomen on the top side. About 1 cm long incision was operated under the muscle in

the abdomen skin by scalpel and scissor. Oocytes were stripped out from the ovary of the frog and cut by scissor. In 50 mL falcon tube, ND96 buffer (96 mM NaCl, 2 mM KCl, 1 mM CaCl$_2$, 1 mM MgCl$_2$, 5 mM HEPES and pH 7.4) without Ca^{2+} was used to put oocytes with 50 µg/mL gentamycin. The incision was sewed by Surgical Medical Monofilament Non absorbable Nylon Thread, and then the frog was put back into fresh water to recover.

Oocytes were then washed two times by ND 96 buffer (without Ca^{2+}) and then digested by 5 mg/mL collagenase for ~1.5 h. Finally, oocytes were washed by ND96 buffer (with Ca^{2+}) few times to make it clean enough and put in culture dishes at an 18 °C incubator for next step.

2.3.2 Exogenous proteins expression in *Xenopus* oocytes

RNA microinjection in *Xenopus* oocytes was operated by Drummond Nanoinjector (Drummond Scientific, Broomall, PA, USA). A vertical Puller (PP-83; Narishige) was employed to generate injection capillary (3.511 Drummond #3-000-203-G/X; Drummond Scientific Company). To express exogenous proteins in *Xenopus* oocytes, certain amounts of RNA were injected in oocytes for individual tests. The amount of RNA injection was adjusted to less than 40 ng in total, while the injected volume was regulated no more than 50.6 nL in each oocyte. To co-express two or more constructs together, certain amounts of RNA samples were mixed in designed ratios and injected only at one time. After complimentary ribonucleic acid (cRNA) microinjection in oocytes, oocytes were then incubated in ND96 buffer for 2-3 days in an 18 °C incubator. If light-sensitive opsins were compared with or without all-*trans*-retinal (ATR), additional 1 µM ATR was then directly prepared or not in ND96 buffer. Oocytes expressing enzymes can be frozen in liquid N$_2$ first and stored at -80 °C for further tests.

2.3.3 Extraction of the membrane or cytosolic proteins from *Xenopus* oocytes

If a targeted protein was fused with YFP tag, after 2 or 3 days post-injection (dpi), oocytes would be checked fluorescence first under blue light illumination, and green light can be observed in the surface of oocytes with certain absorptive filters. Oocytes with obvious fluorescence were selected out for membrane extraction processes afterward. Solution A was prepared to extract oocytes membrane or cytosol as shown in **Table 2.3**.

Table 2.3 Solution A for extraction

NaCl	100 mM
Tris-HCl	75 mM
DTT	5 mM
$MgCl_2$ (add or not as indicated)	5 mM
Glycerol	add to 5%
Adjust pH with HCl to 7.4	

1) Oocytes membrane extraction processes

To test the activity of membrane-bound enzymes, membranes from oocytes were extracted for further *in vitro* reaction. The procedures for oocytes membrane extraction were displayed as follows (all the processes were operated on ice and 4 ℃ centrifugation):

Step 1: A certain number of *Xenopus* oocytes were homogenized with 500 μL solution A in individual experiments. Vortex for 30 s and put it on ice for 30 s, repeated 2-3 times until samples were homogenous enough. The first centrifugation was used at 500 × g for 15 min at 4 ℃, and the supernatant was transferred to a new 1.5 mL tube.

Step 2: Supernatant from above was then centrifuged at ~30,000 × g for another 15 min at 4 ℃, then the supernatant was discarded, ~500 μL solution A was added to wash pellet (enriched membrane).

Step3: Centrifuged at ~30,000 × g for another 10 min at 4 ℃, and pipette out supernatant carefully to remove lipid. Repeated washing step one more time with 500 μL solution A.

Step4: After the same centrifugation condition, the supernatant was discarded and the pellet was dissolved with a certain amount of solution A (e.g. 4 μL solution A dissolved 1 oocyte membrane for further *in vitro* reaction processes).

Step5: Above mixture was centrifuged at 500 × g for another 15 min at 4 ℃. The pure membrane fraction was taken from supernatant for further use.

2) **Oocytes cytosol extraction processes**

To test the activity of enzymes expressing in the cytosol of oocytes, dialysis was applied to accumulate concentration of soluble proteins and remove background substrates in the following processes.

Step 1: Oocytes were homogenized in a certain volume of solution A (e.g. 500 μL solution A was used to homogenize 5 oocytes, 100 μL/oocyte).

Step 2: After homogenization, samples were centrifuged at ~30,000 × g for 20 min at 4 ℃, gently took lipid out from upper layer of the supernatant, and then certain volume of transparent supernatant was transferred into a filter (Amicon Ultra-0.5 Centrifugal Filter Devices) for dialysis (10 kDa, 30 kDa or 50 kDa) based on molecular weight of targeted protein.

Step 3: Centrifuged at ~10,000 × g for ~12 min at 4 ℃ until less than 100 μL supernatant left in the filter device, discarded solution in the collected microcentrifuge tube, added another 400 μL solution A to dilute and centrifuge again, and then repeated dilution and centrifugation processes for 4-5 times.

Step 4: After centrifugation, put the filter device upside down, centrifuged at ~1,000 × g for ~2 min at 4 ℃. Then the background cGMP or cAMP from cytosol was removed clean enough, not influencing following enzyme activity tests.

2.3.4 Protein quantification with YFP fusion

Before enzyme reaction processes, fluorescence emission values were measured from membrane samples expressing YFP fusion constructs by Fluoroskan Ascent microplate fluorometer (Thermo Electron Corporation). These constructs contain all the YFP-*Cr*2c-Cyclop and its YFP fusion mutants, YFP-*Sr*RhoPDE and its mutants with YFP tag. The standard YFP sample (purified recombinant YFP from *E. coli*, 1 μg/μL) can be applied to generate a linear relationship with YFP amount. This allows us to measure YFP-tagged protein amount from a calibration curve. The protein amount was then calculated by using the YFP fluorescence with $y=(x+0.015)/(27.7*0.02)$ pmol, x represents fluorescence emission value, generating from the same buffer. To calculate the turnover number of YFP-*Cr*2c-Cyclop and YFP-*Sr*RhoPDE in

dark or light illumination, the cyclase activity or PDE activity were correspondingly measured in different experiments.

2.3.5 *In vitro* reaction of membrane or cytosol extracted from oocytes

After oocytes membrane extraction, the pellet was resuspended with certain volume (e.g. 1 µL, 4 µL, 8 µL) of solution A per oocyte for enzymatic reactions. The reaction system comprises a ratio of 1:9 (membrane to reaction buffer), the final concentrations of the reaction buffer as followings:

To test 2c-Cyclop enzyme activity, the final reaction buffer contains 100 mM NaCl, 75 mM Tris-HCl, 5 mM DTT, 5 mM $MgCl_2$ or as indicated, with additional 0.25 mM ATP, 0.2 mM GTP, pH 7.3. For RhoPDE enzyme reaction processes, the substrate in the reaction buffer was used 100 µM cGMP-Na or cAMP-Na or as indicated concentration instead of ATP and GTP.

In individual experiments, the mixture of membrane and reaction buffer (volume ratio1:9) was in the reaction at 20 °C in a water bath. Moreover, certain volume (e.g. 40 µL or 60 µL) of cytosol from 1 oocyte was directly added 2 µL stock or diluted substrate to a final concentration. The reaction time periods were strictly recorded under a dark condition or light illumination. Finally, 10 µL reactive mixture was then transferred to 190 µL sample dilute (0.1 N HCl) to stop the reaction. The samples can be stored at -20 °C for further cGMP or cAMP immunoassay.

2.3.6 Conditions for sensitivity and action spectra tests of *Cr*2c-Cyclop and *Vc*2c-Cyclop

To obtain light sensitivities of 2c-Cyclops, three different wavelengths were applied for Cr2c-Cyclop in several light intensity points. Different light powers were detected by LaserCheck power meter (Coherent Technologies) from the bottom of a black tube (cut out a circular hole), and the bottom area was about 12.56 mm^2, simulating the area which reaction volume can obtain light illumination. Then the light intensities were adjusted as follows: Blue (473 nm) and green light (532 nm) intensities were adjusted to 0.02, 0.09, 0.26, 0.8, 2.39, 7.17, 21.5, 64.5 $µW/mm^2$. Red light (596 nm) intensities were adjusted to 0.03, 0.1, 0.33, 0.92, 2.79, 5.41, 18.71 $µW/mm^2$. Similarly, for *Vc*2c-Cyclop sensitivity test, the 556 nm green light was employed in

several light intensities (1.27, 3.62, 10.51, 30.56, 85.19 μW/mm^2).

On the other hand, to observe action spectra of the two 2c-Cyclops, certain light intensities closed to half-saturated inhibition were employed to do *in vitro* reaction under different wavelengths of light (from UV-light to red light). A white light source Light PhotoFluor II (89North) was applied to obtain different wavelengths by correspondingly narrow bandwidth interference filters (Edmund Optics). For *Cr*2c-Cyclop, the wavelengths were applied at 365, 422, 460, 497, 517, 541, 563, 568, 580, 600, 641, 658, 714 nm separately, and the corresponding light intensities were adjusted to 1.79, 0.88, 1.31, 1.07, 0.91, 0.9, 0.88, 1.39, 1.29, 1.18, 1.13, 1, 1.04 μW/mm^2. The action spectrum was normalized based on inhibition percentage at overlapped 600 nm with the same light intensity in two separated experiments. For *Vc*2c-Cyclop, the wavelengths were used at 473, 532, 556, 596, 635 nm separately, and the corresponding light intensities were adjusted to 6.37, 6.77, 5.25, 11.15, 4.62 μW/mm^2. LEDs were also used with different wavelengths which were further confirmed by a spectrometer (Ocean Optics).

The light inhibition percentage was used to represent the wavelengths effects. For action spectrum test, final activities of the two enzyme opsins were normalized to the similar number of photons from different wavelength powers (total energy, E_{total}= W*s (J); single photon energy, E=hc/λ, h is Planck's constant, c is the speed of light, λ is the wavelength of the photon; photon number= E_{total} /E). Based on different values of cGMP production in dark and light illumination, the light inhibition percentages were calculated by (D-L)/D, quantifying inhibition effects of individual light wavelengths.

2.3.7 cGMP and cAMP Measurement

The concentrations of cGMP or cAMP were detected by DetectX Direct Cyclic GMP or cAMP Enzyme Immunoassay Kit (Cat. No. K019-H5, Cat. No. K020-H5, Arbor assays). The immunoassay kit was based on Enzyme-linked immunosorbent assay (ELISA), free cGMP or cAMP from samples will compete for binding sites with cyclic GMP-peroxidase or cyclic AMP-peroxidase conjugate in monoclonal antibody specific for cyclic GMP or cyclic AMP. After 2 h incubation, the substrate was added to react with bound cyclic GMP or AMP-peroxidase conjugate, stayed in 30 min and stopped the reaction, yellow color solvent in

microplate was detected absorbance at 450 nm wavelength by Thermo Fisher Scientific Multiskan Ascent 96/384 Plate Reader.

Samples from reaction processes were diluted into a reasonable range, 0.5~32 nM for cGMP, 0.56~150 nM for cAMP, as the standard curve from Chemiluminescent Immunoassay Kit. After *in vitro* reaction, samples for measurement were prepared from stop solution with the reaction mix. While *in vivo* assay, cGMP or cAMP levels from oocytes expressing protein can also be directly measured from the transparent supernatant after oocytes homogenization and centrifugation.

To make measurements more convincing, the measured values for calculation normally localized in the middle range of standard curves instead of the border range. Standard curves were normally prepared in each experiment or 96 well microplate due to variable changes.

2.3.8 Western blot of *Cr*2c-Cyclop-YFP expressing in oocytes

To visualize *Cr*2c-Cyclop-YFP expressed in oocytes, anti-GFP antibody (GFP Antibody (FL): sc-8334, a rabbit polyclonal IgG; 200 µg/ml) was applied to detect *Cr*2c-Cyclop-YFP membrane protein by western blot. The procedures were operated as followings:

1) Sample preparation

A certain number of control oocytes and ChR2-YFP expressing oocytes membrane was extracted as a negative and positive control respectively. Membrane extraction procedures were following treatment in **2.3.3** section.

For sonication treatment, in the first step, oocytes were homogenized and put in ~200 µL extraction solution A to sonicate in water bath sonicator (Shyam Industries) for 30 s, then put on ice for 30 s, repeated sonicating and cooling cycles for 3 times. In the final step, the membrane pellet was resuspended and added to final 30 µL (containing 3% SDS, 100 mM DTT), stayed at room temperature for ~2 h. For heating treatment, extracted membrane samples and 40 ng YFP standard protein were added an aliquot of 5× SDS loading buffer and boiled at 95 °C for ~5 min before loading samples in an SDS-PAGE gel.

2) SDS-PAGE gel electrophoresis

10% precise Tris-Glycine gel (Thermo Scientific™) was used as a protein running system with $1\times$ running buffer in the gel running tank (Bolt Mini Gel Tank). PageRuler Plus Prestained Protein Ladder was used as a protein marker. Then samples above and marker were loaded in the SDS gel lanes separately. The gel was running at 60 V for ~30 min, then increasing to 100-150 V for about 2 h, until the loading dye was close to the bottom of the gel. Different percentages of SDS-PAGE gels were also prepared for further protein visualization and purification. Tris-SDS buffer systems were prepared for SDS-PAGE gel (**Table 2.4 and 2.5**).

Table 2.4   4x Tris-SDS pH 8.8

Tris base	1.5 M
SDS	0.4 %
Adjust pH 8.8 with HCl	

Table 2.5   4x Tris-SDS pH 6.8

Tris base	1.5 M
SDS	0.4 %
Adjust pH 6.8 with HCl	

10% APS was prepared with 1 g Ammonium persulfate in 10 mL ddH_2O. An aliquot was stored at -20 ℃ for long-term and at 4 ℃ for the usual application. The 10% resolving mini gel was prepared in **Table 2.6**.

Table 2.6 10% resolving gel preparation (8 mL)

H_2O	4 mL
4x Tris-SDS pH 8.8	2 mL
40 % Acrylamide	2 mL
10 % APS	40 μL
TEMED	20 μL

To prepare for stop gel, 0.5 mL of resolving gel solution without APS and TEMED can be added

2.5 µL APS and 1.0 µL TEMED firstly.

4% stacking mini gel and 10× running buffer preparations followed in **Table 2.7 and 2.8**.

Table 2.7   4% stacking gel preparation (5 mL)

H_2O	3.3 mL
4x Tris-SDS pH 6.8	1.25 mL
40% Acrylamide	0.5 mL
10% APS	25 µL
TEMED	10 µL

Table 2.8   10× running buffer

Tris base	250 mM
Glycine	1.9 M
SDS	1%
Adjust pH to 8.3	

3) Transfer proteins from gel to a nitrocellulose membrane

The gel was soaked in 1× transfer buffer (**Table 2.9**) for 5-10 min and then transferred on top of nitrocellulose transfer membrane (pre-wet in 1× transfer buffer, Whatman™ Protran™ Nitrocellulose Blotting Membranes). Three pieces of 1× transfer buffer soaked filter paper were put on both sides, and bubbles inside were removed with a clean tube. This sandwich was placed in a transfer system (Western Blot Transfer System, PEQLAB Biotechnology GmbH) with gel in anode and membrane in cathode side, adjusted to 0.8 mA/cm^2, and ran for ~ 1 h.

Table 2.9   1× transfer buffer (semi-dry)

Tris base	48 mM
Glycine	39 mM
Methanol	20%
SDS	0.04%
Adjust pH to 8.3	

4) Antibody incubation

After membrane transfer, the membrane was blocked with 3%-5% milk or 1% BSA (bovine serum albumin) at room temperature for 1 h or at 4 ℃ for overnight. GFP-HRP antibody was diluted to 1:1000 in 1× TBST buffer (**Table 2.10**) and incubated in time conditions as above. Washed the membrane with TBST 4-5 times, 5-10 min for each time.

Table 2.10   1× Tris-buffered saline with Tween 20 (TBST) buffer

Tris base	20 mM
NaCl	150 mM
Tween20	0.1%
BSA or milk (for blocking)	1% or 3%
Adjust pH to 7.5	

5) Imaging and data analysis

The washed membrane was then put in a transparent plastic bag, then added 1:1 ratio of two ECL solutions (Pierce™ ECL Western Blotting Substrate) to the membrane and sealed plastic bag for 5 min incubation, then detected Chemiluminescent at different exposure time by a Bio-RAD imaging system (Bio-Rad ChemiDoc™ XRS+ with Image Lab™ Software).

2.4 Recombinant protein overexpression and purification in *E. coli* BL21 (DE3)

The DNA sequence of histidine kinase domain of *Cr*2c-Cyclop (briefly named *Cr*2c-HK) was selected to clone in PET-28b vector for overexpression and inclusion body purification. The restriction sites in 5' heads of both primers were used NcoI and XhoI. The stop codon TAA was added before his-tag in the 3' cloning site of the vector. Overexpression processes as follows:

Step1: Pre-culture: One positive colony of *E. coli* BL21(DE3) with PET-28b-*Cr2c-HK* recombinant DNA was transferred into 5 mL LB (+kanamycin), shaking at 200 rpm for overnight at 37 ℃.

Step2: Induction: 1 mL pre-culture was transferred into ~100 mL LB (+kanamycin), grew until OD600=0.6-0.8 (~2.5 h). Then 4 copies were divided into 15 mL each in 50 mL falcon tubes,

and different IPTG concentrations (0, 0.5 mM, 1 mM, 2 mM) were used to induce protein expression.

Step3: Harvest and lysate: After adding IPTG, shaking at 200 rpm, 3 h at 37 °C, centrifuged at 4,000 × g for 10 min at 4 °C, suspended pellet with 700 μL lysis buffer (**Table 2.11**) in 1.5 mL tubes with protease inhibitor cocktail (dilution 1:1000).

Step 4: Sonicate: *E.coli* cells were lysed on ice for ~30 min, then frozen in liquid N_2 and thawed on a water bath sonicator at 40 °C, repeated 4 times. 30 μL sample was taken out as total cell lysate, centrifuged for 30 min, 10,000 × g at 4°C to get inclusion bodies. 30 μL supernatant was taken out as soluble fractions.

Step5: Wash the pellet: The pellet was washed 4 times with 700 μL buffer each time, and centrifuged at 10,000 × g for 10 min at 4 °C between washing steps. The detergent buffer was firstly used to wash pellet (**Table 2.12**), the pellet was resuspended in washing buffer and centrifuged twice(**Table 2.13**), and then the pellet was washed with Tris buffer (**Table 2.14**).

The pellet was finally resuspended in 80 μL 20 mM Tris Buffer at pH 8.0 and taken 30 μL out as insoluble fractions.

Table 2.11   Lysis buffer

Tris-HCl	50 mM
NaCl	100 mM
EDTA	5 mM
Triton X-100	0.5 %
DTT	5 mM
Lysozyme	2 mg/ml
Adjust pH to 8.0	

Table 2.12   Detergent buffer

Tris-HCl	20 mM
NaCl	200 mM
Deoxycholic acid	1 % (w/v)
Nonidet P-40	1 % (v/v)
DTT	5 mM
EDTA	2 mM
Adjust pH to 7.5	

Table 2.13 Washing buffer

Tris-HCl	20 mM
Triton X-100	0.5 %
DTT	5 mM
EDTA	1 mM
Adjust pH to 7.5	

Table 2.14 Tris buffer

Tris-HCl	20 mM
DTT	5 mM
EDTA	1 mM
Adjust pH to 8.0	

Based on the above small volume of sample preparation, a larger volume of sample was treated with an enlarged buffer system. The pellet from 0.5 L recombinant *E. coli* was suspended and sonicated in 200 mL lysis buffer. Strong sonicator (Q500 Sonicator System) was applied to replace the "freeze and thaw" mild method. The suspended *E. coli* cells were sonicated on ice for ~10 min, with 50% intensity and 50% duty cycle. In detergent and washing steps, 30 mL buffer was used to resuspend pellet each time. The final pellet was resuspended by 2.5 mL Tris buffer or PBS buffer (150 mM NaCl, 10 mM NaH_2PO_4, pH 7.4 with NaOH). Then all samples

were mixed with 5× SDS loading dye (**Table 2.15**) and loaded in an SDS-PAGE gel (prepared with 4% stacking gel and 12% resolving gel). After running the gel, it was put in Coomassie stain solution (**Table 2.16**), shaking for ~1-2 h at room temperature. Then Coomassie background was removed by destain solution (**Table 2.17**) until the protein bands were clearly observed.

Table 2.15   5× SDS loading dye

Tris-HCl	250 mM
SDS	10 %
Glycerol	30 %
β-Mercaptoethanol	5 %
Bromophenol blue	0.02%
Adjust pH to 6.8	

Table 2.16   Coomassie stain solution

H_2O	250 mL
Coomassie-R250	0.1 g
Acetic acid	50 mL
Methanol	200 mL

Table 2.17   Coomassie destain solution

H_2O	350 mL
Acetic acid	50 mL
Methanol	100 mL

2.5 BiFC (Bimolecular fluorescence complementation) experiment and imaging

BiFC analysis can be widely applied to explain whether two different proteins would interact with each other in direct visualization of living cells. This may exclude possibilities that two

interacted proteins could be disrupted by artificial cell lysate or fixation. Two non-fluorescent fragments were selected from fluorescent protein like YFP, which would recover fluorescence after co-expression in cells.

According to reference[102], YFP truncation at the 155 position can generate one YFP-N-terminal part 1-154 and YFP-C-terminal part 155-238. Both non-fluorescent fragments would fuse together and show strong fluorescence when fusion with two interaction partners (**Figure 2.2**), but weak or no fluorescence can be detected if two proteins would not interact. On the other hand, this method can be used in the identification of transmembrane helices numbers of opsins. Here we chose to use BiFC analysis to confirm whether there is an even or odd number of transmembrane helices in rhodopsin domain of *Cr*2c-Cyclop, *Vc*2c-Cyclop, COP5 (HKR) as well as *Sr*RhoPDE. In each construct, the YFP-N (1-154) part was fused to C-terminal of rhodopsin domain, while the YFP-C (155-238) part was fused to N-terminal of rhodopsin domain.

The fluorescence of YFP-fused proteins was detected by a confocal microscope (Leica DM6000) after 3 days RNA injection in oocytes. The excitation laser wavelength was adjusted to 515 nm, while 535 nm light emission intensity was detected.
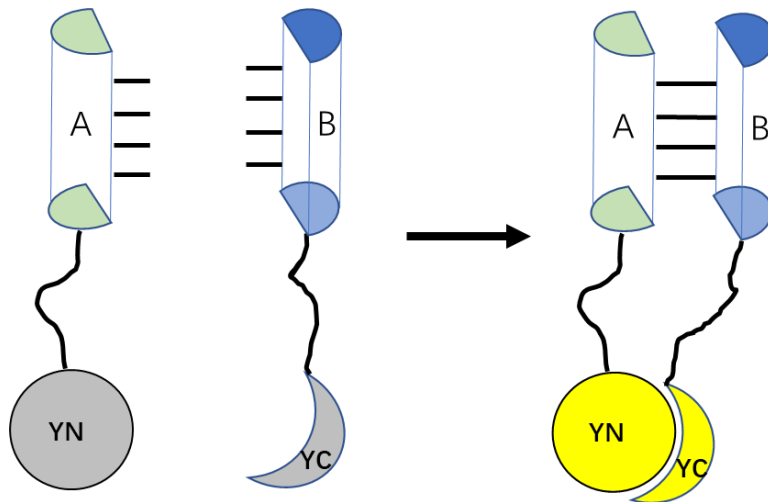

Figure 2.2 Working model of BiFC experiment.

Two putative interaction proteins A and B fused with non-fluorescent YFP-N and YFP-C part separately. After co-expression in cell systems, the interaction between A and B would recover bright fluorescence, detected by a confocal microscope.

2.6 Bioinformatics

Vector NTI or BioEdit and Benchling were applied in all experiments to manipulate DNA, protein sequences, primer design and DNA sequencing check.

In alignment analysis, homologous proteins from different organisms were selected by using NCBI or JGI database Blast. Some known protein structures and motif analysis were based on PDB protein database. To obtain alignment results, Clustal Omega was used to export clustal files which were then edited in GeneDoc for sequences alignment export. Transmembrane helices numbers in rhodopsin domains of *Be*CyclOp, *Cr*2c-Cyclop, *Vc*2c-Cyclop, *Sr*RhoPDE were predicted in CCTOP, a topology prediction web server[103]. Paircoil2[104], as a Paircoil program available online, can be applied to predict coiled-coil motifs or sequences in linker regions of photoreceptors.

3. Results

In my study, I mainly characterized two classes of enzyme rhodopsins. One study involved *Cr*2c-Cyclop and *Vc*2c-Cyclop, demonstrating that they exhibit light-inhibited and ATP-dependent guanylyl cyclase activity. Another investigation was rhodopsin phosphodiesterase (RhoPDE), showing light-activated cGMP and cAMP hydrolysis activity.

3.1 Classification of three enzyme opsins

Enzyme rhodopsins, as one of rhodopsin members, mainly come from three microorganisms including fungi, green algae as well as protist nowadays. With two or more integrated domains, most of them are closely related with cGMP or cAMP regulation by light.

They can be divided into three classes (**Figure 3.1**). In my study, *Cr*2c-Cyclop and *Vc*2cCyclop from unicellular *Chlamydomonas reinhartii* and multicellular *Volvox carteri* were characterized as light-inhibited guanylyl cyclases. Two components (histidine kinase and response regulator) locate in the middle. This may suggest that opsins with four domains from green algae are proposed to be light-inhibited enzymes. The second class belongs to light-activated guanylyl cyclase opsins (Cyclop) from fungi. In 2014, BeGC1 or *Be*CyclOp was firstly identified from *Blastocladiella emersonii* and then characterized as the light-activated guanylyl cyclase by Gao et al. in 2015[7, 54]. Meanwhile, other members from different fungi species were identified as light-activated enzymes, such as *Am*CyclOp and *Ca*CyclOp. The third class was confirmed by us as light-activated phosphodiesterase (PDE) from the protist *Salpingoeca rosetta*.
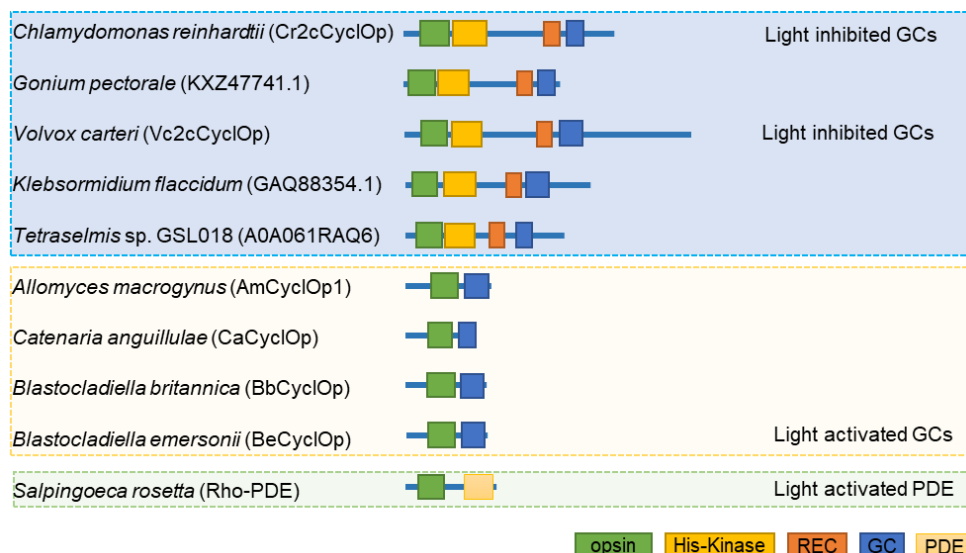
Figure 3.1 Three classes of enzyme rhodopsins.

The functional domains are labeled with five different colored boxes. The orange box (REC) indicates a receiver domain or response regulator (RR) domain. The blue box (GC) indicates the guanylyl cyclase domain, specifically catalyzing GTP to cGMP in these enzyme opsins. Three classes of enzyme opsins were divided into light blue, yellow and green color background respectively.

3.2 Characterization of 2c-Cyclops from *C. reinhardtii* and *V. carteri*

The cDNA sequence was cloned from *C. reinhardtii* by us and then translated to a large protein (1605 aa). A conserved domain search from NCBI[105] illustrates bacteriorhodopsin-like domain with aa95-aa308, histidine kinase domain with aa342-aa566, response regulator domain with aa1043-aa1159, and guanylyl cyclase domain with aa1211-aa1389. From schematic and alignment between full lengths of *Cr*2c-Cyclop and *Vc*2c-Cyclop (**Figure 3.2 and Figure S1**), it is suggested that the four domains are very conserved (Identities=84%) in the two enzyme opsins.
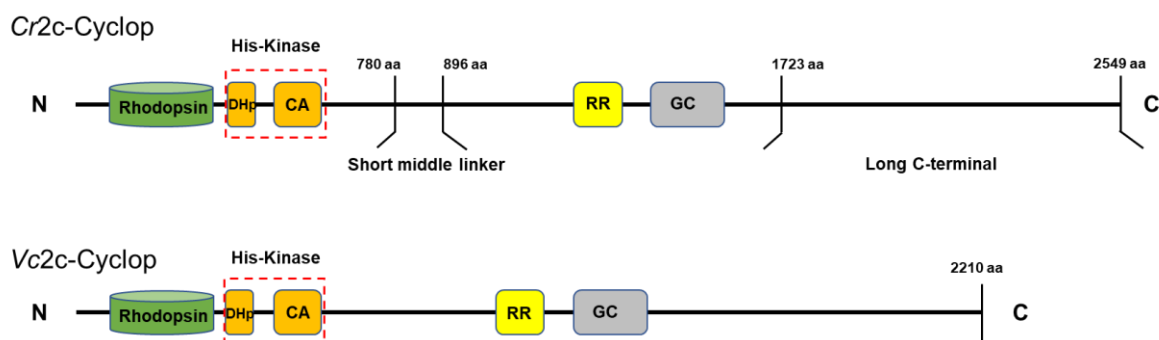
Figure 3.2 Schematic models of *Cr*2c-Cyclop and *Vc*2c-Cyclop.

Conserved domains analysis of full lengths of *Cr*2c-Cyclop (Cre11.g467678) and *Vc*2c-Cyclop (Vocar.0009s0380.1) protein sequences taken from JGI genome database. Four domains are labeled in different colored boxes, consisting of Rhodopsin domain in green box, histidine kinase including DHp domain (**D**imerization and **H**istidine-containing **p**hosphotransfer) and CA domain (**C**atalytic and **A**TP-binding) in two orange boxes respectively, RR domain (**R**esponse **R**egulator) in yellow box and GC domain (**G**uanylyl **C**yclase) in gray box.

The two-component cyclase opsins *Cr*2c-Cyclop and *Vc*2c-Cyclop were characterized in detail by different methods as followings.

3.2.1 Optimization of functional *Cr*2c-Cyclop sequences

The full-length of *Cr*2c-Cyclop sequence information comes from *C. reinhardtii* genome database, Cre11.g467678. Guanylyl cyclase activities of four different truncated sequences were compared under dark and light illumination. From the test, it is clear to see that *Cr*2c-Cyclop.smc has the best dark/light ratio and the highest dark activity, which was easily quantified with the highest expression and further used for following characterization (**Figure 3.3 A, B**). In the full text, I used the brief name *Cr*2c-Cyclop to replace *Cr*2c-Cyclop.smc applied in the following experiments.
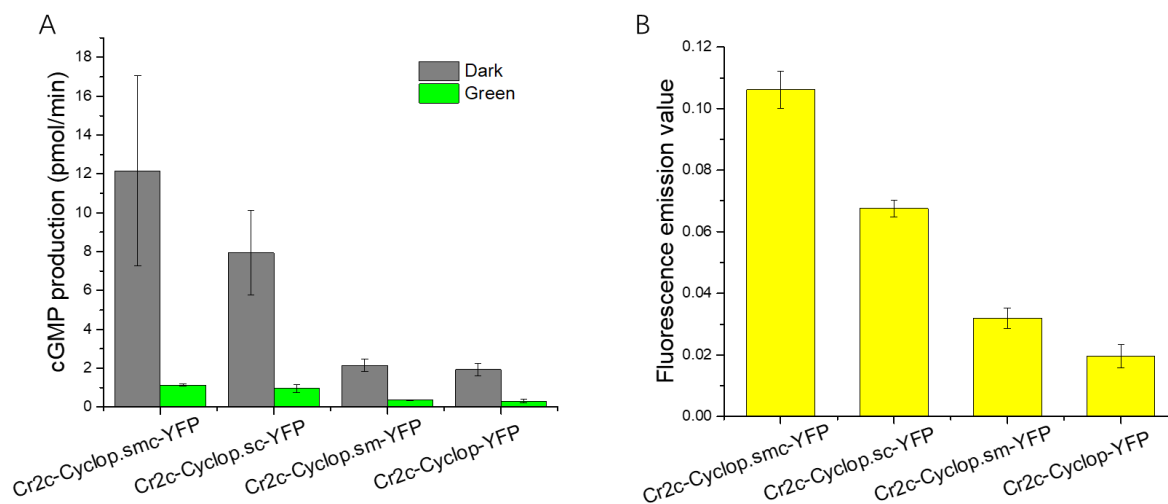
Figure 3.3 Comparison of four lengths of *Cr*2c-Cyclop protein.

A, The dark and light activity measurements from four different lengths of *Cr*2c-Cyclop, activities in dark and light came from one oocyte membrane expressing constructs individually, 30 ng cRNA injection for each construct, 3dpi. n=4, error bars=SD.

B, The fluorescence emission values were measured in four constructs and shown here after subtraction of control emission values, 12.5 oocytes membrane were extracted for each measurement, n=3, error bars=SD.

*Cr*2c-Cyclop full-length comes from database Cre11.g467678.

*Cr*2c-Cyclop.sm represents the full-length *Cr*2c-Cyclop protein without short middle linker (780 aa-896 aa).

*Cr*2c-Cyclop.sc indicates C-terminus polypeptide (1723 aa-2549 aa) is deleted from the full-length protein.

*Cr*2c-Cyclop.smc indicates middle linker (780 aa-896 aa) is further deleted from *Cr*2c-Cyclop.sc sequence. The above four sequences are depicted in **Figure 3.2**.

3.2.2 Optimized reaction system for *Cr*2c-Cyclop

To uncover the mechanisms of *Cr*2c-Cyclop more efficiently, I first tried to optimize the reaction systems to get a higher dark to light ratio. Reaction buffer included 75 mM Tris-HCl, 10 mM $MgCl_2$, 5 mM DTT, different concentrations of GTP or ATP, pH adjusted to ~7.35. Moreover, light sensitivity and action spectrum can be more convincing when using the best reaction system with a stronger light regulation.

When increasing different substrate GTP concentrations (from 0.02 mM to 1.62 mM) with 1 mM ATP, the dark activity strongly increased while light activity increased slightly. The D/L ratio peaked at ~13 fold with 0.18 mM GTP. Interestingly, the dark activity points fitted Michaelis-Menten equation well, and the K_m value for GTP in dark is ~0.13 mM (**Figure 3.4 A**). Here I further applied 0.18 mM GTP in reaction system to quantify cGMP production in

different ATP concentrations (from 0.01 mM to 0.81 mM). Last three concentration points showed similar D/L ratio, keeping at ~7 fold (**Figure 3.4 B**). I chose to use 0.2 mM GTP and 0.25 mM ATP for next optimized processes.
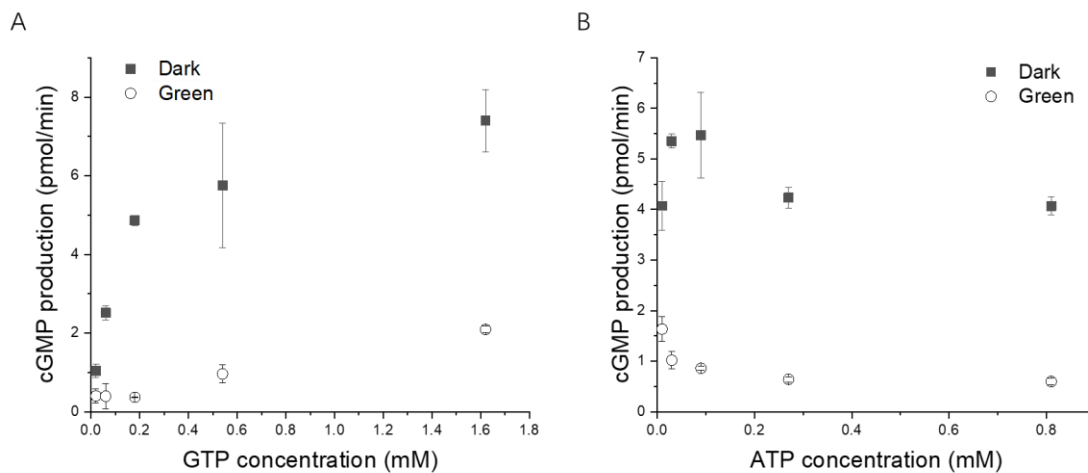

Figure 3.4 cGMP production in different GTP and ATP concentrations.

A, cGMP production in dark and light with different substrate GTP concentrations (starting at 0.02 mM, 0.06 mM, 0.18 mM, 0.54 mM, 1.62 mM individually).

B, cGMP production in dark and light with different ATP concentrations (0.01 mM, 0.03 mM, 0.09 mM, 0.27 mM, 0.81 mM respectively).

For both graphs, activities came from one oocyte membrane expressing Cr2c-Cyclop, 30 ng cRNA injection, 3dpi. n=3, error bars=SD.

Different cation ions like Mn^{2+} or Ca^{2+} should also have effects on Cr2c-Cyclop activities, so cyclase activities were compared at different concentrations of metal ions for further optimization. Reaction system was regulated as follows: 75 mM Tris-HCl, 5 mM DTT, 0.2 mM GTP and 0.25 mM ATP, pH adjusted to ~7.35, and certain concentration of $MgCl_2$, $MnCl_2$ or $CaCl_2$ was added in individual tests. As presented in **Figure 3.5 A**, additional 2 mM Mn^{2+} can strongly increase cyclase activity in dark and light, but the D/L ratio decreased to ~4 fold due to strong light activity. Different Mg^{2+} concentrations illustrated no obvious difference, and D/L ratio at 5 mM Mg^{2+} was ~11 fold, better than that of at 10 mM or 20 mM Mg^{2+} with ~9 fold. However, additional 2 mM Ca^{2+} abolished light regulation by dramatically decreasing dark activity. This suggested that Ca^{2+} could compete with Mg^{2+} binding sites in the cyclase domain of Cr2c-Cyclop. In the next experiments, 5 mM Mg^{2+} was applied in the reaction buffer system.
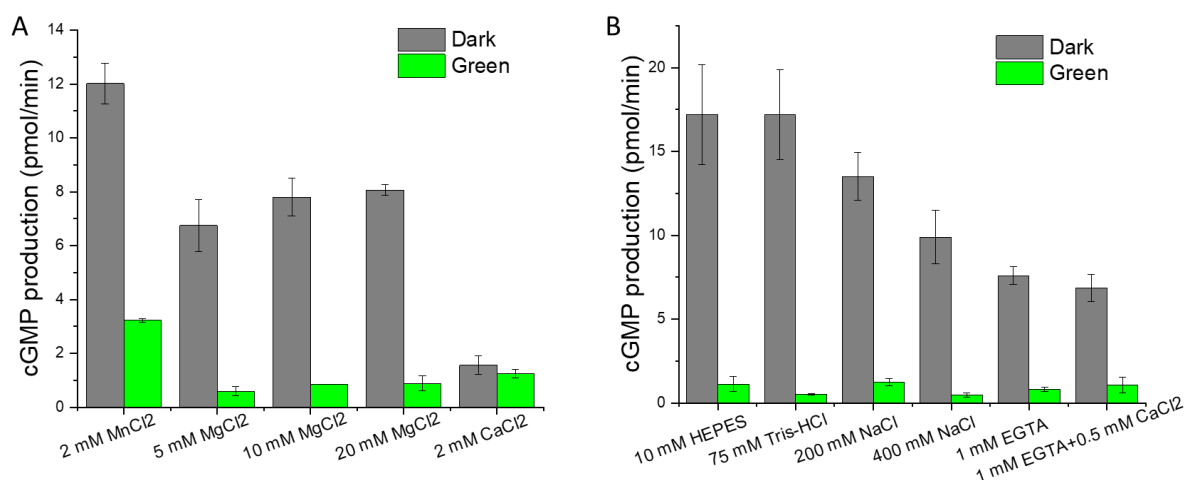
Figure 3.5 Comparison of mental ions effects on *Cr*2c-Cyclop activities.

A, 2 mM MnCl$_2$ also included 5 mM MgCl$_2$ in reaction buffer, while 2 mM CaCl$_2$ included 10 mM MgCl$_2$ in reaction buffer.

B, 10 mM HEPES or 75 mM Tris-HCl buffer system were compared containing 100 mM NaCl, while 1 mM EGTA or 1 mM EGTA + 0.5 mM CaCl$_2$ were compared containing 100 mM NaCl.

For both graphs, GC activities were calculated from one oocyte membrane expressing *Cr*2c-Cyclop, 30 ng cRNA injection, 3dpi. n=3, error bars=SD.

Interestingly, when adding a certain concentration of NaCl, it was helpful to increase cyclase activity as well as light regulation (**Figure 3.5 B**). At 100 mM NaCl, 75 mM Tris-HCl contained reaction buffer showed the highest light-inhibited effect (D/L ratio: ~34) compared with HEPES system (D/L ratio: ~15). Moreover, compared with the previous buffer system with less than 10 mM NaCl, the *Cr*2c-Cyclop dark activity stayed around 6 pmol/min cGMP production from one oocyte membrane. Here the cyclase activity in dark stage increased to ~17 pmol/min cGMP production from one oocyte membrane. Therefore, 100 mM NaCl in the reaction buffer system could accumulate more cGMP production excluding the influence of expression. When increasing NaCl concentration to 200 mM or 400 mM, the dark activity decreased as well as light regulation effect. Compared cyclase activities between 1 mM EGTA and 1 mM EGTA + 0.5 mM CaCl$_2$, they were observed with similar lower dark activity and light regulation. Therefore, 100 mM NaCl was further applied in the reaction system. The final optimized reaction system can be seen in **2.3.5** section.

3.2.3 Eight transmembrane helices (8-TM) in opsin part of 2c-Cyclops

Constrained Consensus TOPology prediction server (CCTOP) was applied to analyze the opsin domains of *Cr*2c-Cyclop and *Vc*2c-Cyclop, predicting a similar 8-TM topology. Meanwhile, based on alignment with opsin domains of two other experimentally confirmed 8-TM *Be*Cyclop and *Sr*RhoPDE, as well as classical 7-TM microbial opsins Channelrhodopsin-2 (ChR2), Bacteriorhodopsin (BR) and Halorhodopsin (HR), it was suggested that 2c-Cyclops could have 8 transmembrane helices with an additional transmembrane helix TM0 in N-terminal. Their topology could be more similar with *Be*Cyclop and *Sr*RhoPDE (**Figure 3.6**), with both N- and C- termini locating in the same cytosolic side.

Figure 3.6 Alignment of rhodopsin domains from *Cr*2c-Cyclop, *Vc*2c-Cyclop, *Be*CyclOp, and *Sr*RhoPDE with other single-domain microbial opsins, including ChR2, BR, HR.

Transmembrane helices are labeled with green or yellow color background.

For ChR2, BR and HR, transmembrane helices are labeled based on their structures (ChR2 PDB ID: 6EID, BR PDB ID: 5AZD, HR PDB ID: 3A7K). Gene bank accession No.: *Be*CyclOp: AIC07007.1, *Sr*RhoPDE: XP_004998010.1, Channelrhodopsin-2 (ChR2): ABO64386, Bacteriorhodopsin (BR): WP_016329665, Halorhodopsin (HR): AAA72222.1. Alignment by Clustal Omega.

Membrane-spanning regions were predicted by CCTOP (Constrained Consensus TOPology prediction server) and APSSP (Advanced Protein Secondary Structure Prediction Server).

To confirm both N- and C- termini of the opsin part of 2c-Cyclop located on the same side, I generated BiFC (Bimolecular fluorescence complementation) constructs. The working model is depicted in **Figure 3.7 A**. Two parts of YFP can fuse together and show strong fluorescence after expressing those constructs in *Xenopus* oocytes (**Figure 3.7 B**).

Considering opsin part alignment, CCTOP prediction, BiFC experimental results, and cytosolic cGMP production, it was then suggested that both *Cr*2c-Cyclop and *Vc*2c-Cyclop have 8-TM in the rhodopsin domains. Meanwhile, the opsin part of Cop5 (or "HisKR") was also proved as an 8-TM opsin.
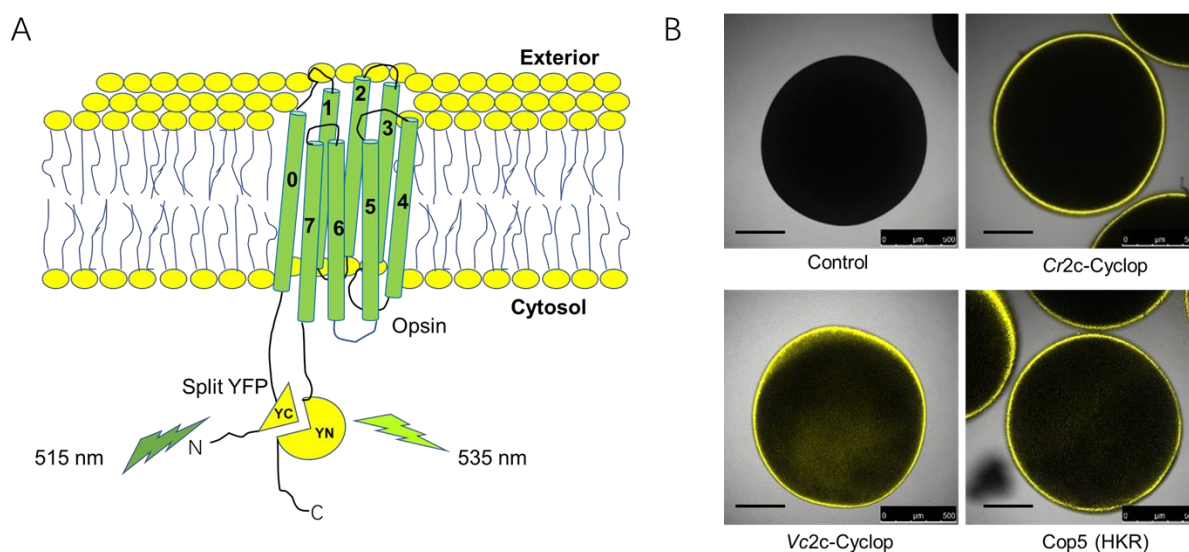
Figure 3.7 2c-Cyclops share with a similar topology of 8-TM.

A, Schematic working model of BiFC experiments. The opsin domains of 2c-Cyclops were individually fused with split YFP (YFP-C-terminal 155-238 aa, YFP-N-terminal 1-154 aa) in both termini.

B, Fluorescence graphs are illustrated as control oocytes, oocytes expressing BiFC-*Cr*2c-Cyclop-Rhodopsin, BiFC-*Vc*2c-Cyclop- Rhodopsin, BiFC-Cop5-Rhodopsin constructs respectively. Fluorescence images were taken in a confocal microscope after 30 ng cRNA injection for each construct, 3dpi.

Black scale bars: 300 μm.

3.2.4 2c-Cyclops are light-inhibited guanylyl cyclases (GC)

After doing the alignment of guanylyl cyclase domains in 2c-Cyclops together with CYG12, bPAC, and Cya2, key residues with labeled colors show high identities (**Figure 3.8**). This suggests that 2c-Cyclops function as a GC. Moreover, based on the crystal structure of class III cyclase Cya2 (from cyanobacterium *Synechocystis* PCC6803), guanylyl cyclase normally functions as a dimer, with high affinity for GTP than for ATP[106].


```
Cr2cCyclop : ------HVVILFSDIVGFTSLSSKLPTAEVFLMLSNMFTAFDKLTDRF-SVYKVETIGDAYMVAAGHDE : 1272
Vc2cCyclop : ------HVVILFSDIVGFTSLSSKLPTAEVFLMLSNMFTAFDKLTDRF-SVYKVETIGDAYMVAAGHDE : 1208
BeCyclop   : -EAKEYESVTVFFSDITNFTVISSRTSTKDMMATLNKLWLEYDAIAKRW-GVYKVETIGDAYLGVTGAPD : 511
CYG12      : -PAQEHPEATVLFSDIVGFTEIASRSSPLEVCSLLDELYQRFDAAIEEYPQLYKVETIGDAYMVVCNVTV : 537
Cya2       : -------ITILTSDLRGFTSTSEGLNPEEVVKVLNIYFGKMADVITHH-GGTIDBFMGDGILVLFGAPT : 502
bPAC       : TVEPQLVEKIIFFSDILAFSTLTEKLPVNEVVILVNRYFSICTRIISAY-GGEVTKFIGDCVMASFTKEQ : 211


Cr2cCyclop : DEDKEAKGSPLMRVLGFARAMLDVVRNITA-----PNGERLRIRIGVHCGPAFAGVIGM-KCPRYCFLGD : 1336
Vc2cCyclop : DEDKARKGSPLTRVLGFAKAMLDVVRNITA-----PNGERMRIRIGVHCGPAFAGVIGM-KCPRYCFLGD : 1272
BeCyclop   : VVPDHA-----ERACNFAVDIIEMIKSFKT-----ITGESINIRIGLNSGPVTAGVLGD-LNPHWCLVGD : 570
CYG12      : PCDDHA-----DVLLEFALRMHEEASRVAS-----SLGEPVRIRVGMHSGPVVAGVVGR-KMPRFCLFGD : 596
Cya2       : SQQDDA-----LRAVACGVEMQLALREVNQQVT-GLGLQPLEMGIGINTGEVVVGNIGSEKRTKYGVVGA : 566
bPAC       : G--DAA-----IRT---SLDIISELKQLRHHVEATNPLHLLYTGIGLSYGHVIEGNMGSSLKMDHTLLGD : 271


Cr2cCyclop : TVNTASRMESTGFPMCIHVSENVFKHHPAA-EAELQEVGERDIKGKGHMRTYVV--------------- : 1389
Vc2cCyclop : TVNTASRMESTGFPMCIHVSEDVYQHHPNM-EGEFVEVGEREVKGKGRMRTYLV--------------- : 1325
BeCyclop   : TVNTASRMESTSKAGHIHISESTYHFIKSK--FVTQPLDVMEVKGKGKMQTYWVLGRK----------- : 626
CYG12      : TVNTASRMESHGEAGQIHISEACYCCLRSKERFEIRERGNITVKGKGTMRTYLLSP------------- : 652
Cya2       : QVNLTYRIESYTTGGQIFISSTT--LEAAGDRVHVNGNRTVQPKGVKDPVVIWDV------------- : 619
bPAC       : AVNVAARLEALTRQLPYALAFTAGVKKCCQAQWTFINLGAHQVKGKQEAIEVYTVNEAQKYYDTLQITQL : 341
```

Figure 3.8 Alignment of 2cCyclops GC domains with other proteins.

An alignment of GC domains of *Cr*2c-Cyclop, *Vc*2cCyclop, *Be*Cyclop, CYG12, Cya2, and bPAC (*Bg*PAC). Blue background residues: metal binding, red ones: base recognition (bold red letter represents base recognition for bPAC), green ones: ribose orienting residue, purple ones: transition state stabilizing residue. Accession No.: BeCyclOp: AIC07007.1, CYG12: EDP07101.1, Cya2: WP_010871597.1, bPAC: ADC33127.1

To do *in vitro* reactions, *Cr*2c-Cyclop and *Vc*2c-Cyclop were separately expressed in *Xenopus* oocytes, and then the extracted membrane from oocytes was mixed with reaction buffer to produce cGMP under dark treatment and green light illumination (532 nm).

It was clearly suggested that *Cr*2c-Cyclop has high GC activity in the dark while green light (532 nm) can strongly inhibit GC activity. In dark, *Cr*2c-Cyclop-expressing membranes from 1

oocyte could produce 11 ± 0.3 pmol cGMP/min in the reaction mix. Under green light (20 μW/mm^2, 532 nm) illumination, GC activity reduced to 0.3 ± 0.09 pmol cGMP/min. The dark to light ratio could be ~35. Then the C-terminal YFP fusion construct decreased the GC activity to ~65% from extracted one oocyte membrane, possibly due to less protein amount. But the D/L ratio was not changed obviously (**Figure 3.9 A**).

After extraction of *Vc*2c-Cyclop-expressing oocytes membrane and reaction processes, *Vc*2c-Cyclop also exhibited light-inhibited GC activity with dark/light ratio ~5 (**Figure 3.9 B**). But the dark activity of *Vc*2c-Cyclop expressing in one oocyte membrane was over 10 times lower than that of in *Cr*2c-Cyclop expressing membrane, primarily because the full-length of *Vc*2c-Cyclop is a larger protein with a lower expression level in each oocyte.

The all-*trans*-retinal (ATR) is necessary for opsins to improve the light-regulated efficiency. Without ATR in the cultural medium ND96, *Cr*2c-Cyclop exhibited higher GC activity under green light illumination than that of with additional 1 μM ATR, even though the dark activity is similar (**Figure 3.9 A**). The dark to light ratio changed from ~30 to ~4.5. This suggested that *Cr*2c-Cyclop without completely bound ATR was not degraded but became less light-sensitive.
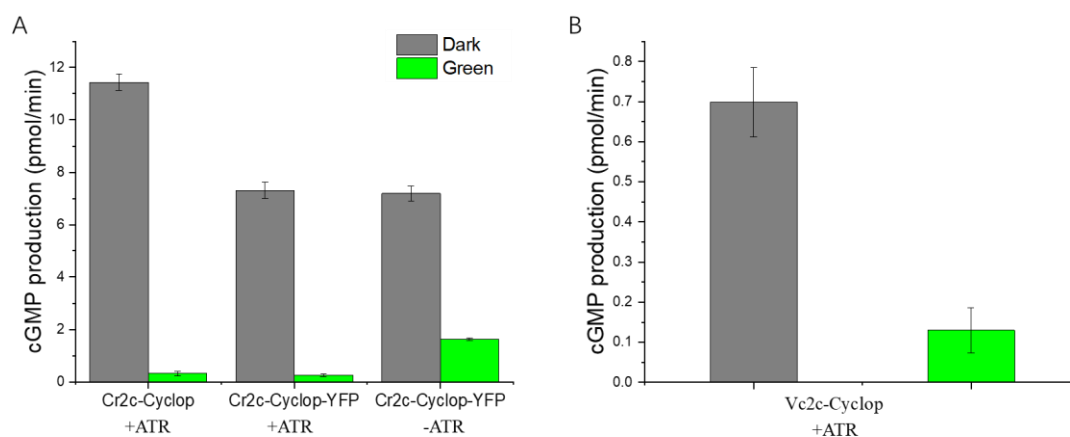
Figure 3.9 Enzymatic activities of *Cr*2c-Cyclop and *Vc*2c-Cyclop under dark and light illumination.

A, Comparison of light and dark activities among *Cr*2c-Cyclop +1 μM ATR, *Cr*2c-Cyclop-YFP +1 μM ATR and *Cr*2c-Cyclop-YFP without ATR. n=3, error bars=SD.

B, *Vc*2c-Cyclop activities under dark and light illumination with 1 μM ATR. n=4, error bars=SD.

For both A and B graphs, in each construct, the same amount of 30 ng cRNA was injected in the same batch of *Xenopus* oocytes. The cGMP production in y-axis came from one oocyte membrane. The intensity of green light (532 nm) was adjusted to ~20 μW/mm^2.

3.2.5 Light sensitivity of *Cr*2c-Cyclop and *Vc*2c-Cyclop

Light sensitivity was compared between *Cr*2c-Cyclop and *Vc*2c-Cyclop. It is clear to see that *Cr*2c-Cyclop shows the most sensitive to light, compared with other light-sensitive nucleotidyl cyclases such as *Vc*2c-Cyclop, *Be*Cyclop, mPAC, and bPAC. The half-maximal inhibition ($K_{0.5}$) was determined to be ~0.2 $\mu W/mm^2$ with both 532 nm and 596 nm light (**Figure 3.10 A**). The $K_{0.5}$ value increased to ~0.5 $\mu W/mm^2$ under blue light (473 nm) illumination, showing less sensitivity. *Vc*2c-Cyclop showed less sensitive to light, and the half-maximal inhibition is ~1.3 $\mu W/mm^2$ at 556 nm light around its action spectrum peak (**Figure 3.10 B**). While the half-maximal activation ($K_{0.5}$) was observed at 55 $\mu W/mm^2$, 4 $\mu W/mm^2$, and 6 $\mu W/mm^2$ for BeCyclop[7], bPAC[53], and mPAC[107] around the peak of their own action spectra respectively.
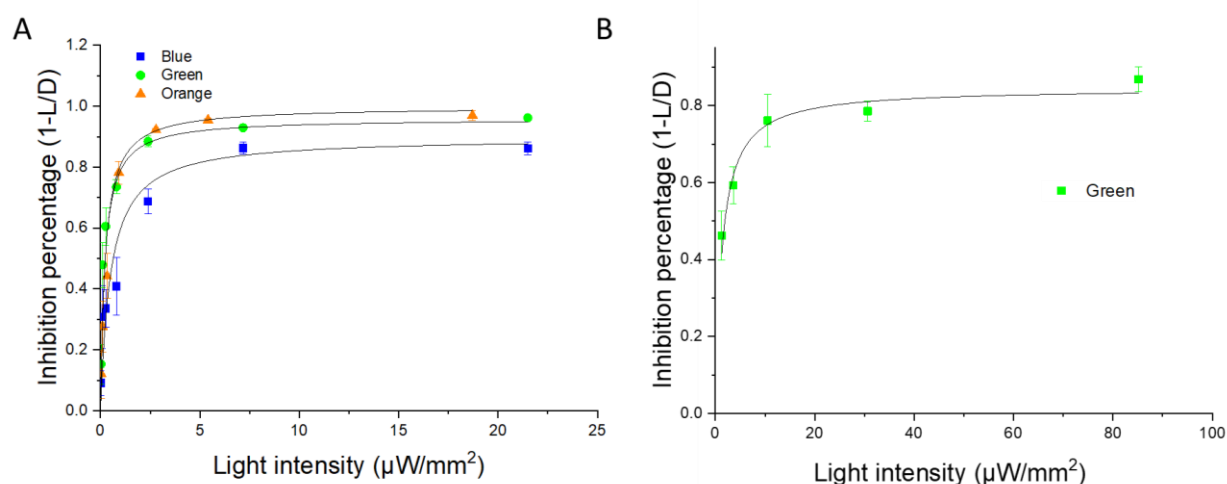
Figure 3.10 Comparison of light sensitivity between *Cr*2c-Cyclop and *Vc*2c-Cyclop.

A, Three different wavelengths of light were employed to test the sensitivity of *Cr*2c-Cyclop. The blue light was used at 473 nm with a range of light intensities similar with green light (532 nm), Orange light (596 nm) was applied in a different range of light intensities. n=3, error bars=SD.

B, A range of green light intensities at 556 nm were applied to obtain the sensitivity of *Vc*2c-Cyclop. n=4, error bars=SD. For A and B graphs, inhibition percentage was calculated by (D-L)/D ratio in each light intensity. 30 ng cRNA injection for each construct, 3dpi. The detailed light intensities see **2.3.6** section.

3.2.6 Action spectra of *Cr*2c-Cyclop and *Vc*2c-Cyclop

Based on above sensitivity test, the action spectrum investigation was operated near the half-maximal inhibition light intensity. Different wavelengths from UV light to red light were applied to confirm the *Cr*2c-Cyclop cyclase activity from similar amounts of photons in individual tests, and then the inhibition percentage 1-L/D was represented as the inhibition effect. The action spectrum of *Cr*2c-Cyclop peaked at ~541 nm (**Figure 3.11 A**), similar with *Be*Cyclop action spectrum peaking at ~530 nm from fungi. While *Vc*2c-Cyclop action spectrum showed a slight red shift, peaking at ~556 nm (**Figure 3.11 B**).
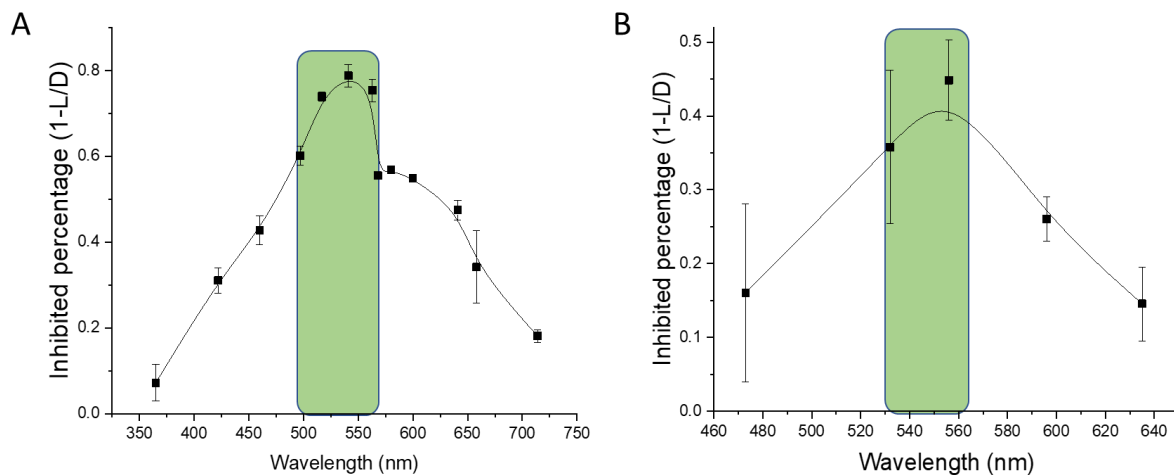
Figure 3.11 Investigation of action spectra in both 2c-Cyclops.

A, Action spectrum of *Cr*2c-Cyclop was investigated from UV-light 365 nm to red light 714 nm with individual light intensities near half-saturated inhibition. The action spectrum peaked at ~541 nm.

n=3, error bars=SD.

B, Action spectrum of *Vc*2c-Cyclop was tested from blue light 473 nm to red light 635 nm with individual light intensities close to half-saturated inhibition. The action spectrum peaked at ~556 nm.

n=4, error bars=SD.

For A and B, 30 ng cRNA injection for each construct, 3dpi. The inhibited percentage in each wavelength light was calculated from similar amounts of photons effects. Detailed applied wavelengths and light intensity values see **2.3.6** section.

3.2.7 GC activity of *Cr*2c-Cyclop under different reaction conditions

To identify the dynamic activity of *Cr*2c-Cyclop, the reaction processes were applied at different time points with or without light illumination. After 30 seconds dark, the first treatment was kept on constant dark, the second treatment was transferred to constant light illumination, and the third one was illuminated for 3 min and then transferred to the dark stage. From cGMP measurement, we could observe that the reaction velocity changed back to a dark condition within 30 s in the third sample treatment (**Figure 3.12**). Therefore, this suggests that the photocycle of *Cr*2c-Cyclop can recover from light to the dark state less than 30 s.
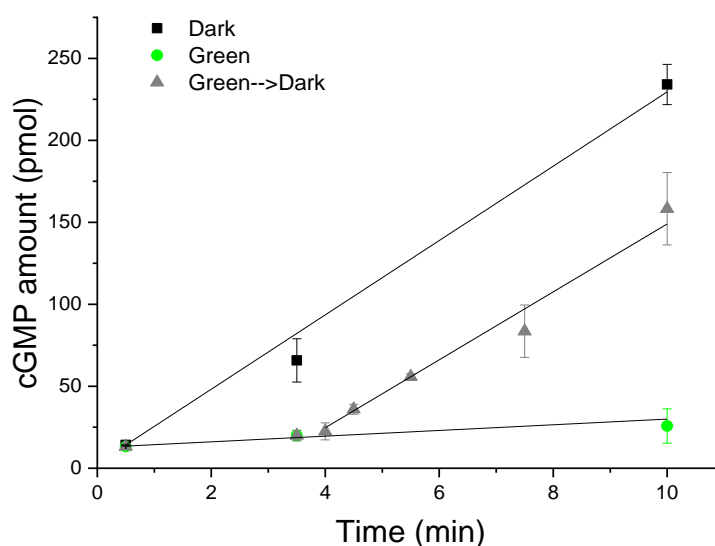
Figure 3.12 Dynamic activity of *Cr*2c-Cyclop.

All three individual treatments were under the dark stage for 30 s to collect some cGMP production. Then the first constant dark condition was illustrated in dark square symbol■, while the second treatment was illuminated under green light constantly, labeled with green dot symbol●. The third treated samples were illuminated with green light for 3 min then transferred to dark condition, and finally stopped in certain time points (4, 4.5, 5.5, 7.5, 10 min, with gray triangle symbol▲). cGMP production was calculated from one oocyte membrane expressing *Cr*2c-Cyclop, 30 ng cRNA injection, 3dpi. n=3, error bars=SD.

Based on previous tests with different metal ions conditions, nucleotide cyclase activity can be strongly affected by certain metal ions like Mn^{2+}, Mg^{2+}, Ca^{2+} etc. The standard reaction buffer was performed with 5 mM Mg^{2+}. The cyclase activity under dark and light conditions was not detected in absence of Mg^{2+} with additional 1 mM EDTA. When 5 mM Ca^{2+} was added in the

standard reaction buffer, the cyclase activity in dark strongly reduced as light activity in the standard condition, and the dark to light ratio decreased to ~2 fold. This suggested Ca^{2+} probably can compete for Mg^{2+} binding sites in the cyclase domain (**Figure 3.13**). The enzyme activities in both dark and light conditions dramatically reduced without ATP (as phosphoryl group donor), suggesting that phosphorylation could be required for the catalytic process.
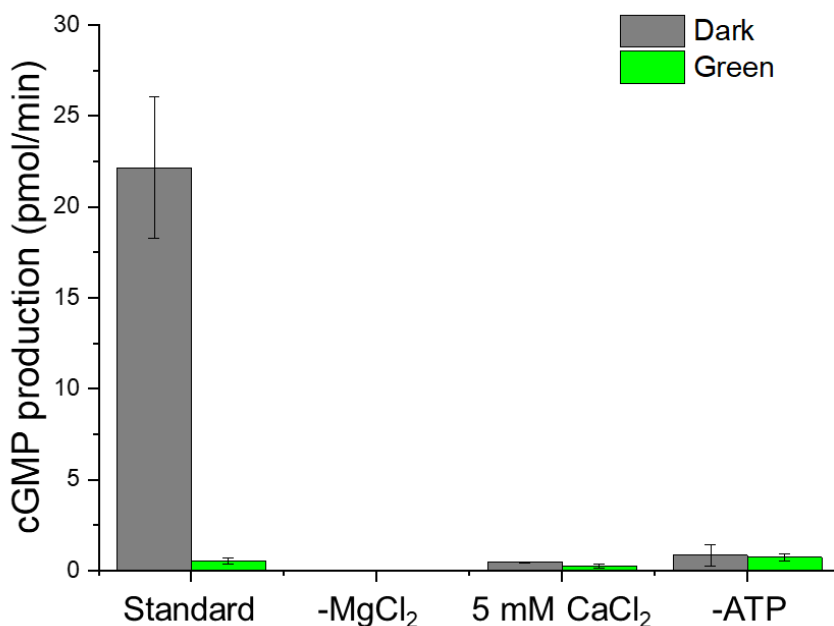
Figure 3.13 *Cr*2c-Cyclop activity depends on ATP, Mg^{2+}, and Ca^{2+}.

The standard reaction buffer was used as previous optimization system, including 100 mM NaCl, 75 mM Tris-HCl, 5 mM $MgCl_2$, 5 mM DTT, 0.2 mM GTP, 0.25 mM ATP, pH adjusted to 7.3. -$MgCl_2$ means Mg^{2+} was removed in the extracted buffer and above reaction buffer. 5 mM $CaCl_2$ was further added in standard reaction buffer. -ATP represents that ATP was depleted in reactive solutions.

All cGMP production was calculated from one oocyte membrane expressing Cr2c-Cyclop, 30 ng cRNA injection, 3dpi. n=4, error bars=SD.

*Cr*2c-Cyclop activity and D/L ratio can also depend on other influenced elements such as pH and temperature. From 10 °C to 30 °C temperature, the dark and light activity enhanced with increasing temperature, peaking at 30 °C, but the D/L ratio kept the highest at 20 °C. Further increasing temperature to 40 °C dramatically dropped the dark activity and D/L ratio (**Figure 3.14 A**). At pH 7.3, both D/L ratio and dark activity were remained the highest value, while the dark activity decreased and D/L ratio dropped to < 10 at pH 6.3 or 8.3 (**Figure 3.14 B**).
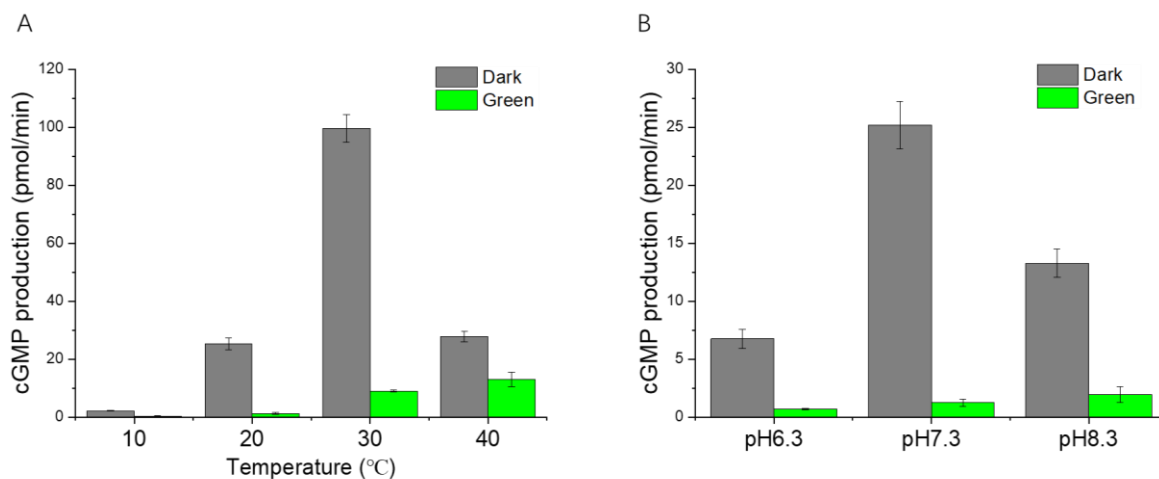
Figure 3.14 Temperature and pH-depended *Cr*2c-Cyclop activity.

A, Four different temperatures (10, 20, 30, 40 ℃) were applied to test cyclase activities in dark and light. B, Three different pH values (pH 6.3, pH 7.3, pH 8.3) were used to compare enzyme activities in both treatments. For both A and B, cGMP production was calculated from one oocyte membrane, 30 ng cRNA injection, 3dpi. n=3, error bars=SD.

3.2.8 Light-regulated mechanism of *Cr*2c-Cyclop

Based on alignment with other microbial opsins, the lysine residue K298 is very conserved and plays crucial roles in all-*trans*-retinal binding. The two components histidine kinase domain and response regulator are naturally integrated into the full-length of *Cr*2c-Cyclop protein, which also contain conserved residues and motifs after alignment with typically separated histidine kinases and response regulators (e.g. HK853-RR468 pair, EnvZ-OmpR pair) in two different bacteria (**Figure 3.15**).

Histidine Kinase

```
                                                   H-box
Cr2cCyclop   : ---MIQDLRDSVTRKDQFMSLMSHELRTPLNGIIQLSDAL :   368
Vc2cCyclop   : ---MIQELRDAVTRKDQFMSLMSHELRTPLNGIIQLSDAL :   352
HK853        : ESKELERLKRIDRMKTEFIANISHELRTPLTAIKAYAETI :   276
EnvZ         : FNHMAAGVKQLADDRTLLMAGVSHDLRTPLTRIRLATEMM :   259
```

```
                      G-box
Cr2cCyclop   : TRKYGGTGLGLNIVKQLVEAHEGTIEVASVEGRGTTFTVE :   563
Vc2cCyclop   : TRKYGGTGLGLNIVKQLVEAHEGKIEVQSAEGRGTTFTVT :   547
HK853        : TYEVPGTGLGLAITKEIVELHGGRIWVESEVGKGSRFFVW :   475
EnvZ         : -RTISGTGLGLAIVQRIVDNHNGMLELGTSERGGLSIRAW :   435
```

Response regulator

```
Cr2cCyclop   : PDLILLDCMMPVMSGHEFCATLRKVIPGNVLPVIMVSAKSDEENIVEGLR :   1135
Vc2cCyclop   : PDLILLDCMMPNMSGHEFCATLRKVIPGNVLPVIMVSAKSDEDNIVEGLR :   1070
RR468        : PDLIVLDIMMPVMDGFTVLKKLQEKEEWKRIPVIVLTAKGGEEDESLALS :   96
OmpR         : FHLMVLDLMLPGEDGLSICRRLRSQ--SNPMPIIMVTAKGEEVDRIVGLE :   96
```

Figure 3.15 Alignment of His-Kinases (HK) and Response regulators (RR) in both 2cCyclOps and two typical HK-RR pairs from bacteria.

Residues with black background color are key residues to accept or release phosphoryl group, H-box and G-box are labeled with black boxes. HK853 and RR468 are partners coming from *Thermotoga maritima* MSB8, while EnvZ and OmpR were identified from *Escherichia coli*. Accession No.: HK853: NP_228662.1, EnvZ: WP_069357419.1, RR468: AAD35552.1, OmpR: CDZ22180.1

The H-box and G-box motifs in histidine and kinase domain play roles as phosphorylation and ATP binding regions. In detail, the histidine H352 located in H-box and glycine G533 in G-box could be proposed to determine autophosphorylation and ATP binding, respectively. The aspartic acid D1092 is the key residue in response regulator to accept phosphoryl group from H352 and could regulate downstream cyclase activity. Therefore, ATP deprivation and single mutagenesis analysis in these points will help to prove the above hypothesis. More importantly, YFP tag was fused in the C-terminal of *Cr*2c-Cyclop wild type as well as other single mutants. This would be helpful to quantify cGMP production and normalize activities to the same expression level of proteins based on fluorescence emission values.

The typical two-component system should require ATP as phosphoryl group donor, thus transferring phosphate from histidine residue H352 to aspartic acid residue D1092. This will further influence output domain cyclase activity. From *in vitro* assay, *Cr*2c-Cyclop-YFP exhibited no obvious light-regulated effect in absence of ATP, and the dark activity dramatically decreased similar with *Cr*2c-Cyclop-YFP activity at 0.25 mM ATP under light illumination (20 µW/mm^2, 532 nm). In addition, non-hydrolyzable ATP analog AMP-PNP was applied to mimic ATP to bind the kinase domain, and this cannot recover the activity (**Figure 3.16**). Therefore, it is suggested that ATP is required in dark to activate the GC activity.
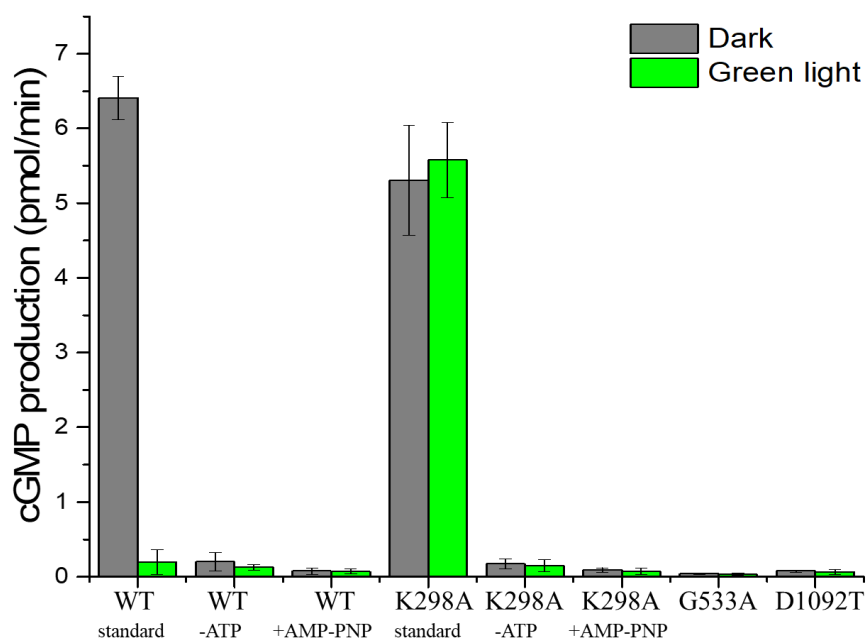
Figure 3.16 Mutant analysis and ATP-dependence of *Cr*2c-Cyclop.

All constructs used here were fused with YFP tag in the C-terminal for comparison. In *Cr*2c-Cyclop-YFP wild type (represents as WT), the dark and light activity tested under standard reaction buffer was compared with activity without ATP or with AMP-PNP. Similarly, *Cr*2c-Cyclop-YFP K298A mutant was under the same treatment and measurement as wild type. Moreover, two *Cr*2c-Cyclop-YFP G533A and D1092T single mutants were tested for comparison in parallel. cGMP production was measured from one oocyte membrane. All constructs were injected the same amount of ~30 ng cRNA, 3dpi, and final cyclase activities were normalized to the same fluorescence emission value, n=4, error bars=SD.

From mutation analysis, K298A mutant can destroy the light-regulation of *Cr*2c-Cyclop. This mutant recovered the light activity similar with the dark activity of wt. Deprivation of ATP in reaction buffer strongly reduced K298A activity to the light-inhibited level of wt. This proved that K298A mutant would remain conformation as wild type in dark and cannot be influenced by light illumination. It was demonstrated that phosphoryl group transfer in dark is a necessity to activate GC continuously.

The H352F mutant destroyed phosphoryl transfer and decreased GC activity lower than the activity of wt under light illumination. Another G533A mutant cannot bind and hydrolyze ATP, thus leading to inhibit cyclase activity (**Figure 3.16**). Both mutants have inhibition effects as light activity in wt primarily because of ATP hydrolysis defect or phosphorylation site blocking. Further D1092T mutant would be expected to block the phosphorylation site in response regulator. Obviously, the activity of D1092T mutant is also similar with the light-inhibition effect of wt. This suggested that D1092 phosphorylation is necessary for cyclase activity.

3.2.9 *Cr*2c-Cyclop may function as a dimer

Normally, nucleotide cyclases function as dimers. To briefly test the dimerization of *Cr*2c-Cyclop, the wt and H352F mutant were co-expressed to detect the GC activity (**Figure 3.17 A**). The YFP fusion proteins can be quantified by fluorescence emission values. The wt and H352F mutant showed similar expression levels after injection of 10 ng and 20 ng amounts of RNA separately (**Figure 3.17 B**). When half amount of wt and half amount of H352F proteins were co-expressed, the dark activity was ~1/4 of activity from only wt expressing membrane (**Figure 3.17 A**). This suggested the homo-dimer of wt mainly contributed that 1/4 activity, and the heterodimer of wt and H352F revealed a less activity like the homo-dimeric mutant. This gives hints that *Cr*2c-Cyclop needs to function as a dimer, while H352F mutant and wt co-expression would inhibit the activity.
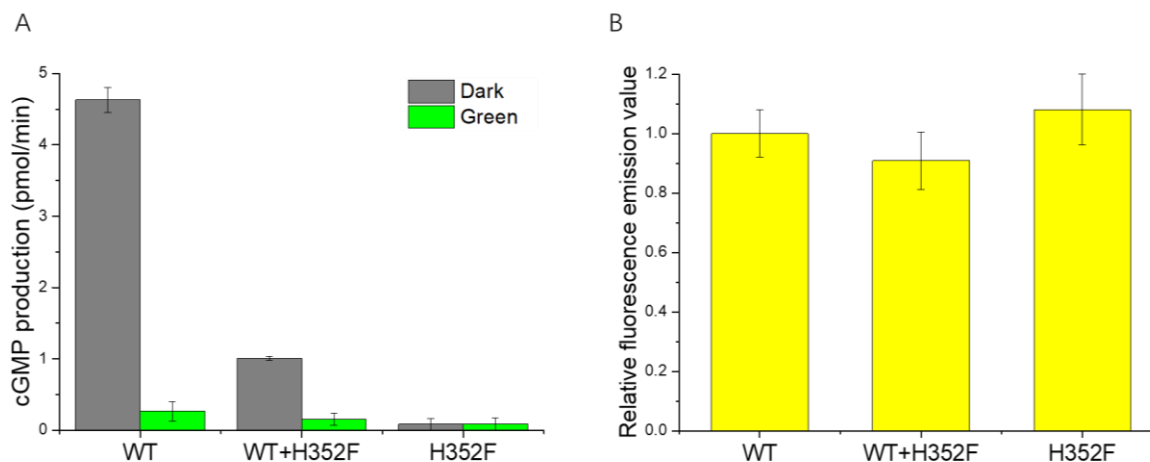
Figure 3.17 Co-expression of *Cr*2c-Cyclop-YFP and its H352F mutant.

A, Dark and light activities were compared among *Cr*2c-Cyclop-YFP, H352F mutant, and co-expression constructs. *Cr*2c-Cyclop-YFP cRNA was injected 10 ng, while H352F mutant cRNA was injected 20 ng in each oocyte, 3dpi. For co-expression, 10 ng *Cr*2c-Cyclop-YFP cRNA and 20 ng H352F mutant cRNA was mixed to inject in each oocyte. All the activities came from one oocyte membrane and were normalized to the similar expression level, n=3, error bars=SD.

B, Relative fluorescence emission values were normalized with wild type as 1 after subtraction of control emission values, n=3, error bars=SD. Because co-expression of WT+H352F constructs showed ~2 times emission value than the other two constructs. The emission value of co-expression was reduced half, so as the activities from co-expression of WT+H352F constructs in graph A.

3.2.10 Investigation of truncated response regulator and cyclase fragment

Mutating *Cr*2c-Cyclop D1092 was predicted to hamper the phosphorylation of the response regulator. The D1092 is close to the cyclase domain, the mutation here has the probability of directly changing the cyclase conformation, thus influencing the GC activity. To exclude this, we generated the truncated *Cr*2c-Cyclop with only the response regulator and cyclase domain (RC) and the same point mutant RC D1092T (**Figure 3.18 A**). After extraction of those two soluble proteins by dialysis and *in vitro* assay, the RC D1092T showed similar activity to the wt RC (**Figure 3.18 B**). The results revealed that truncated cyclase activity was similar with light-inhibited GC activity of full-length wt, because both truncated proteins exhibited over 10 times expression level than the *Cr*2c-Cyclop full-length protein. Meanwhile, GC activities of

truncated RC wt were similar with or without ATP, which suggested D1092 would require phosphoryl group from the upstream H352 residue in his-kinase to further activate guanylyl cyclase activity. This proved that the aspartic acid to threonine substitution will not affect the conformation of the cyclase part but only the phosphorylation.
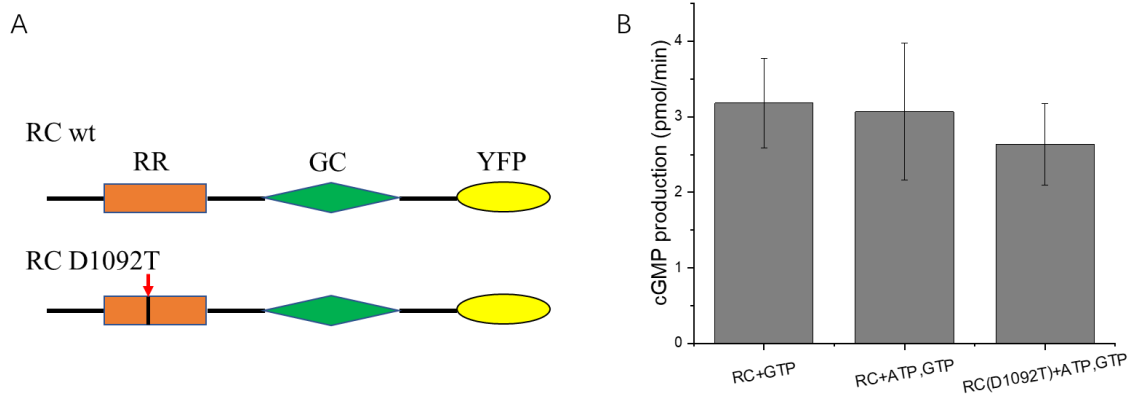
Figure 3.18 Comparison of activities between truncated RC and RC D1092T.

A, The model illustrates truncated fragments including response regulator (RR), guanylyl cyclase (GC) and YFP tag in the C-terminal. RC wt means the wild type of truncated fragment. The position of RC D1092T mutant referred to D1092 position in the full-length of *Cr*2c-Cyclop.

B, *In vitro* assay after dialysis from the cytosol of oocytes. 50 kDa dialysis filter columns were used in dialysis processes. RC+GTP: reaction with 0.2 mM GTP at room temperature; RC+ATP, GTP: reaction with 0.25 mM ATP, 0.2 mM GTP; n=4, error bars=SD. cGMP production (pmol/min) was produced from one oocyte cytosol. All activities were normalized to the same expression level.

3.2.11 The schematic working model of *Cr*2c-Cyclop

From the above results, the working mechanism of *Cr*2c-Cyclop could be clarified. In the dark state, the phosphorylation and phosphoryl transfer can be continuous from histidine kinase H352 to response regulator D1092 with ATP as phosphoryl group donor (**Figure 3.19**). This enables the cyclase to maintain the activity. In contrast, light illumination can initiate a conformational change of rhodopsin domain and then interrupt phosphorylation and phosphate transfer, thus leading to decrease cyclase activity.
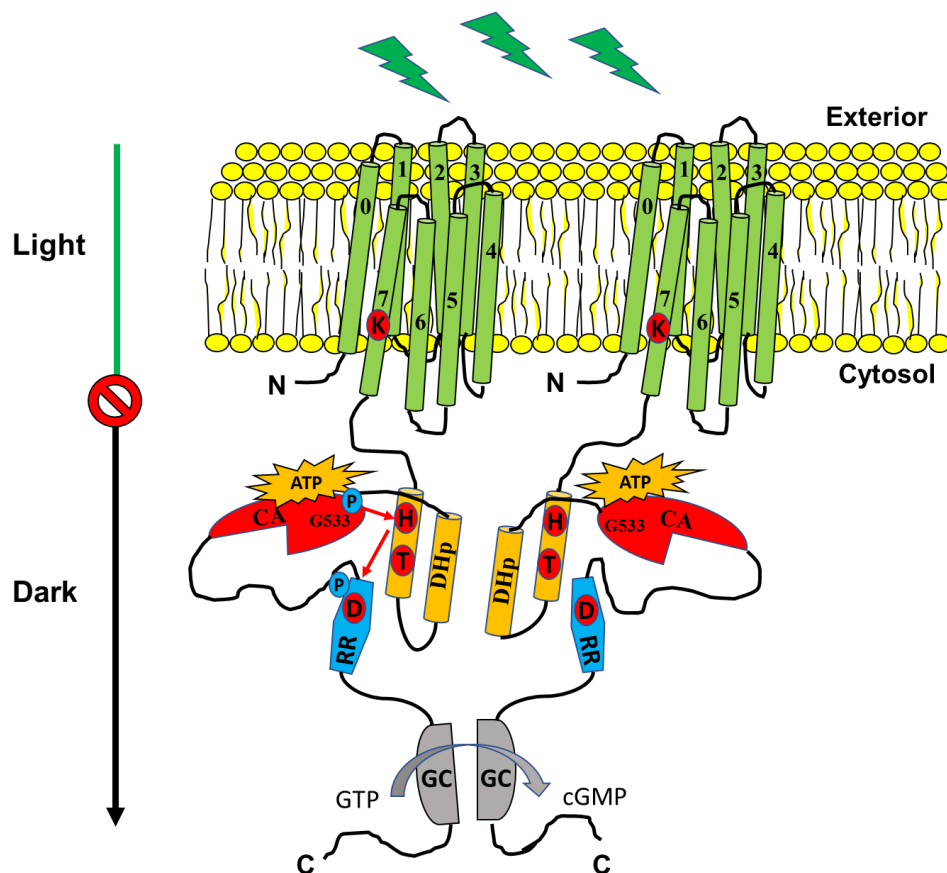
Figure 3.19 The working model of *Cr*2c-Cyclop.
Rhodopsin domain integrates into the membrane with the key K298 residue in the last transmembrane helix binding retinal. Histidine kinase domain was depicted with DHp (Dimerization and histidine phosphotransferase domain) in the orange module and CA (Catalytic and ATPase domain) in the red module, including key residues H352, T356, and G533. Response regulator (RR) was drawn in a blue module with key D1092 residue to accept phosphoryl group. Guanylyl cyclase (GC) domain was depicted in the gray module, producing cGMP from the substrate GTP.

In the following **3.3** section, all figures are based on the following publication:

Statement of individual author contributions in the publication:

Publication (complete reference):
Yuehui TIAN, Shiqiang GAO, Shang YANG and Georg NAGEL. A novel rhodopsin phosphodiesterase from *Salpingoeca rosetta* shows light-enhanced substrate affinity. Biochemical Journal, Mar 26, 2018, 475 (6) 1121-1128. (DOI: 10.1042/BCJ20180010)

Participated in	**Author Initials**, Responsibility decreasing from left to right				
Study Design	Y.T.	S.G	G.N.		
Data Collection	Y.T.	S.G	S.Y.		
Data Analysis and Interpretation	Y.T.	S.G	G.N.		
Manuscript Writing	S.G	G.N.	Y.T.	S.Y.	

All figures and tables included in this section are also illustrated in the publication.

I confirm that I have obtained permission from both the publishers and the co-authors for legal second publication and the correctness of the above mentioned assessment.

I also confirm my primary supervisor's acceptance.

Yuehui Tian	23.07.2018	Würzburg	
Doctoral Researcher's Name	Date	Place	Signature

3.3 A novel rhodopsin phosphodiesterase (*Sr*RhoPDE) from *Salpingoeca rosetta*

After the genomic sequence of the protist *Salpingoeca rosetta* was annotated in the gene bank, another light-induced phosphodiesterase was identified. In my study, I confirmed that *Sr*RhoPDE belongs to light-activated PDE, hydrolyzing both cGMP and cAMP. This enzyme opsin could be a new optogenetic tool to manipulate second messengers by light.

3.3.1 *Sr*RhoPDE is an 8-TM microbial opsin

In our previous experiment, the rhodopsin domain of *Be*Cyclop was proved with N- and C-termini in the cytosolic part, most likely containing 8 transmembrane helices. This *Sr*RhoPDE was then confirmed as an 8-TM opsin by BiFC experiment.

Based on sequence analysis, *Sr*RhoPDE was also predicted as an 8-TM protein with an extra transmembrane helix TM0 in the N-terminal, with PDE domain in the C-terminal (**Figure 3.20 A**). A BiFC construct was generated by fusing two separated parts of YFP (C-terminal part of YFP, YC; N-terminal part of YFP, YN) to both N- and C- termini of opsin part of *Sr*RhoPDE separately, as shown in **Figure 3.7 A**. The N- and C-terminal parts of opsin domain are on the same side of plasma membrane because of the clear fluorescence (**Figure 3.20 B**). Obviously, both termini have to be cytosolic and opsin domain should have 8-TM based on cytosolic localization of PDE activity and alignment with other rhodopsins.
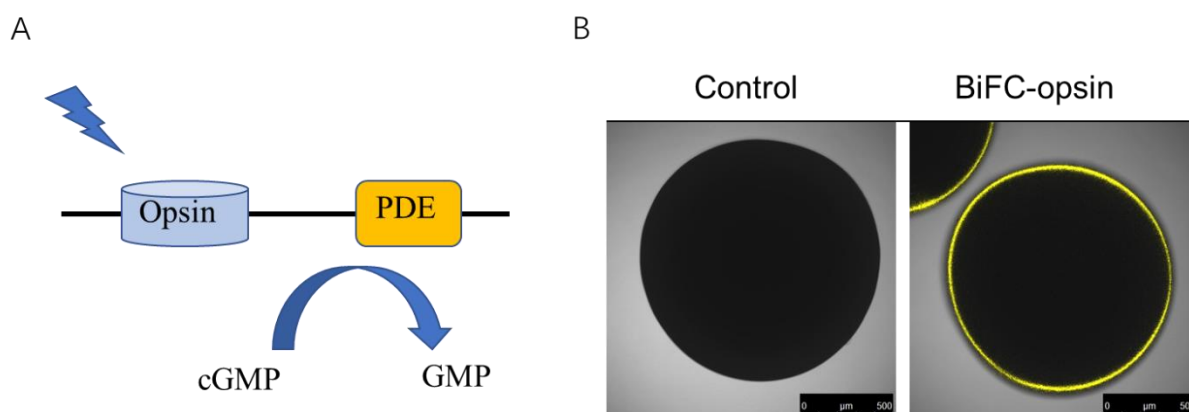
Figure 3.20 *Sr*RhoPDE is an 8-TM rhodopsin.

A, The RhoPDE protein model (including N-terminal opsin domain and C-terminal PDE domain).

B, Strong fluorescence is observed in BiFC-opsin expressing oocytes. After 20 ng cRNA injection for 3 days,

the control oocytes and BiFC-opsin expressing oocytes were taken pictures by a confocal microscope.

3.3.2 *Sr*RhoPDE shows light-activation of cGMP and cAMP hydrolysis

In vitro test, *Sr*RhoPDE expressing oocytes membrane was extracted and mixed with reaction buffer with ~100 μM cGMP or cAMP substrate as beginning concentration. It is clear to see that light can increase *Sr*RhoPDE hydrolysis activity compared with the dark condition. However, no membrane-bound PDE exists in control oocytes membrane without *Sr*RhoPDE expression because no cGMP or cAMP hydrolysis activity can be detected. Here we observed L/D ratio ~2 with 100 μM cGMP at the beginning compared with ~1.4 fold light activation in a previous publication (**Figure 3.21 A**). Obviously, the L/D ratio can reach to ~3 fold starting with 25 μM cGMP. cGMP hydrolysis activity in light is >100 times higher than cAMP hydrolysis activity. The cAMP hydrolysis activity of *Sr*RhoPDE was increased by light ~5 fold (**Figure 3.21 B**). The intensity of blue laser light (at 473 nm) was adjusted to ~0.1 mW/mm^2 in all experiments of the **3.3** section.
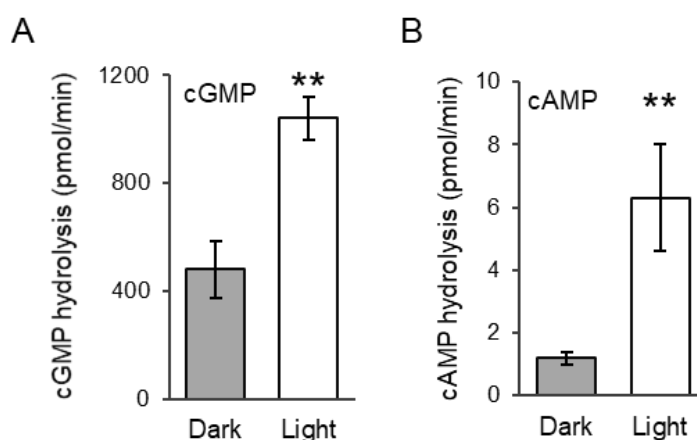
Figure 3.21 RhoPDE is a light-activated enzyme opsin.

A, Comparing cGMP hydrolysis activity of RhoPDE under dark and light, starting with 100 μM cGMP.

B, Comparing cAMP hydrolysis activity of RhoPDE under dark and light, starting with 100 μM cAMP.

For both A and B, final activities were calculated from one oocyte membrane (30 ng cRNA injection, 3dpi), n=4, error bars=SD. Statistics were done by ANOVA one-way test, **P<0.01, ***P<0.001, all following figures about RhoPDE activity test are based on this analysis.

3.3.3 Mg^{2+} is an essential cofactor to activate *Sr*RhoPDE activity

Magnesium ions have a crucial effect on *Sr*RhoPDE activity with 25 μM cGMP at the beginning. Without Mg^{2+}, *Sr*RhoPDE activity was dramatically decreasing and there was no obvious light regulation. *Sr*RhoPDE activity and L/D ratio were increasing when Mg^{2+} concentration increased (**Figure 3.22**). Light activation became ~2 fold with 1 mM Mg^{2+}, while PDE activity and L/D ratio (~3 fold) were similar peaking at 2 mM and 5 mM Mg^{2+}.
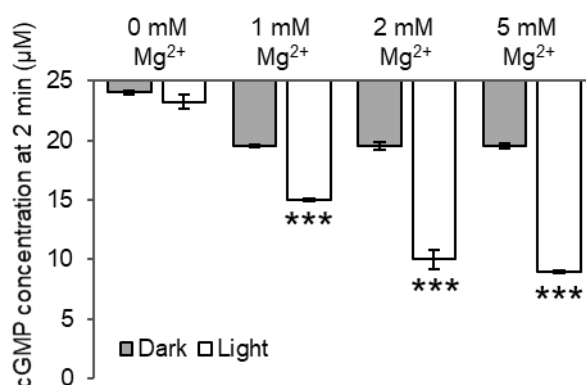
Figure 3.22 Mg^{2+}-dependence of *Sr*RhoPDE activity.
cGMP hydrolytic activity of RhoPDE was tested at different Mg^{2+} concentrations. Without Mg^{2+}, 1 mM EDTA was added. Reactions in dark and light were operated with membrane extracted from one oocyte (30 ng cRNA injection, 3dpi), starting with 25 μM cGMP. n=4, error bars=SD.

3.3.4 YFP fusion of *Sr*RhoPDE

Single site-directed mutagenesis in essential residues can be helpful to explain the mechanism of *Sr*RhoPDE in detail. To exclude the different expression levels in different mutants, it is necessary to quantify amounts of expressed proteins. Therefore, YFP tag was fused in individual wild type or mutants for further comparison.

The YFP tag was fused in the N- or C-terminal of RhoPDE construct. As this is an 8-TM protein in contrast to 7-TM classic rhodopsins with an extracellular N-terminal, the YFP tag should localize in the cytosolic N-terminal side. The results demonstrated that YFP fusion on N- or C-terminal has distinct influences. The YFP-*Sr*RhoPDE construct showed over two times higher expression level than *Sr*RhoPDE-YFP construct (**Figure 3.23 A**). The light activation effects between YFP::*Sr*RhoPDE and *Sr*RhoPDE were similar with cGMP hydrolysis. However, the

expression of *Sr*RhoPDE-YFP reduced and further influenced light activation due to increased dark activity (**Figure 3.23 B**). When measuring cAMP hydrolysis starting at 100 μM, YFP::*Sr*RhoPDE construct showed ~2.6 fold light activation, instead of ~5 L/D ratio without YFP fusion (**Figure 3.23 C**). Therefore, it is reasonable to further compare YFP::*Sr*RhoPDE construct and its mutants.
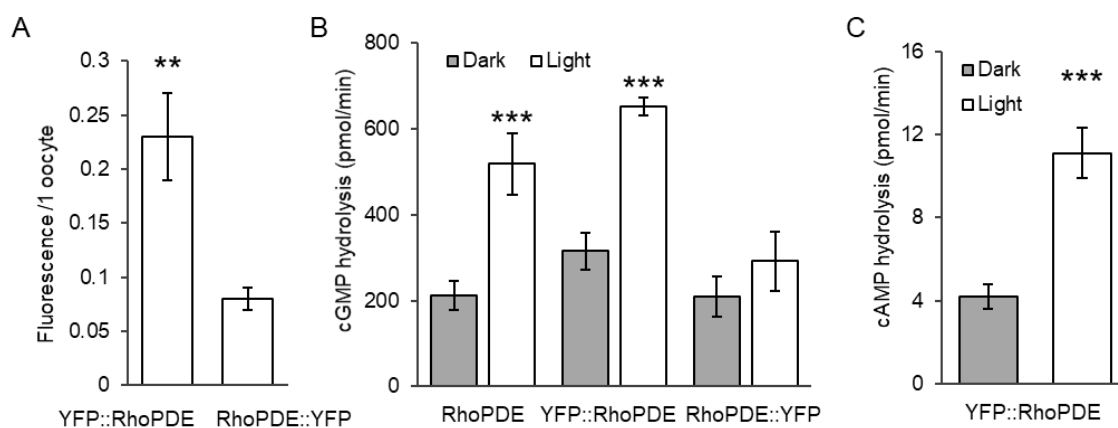
Figure 3.23 Expression and activity of N- or C- terminal YFP fusion RhoPDE.

A, Comparison expression levels between YFP::RhoPDE and RhoPDE::YFP by different fluorescence emission values. Each measurement came from one oocyte membrane, n=3, error bars=SD.

B, cGMP hydrolytic activities comparison among RhoPDE, YFP::RhoPDE and RhoPDE::YFP under dark and light. n=4, error bars=SD.

C, cAMP hydrolysis of YFP::RhoPDE, n=3, error bars=SD.

For B and C, final activities were calculated from one oocyte membrane proteins, starting with 100 μM cGMP or cAMP. 20 ng cRNA of RhoPDE was injected in each oocyte, 3 dpi. 30 ng cRNA of YFP::RhoPDE or RhoPDE::YFP was separately injected in each oocyte, 3 dpi.

3.3.5 Mutation of the retinal-binding lysine abolishes light regulation of *Sr*RhoPDE

According to alignment with other typical opsins (*Be*Cyclop, Channelrhodopsin-2, and Bacteriorhodopsin), the lysine residue (K296) of *Sr*RhoPDE was predicted as a highly conserved amino acid for covalent binding of all-*trans*-retinal (**Figure 3.24 A**).

Two YFP::*Sr*RhoPDE K296A and K296M mutants were generated and tested together with YFP::*Sr*RhoPDE wild type (YFP::wt). Both mutants showed similar expression levels as wild type (**Figure 3.24 B**). As expected, there was no light regulation in K296A and K296M mutants

disturbing retinal binding (**Figure 3.24 C**). The wild type YFP::*Sr*RhoPDE showed ~2.3 fold light activation from the same batch of oocytes and reaction conditions (starting with 100 μM cGMP).
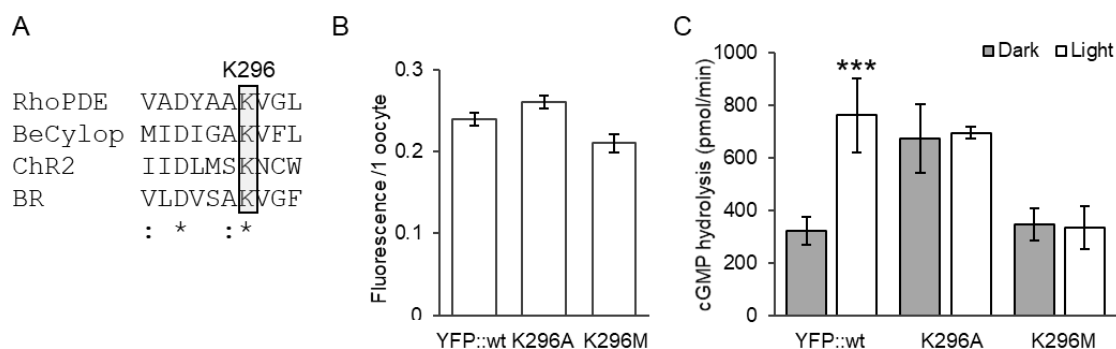
Figure 3.24 Effects on activities of YFP::RhoPDE K296A or K296M mutants.

A, The conserved lysine K296 from alignment are shown to bind chromophore all-*trans*-retinal by forming Schiff base.

B, Fluorescence emission values of YFP::RhoPDE and K296A, K296M mutants.

C, cGMP hydrolytic activity of YFP::RhoPDE and K296A, K296M mutants. Final activities were calculated from one oocyte membrane, starting with 100 μM cGMP. 30 ng cRNA of those three constructs was injected separately in each oocyte. For B and C, n=3, error bars=SD.

The fluorescence intensity of wild type or mutants expressing membranes were detected to quantify the amount of expressed proteins. Based on materials and methods in **2.3.4**, we used fluorescence intensities of standard YFP amounts to obtain a calibration curve. Wild type and its two mutants expressed about 0.5 pmol per oocyte individually. The turnover number of YFP::*Sr*RhoPDE was calculated to be 25 ± 5 s^{-1} at 100 μM cGMP in light, similar with previous determination from Lamarche et al, ranging from 19-28 s^{-1}. We obtained 12 ± 2 s^{-1} turnover number under dark condition (at 100 μM cGMP). Interestingly, under dark treatments, YFP::*Sr*RhoPDE K296A mutant showed a higher turnover (~23 s^{-1}), which is similar with wt under light illumination, and K296M mutant showed a lower turnover about 13 s^{-1}, more close to wt in the dark (**Table 3.1**). This gave hints that the two lysine mutants could keep as two conformations to influence *Sr*RhoPDE activity.

Table 3.1 Turnover (cGMP hydrolysis) of YFP::*Sr*RhoPDE and K296A, K296M mutants

	Protein amount	Dark turnover (s^{-1})	Light turnover (s^{-1})
YFP::RhoPDE	0.46 ± 0.04 pmol/oocyte	12 ± 2	28 ± 5
K296A	0.50 ± 0.04 pmol/oocyte	22 ± 4	23 ± 0.8
K296M	0.41 ± 0.05 pmol/oocyte	14 ± 2	13 ± 3

Turnover refers to reaction starting with 100 μM cGMP at 20 ℃. Based on a standard YFP fluorescence in the same buffer, fluorescence emission values of YFP::*Sr*RhoPDE K296A and K296M mutants were applied to calculate protein amounts. n=3, errors=SD.

3.3.6 Light-enhanced substrate affinity of *Sr*RhoPDE

We hypothesized that light could have effects on the K_m (Michaelis-Menten constant) value for cGMP because the L/D ratio became larger with lower cGMP concentrations. Therefore, we applied different initial cGMP concentrations ranging from 1 to 250 μM cGMP to test *Sr*RhoPDE activity. Data points represented cGMP hydrolysis activities which were then fitted with a Michaelis-Menten equation. It was demonstrated that the maximum hydrolytic speed was not changing obviously by light, with 1.2 nmol/min in dark and 1.6 nmol/min in light (**Figure 3.25 A and B**). Both calculations came from proteins expressing in one oocyte membrane. The K_m value for cGMP significantly decreased from ~80 μM in dark to ~13 μM in the light. This suggested that the light illumination can enhance substrate affinity. The L/D ratio can increase to ~4 starting with 7.5 μM cGMP (**Figure 3.25 C**).
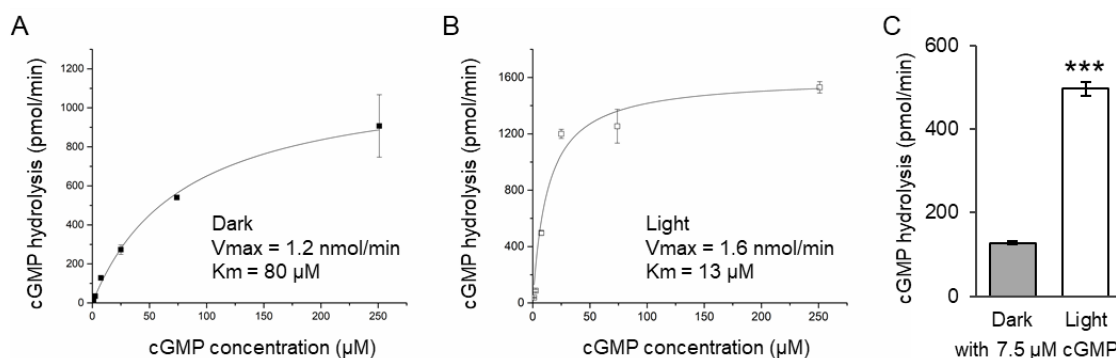
Figure 3.25 Light-enhanced substrate affinity of *Sr*RhoPDE.

A, cGMP-dependence of RhoPDE activity in dark.

B, cGMP-dependence of RhoPDE activity under light illumination.

For both A and B, data points were fitted with Michaelis-Menten equation.

C, Comparison of cGMP hydrolysis of RhoPDE under dark and light, starting with 7.5 μM cGMP as the substrate. All activities were calculated from one oocyte membrane. n=3, error bars=SD.

3.3.7 *Sr*RhoPDE K296A and K296M mutants have different substrate affinities

From previous tests, the K296A mutant showed cGMP hydrolytic activity similar to light activity in wt, while K296M mutant activity was similar to dark activity in wt (**Figure 3.24 C**). This may suggest that both of mutants could have two different conformations. So I further determined cGMP affinity in both light-insensitive K296A and K296M mutants. The reaction of both mutants was individually operated under dark condition because they were completely light-insensitive proteins. As expected, K296A mutant showed a low Km value at ~13 μM in the dark, just the same as wt RhoPDE in light (**Figure 3.26 A**). However, K296M mutant was determined in a Km value of 63 μM (**Figure 3.26 B**), more close to the Km value of 80 μM for wt RhoPDE in dark.
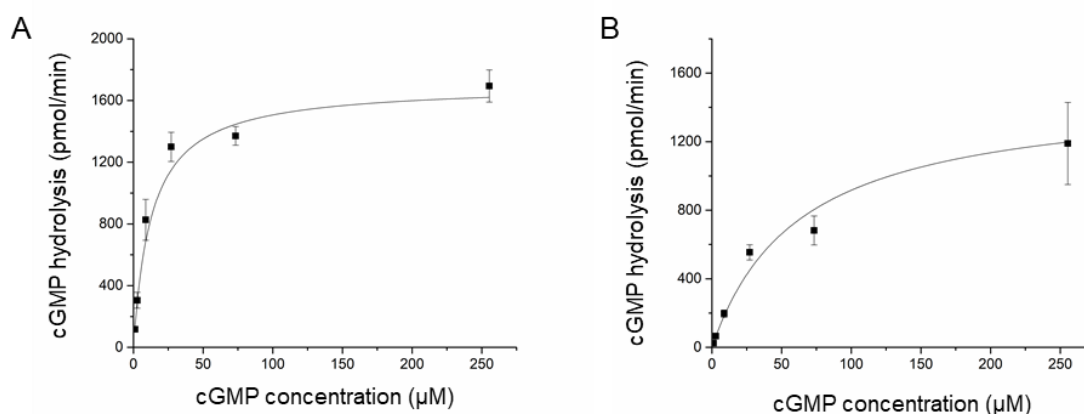
Figure 3.26 Comparison of cGMP affinity between *Sr*RhoPDE K296A and K296M mutants.
A, cGMP affinity tested in RhoPDE/K296A: Km=13 μM, Vmax=1.7 nmol/min.
B, cGMP affinity tested in RhoPDE/K296M: Km=63 μM, Vmax=1.5 nmol/min.
Data points are fitted with Michaelis-Menten equation. Activities were calculated for membranes, extracted from one oocyte. n = 3, error bars = SD.

3.3.8 No membrane-bound phosphodiesterase in *Xenopus* oocytes

To exclude influences from other membrane proteins without RhoPDE expression, control oocytes membrane was also extracted and tested cGMP and cAMP hydrolytic activity. It was

clearly demonstrated that no membrane-bound PDE activity can be detected from control oocytes after longer time reaction in dark and light. No cGMP and cAMP hydrolysis were consumed in **Figure 3.27 A and B**. In contrast, RhoPDE expressing oocytes membrane showed strong cGMP hydrolysis activity, 100 times more than cAMP hydrolysis activity (**Figure 3.27 C and D**).
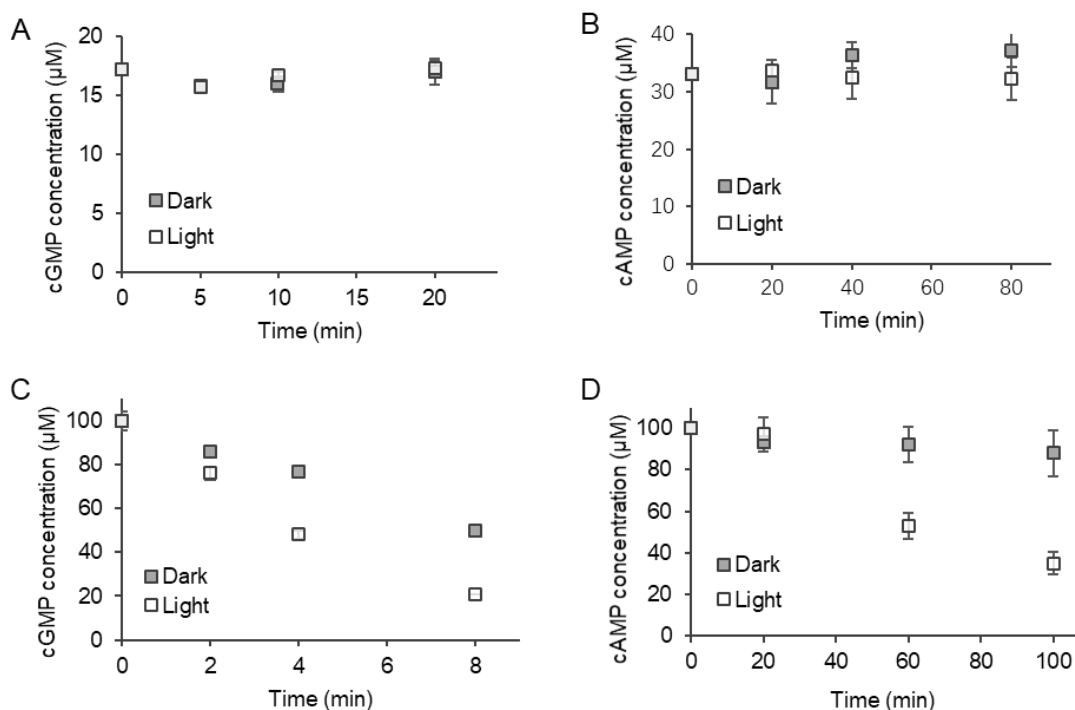
Figure 3.27 Comparison of cGMP and cAMP hydrolysis between control and RhoPDE-expressing membranes

A, No cGMP hydrolysis from control oocytes membranes, without RhoPDE expressing.

B, No cAMP hydrolysis from control oocytes membranes, without RhoPDE expressing.

C, cGMP hydrolysis in dark and blue light, with RhoPDE expression in oocytes membranes.

D, cAMP hydrolysis in dark and blue light, with RhoPDE expression in oocytes membranes.

cAMP was measured with the longer time scale and 8-fold higher membrane concentration (see below) than for cGMP measurement.

For cNMP hydrolysis measurements, membrane pellet from homogenized oocytes was resuspended with 8 µL (A & C) or 1 µL (B & D) solution A per oocyte. The resuspended membrane was mixed with a 9-fold volume of reaction buffer. Reactions were started in the dark or in blue light. 10 µL reaction mix was transferred to 190 µL stop solution at indicated times from dark and light reaction for the cNMP assay. n=3, error bars = SD.

3.3.9 pH-dependence of *Sr*RhoPDE activity

From previous publication, Yoshida *et al.* detected a lower activation effect at pH 6.5 (cGMP hydrolysis L/D: ~1.4) than pH 7.3 (cGMP hydrolysis L/D: ~2) used by us starting from 100 μM cGMP. Here I further compared light-regulated effects of RhoPDE at different pH values (**Figure 3.28**). With 22.5 μM cGMP at the beginning, Light/Dark ratio at pH 7.3 was ~4 fold, while L/D ratio at pH 6.5 was ~3. This also suggested when cGMP concentration increased to 100 μM, the Light/Dark ratio of PDE activity decreased near half. More cGMP concentration will destroy the light-activated effect, that is the reason why Lamarche *et al.* cannot detect light regulation starting with 5 mM cGMP.
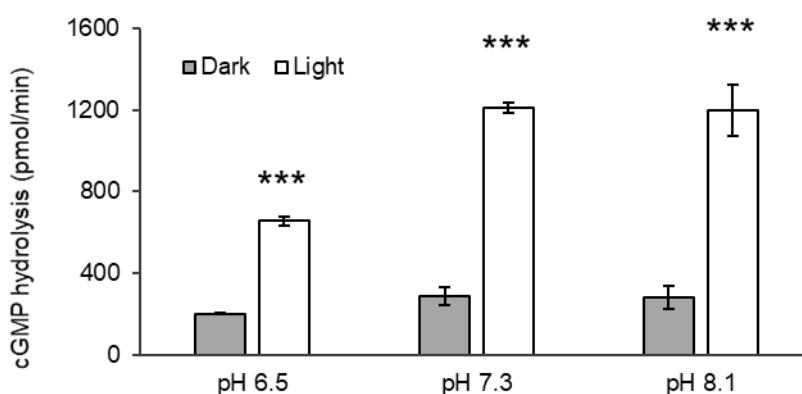
Figure 3.28 pH-dependence of *Sr*RhoPDE activity.
cGMP hydrolysis activities of RhoPDE in dark and light were measured at pH 6.5, pH 7.3, and pH 8.1, starting with 22.5 μM cGMP. Final activities were calculated to membrane proteins extracted from one oocyte. n=3, error bars = SD.

3.4 Other possible strategies to engineer light-regulated phosphodiesterase

There are still many other possibilities to generate light-induced phosphodiesterase. This could unveil some mechanisms of PDE with light-regulated influence during engineering processes. Therefore, it is necessary to do more trials for an efficient optogenetic tool, for example, mutagenesis in opsin domain or PDE catalytic domain of *Sr*RhoPDE to decrease strong dark and light activity, modification of different linker lengths between rhodopsin domain and catalytic PDE domain of *Sr*RhoPDE, chimeras generation fused rhodopsin domain from *Be*Cyclop or *Cr*2c-Cyclop to PDE domain of *Sr*RhoPDE, fusion of LOV domain with PDE2A catalytic domain from human, or fusion of Dronpa domain to PDE2A domain.

3.4.1 Mutant analysis of *Sr*RhoPDE

In the rhodopsin domain of *Sr*RhoPDE, some positions might play roles in light regulation based on alignment with functional mutants of Channelrhodopsin-2. All tested wild type (WT) and mutants were fused with YFP in the N-terminal, and activities under dark and light were normalized to the same expression level except for M564F mutant without expression (**Figure 3.29**). The results showed that cGMP hydrolysis activity of both C169T and T192C mutants increased over 2 times than wt in dark, while less than 2 times light activation led to lower light/dark ratios. D189C mutant activity decreased to 1/3 of wt, but it still remained light to dark ratio at ~3. In addition, in PDE catalytic domain, several mutants closed to substrate binding region were generated based on available PDE structure[92]. F451W mutant activity was strongly reduced, and light/dark ratio decreased to ~1.5. M564F was observed no activity because of no expression. The remaining mutants including M459D, F627W, F627Y still remain high cGMP hydrolysis activity under dark and light, but the L/D ratios lower than wt.
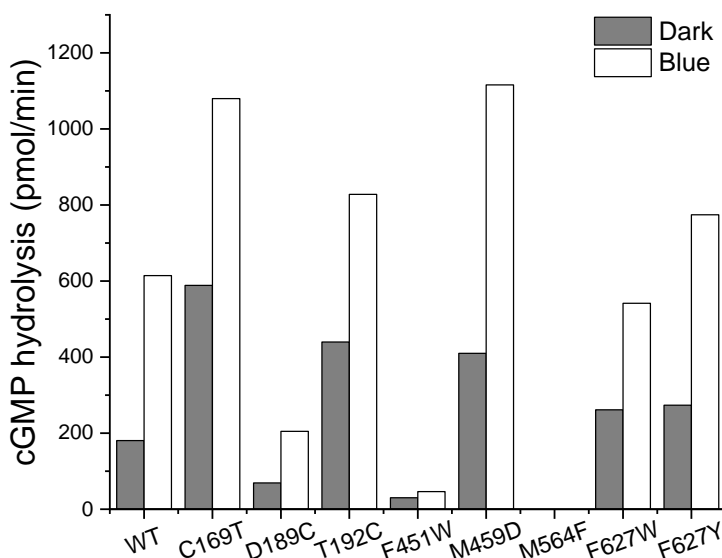
Figure 3.29 Single site-directed mutagenesis analysis of YFP-*Sr*RhoPDE.
All the single mutants were generated by Quickchange PCR with YFP-*Sr*RhoPDE as the template. WT represents YFP-*Sr*RhoPDE wild type. All the activities came from one oocyte membrane expressing individual construct under dark and blue light illumination (~0.1 mW/mm^2), starting with 24 μM cGMP. Moreover, they were normalized to the same level of fluorescence emission value.

Interestingly, two single mutants were further selected with lower cGMP hydrolysis activity (**Figure 3.30 A**). In PDE catalytic domain, L623F mutant activity decreased to ~1/40 of wt in dark and light (L/D ratio ~4), while E657Q mutant activity was ~1/30 of wt (L/D ratio ~5). RhoPDE wild type has very strong activity in both dark and light conditions. Therefore, it is meaningful to obtain less cGMP hydrolytic activity but still keep L/D ratio, and both single mutants can be the candidates for further application. Until now, I tested several double mutants including F451W+L623F, M564F+L623F, S620Y+L623F, and M459D+E657Q, which had obvious expression in oocytes but no cGMP hydrolytic activity can be observed.

Based on alignment with cAMP-specific PDE enzymes, M459D and E657Q mutants might increase cAMP hydrolysis activity. The results indicated that cAMP hydrolytic activity of M459D mutant increased to ~8 times higher than wt (**Figure 3.30 B**). E657Q mutant showed ~2 times higher than wt. This suggested that M459D mutant could be the candidate to be further engineered as cAMP-specific PDE enzyme.
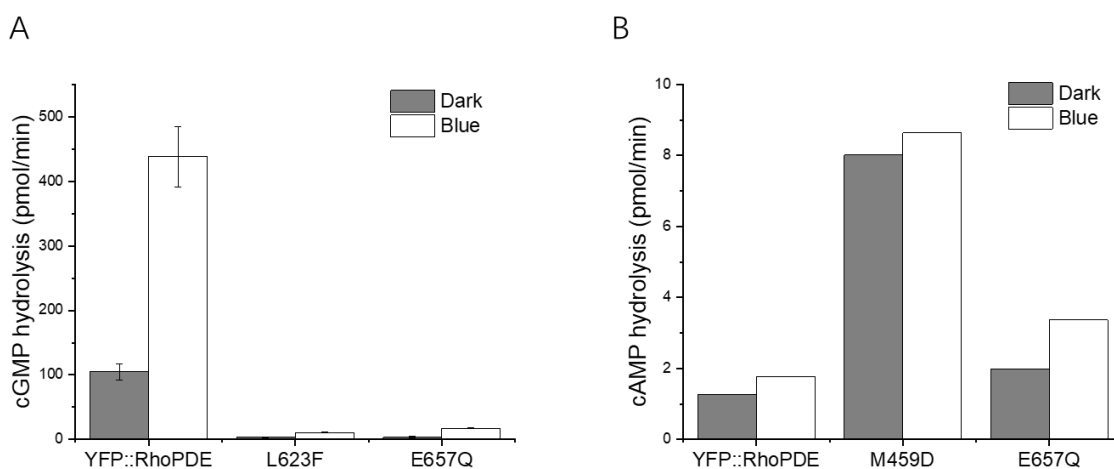
Figure 3.30 Mutants with lower cGMP hydrolysis and higher cAMP hydrolysis.

A, Comparison of cGMP hydrolysis activity among YFP::RhoPDE, L623F and E657Q mutants. Reaction started with 10 μM cGMP. n=3, error bars = SD.

B, Comparison of cAMP hydrolysis activity among YFP::RhoPDE, M459D and E657Q mutants. Reaction started with 20 μM cAMP.

For both A and B, mutants were generated with YFP tag in the N-terminal. Activities came from extracted one oocyte membrane expressing that wt and its mutants under dark and blue light illumination (~0.1 mW/mm^2). They were normalized to the same level of fluorescence emission value.

3.4.2 Different linker lengths between rhodopsin domain and catalytic PDE domain

In previous tests, I did not find an obvious light regulation starting with 1 mM cGMP. Then I tried to modify different linker lengths between opsin domain and PDE domain. Considering the unusual linker of RhoPDE, longer than other enzyme opsins linkers just behind rhodopsin domain, we hypothesized there might be an intron in this long linker region. Therefore, I generated 8 RhoPDE mutants from RhoPDE S1 to S8 with different linker lengths (**Figure 3.31 A**). Then I compared activities of these mutants with wt under dark and green light illumination with 90 μM cGMP at the beginning. The wild type showed slightly ~1.5 fold light activation effect, but RhoPDE S1 mutant with the shortest linker length abolished cGMP hydrolytic activity (**Figure 3.31 B**). RhoPDE S2 revealed the half activity of wt in dark but no light regulation effect. The remaining mutants from S3 to S8 similarly showed the highest cGMP hydrolytic activity but destroyed light regulation. This gives hints that linker modification might obtain a more efficiently light-regulated PDE in the next step.
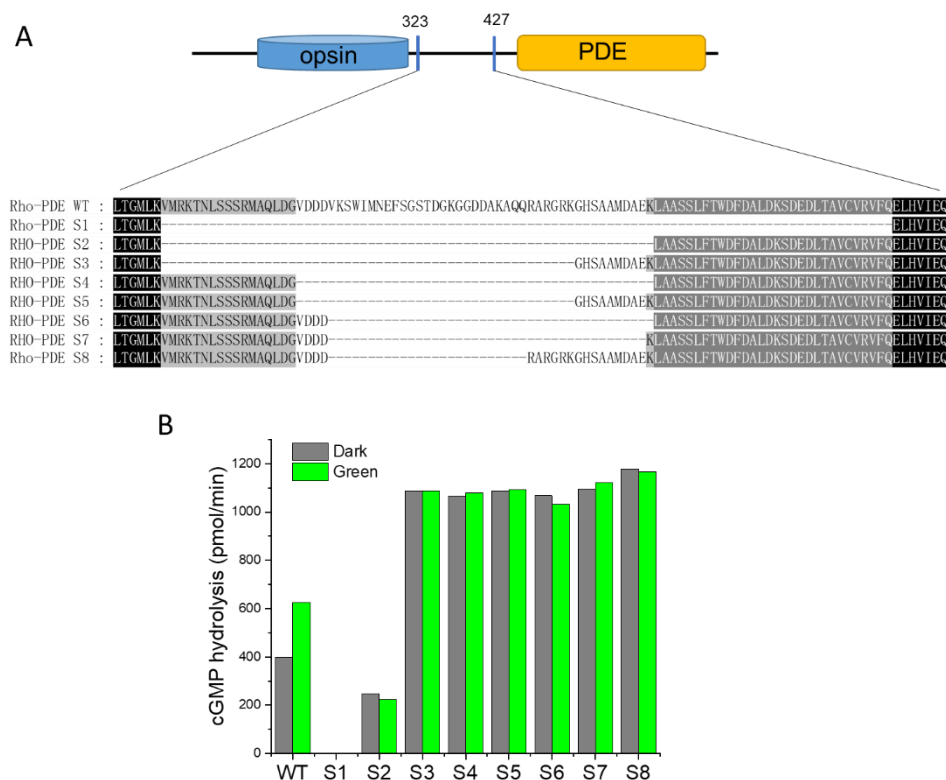
Figure 3.31 Comparison activities of *Sr*RhoPDE with truncated linker lengths.

A, Models and sequences about different truncated constructs in the long linker region.

B, All the mutants were generated from *Sr*RhoPDE wild type by Quickchange PCR. The cGMP hydrolytic

activity was measured under dark and green light illumination (~0.16 mW/mm^2), coming from one oocyte membrane. Reaction started with ~90 μM cGMP.

3.4.3 Fusion of LOV domain or Dronpa to PDE2A

To engineer other light-regulated PDE enzymes is also an important strategy to expand optogenetic tools. *Hs*PDE2A from *Homo sapiens* was applied because it has dual-specificity to hydrolyze both cGMP and cAMP, playing important roles in many important physiological processes. Here I generated a chimera with the blue light-sensor LOV2 domain in the N-terminal, and then PDE2A domain was fused in the middle with YFP tag in the C-terminal. The other construct involved the catalytic domain of PDE2A, smaller fragment named PDE Cat2A, fusion with Dronpa in both terminals (**Figure 3.32 A**). The results revealed that LOV2-PDE2A-YFP construct had a slightly blue light activation effect, possibly because blue light could expose PDE2A domain to increase substrate affinity (**Figure 3.32 B**). More trials in various linker lengths between LOV2 and PDE2A domains could be possible to further increase the light-regulated effect. In addition, based on a novel engineered dimeric fluorescent protein pdDronpa, it can dissociate to monomer in cyan light (500 nm) and reassociate in violet light (400 nm)[108]. Dronpa-PDE Cat2A-Dronpa was generated, but the first trial showed no difference among dark, 410 nm, 505 nm treatment conditions (**Figure 3.32 C**). Both of chimeric constructs were observed the obvious fluorescence emission values from extracted oocytes cytosol (**Figure 3.32 D**).
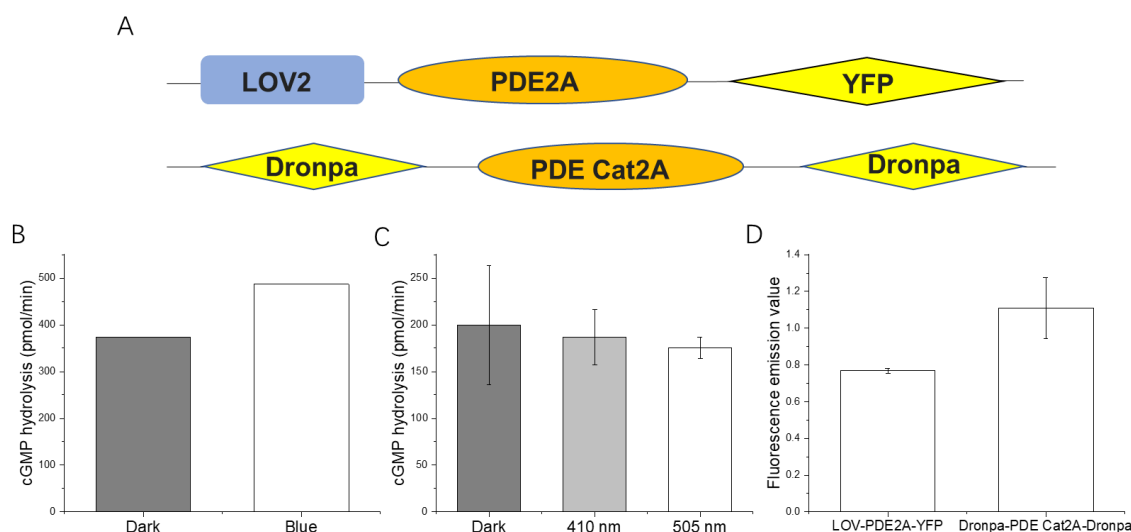
Figure 3.32 Two chimeras were compared under dark and light illumination.

A, Models and strategies to fuse PDE2A with LOV domain or Dronpa domain.

B, cGMP hydrolysis activity of LOV2-PDE2A-YFP in dark and blue light (473 nm, 160 μW/mm^2).

C, cGMP hydrolysis activity of Dronpa-PDE Cat2A-Dronpa in dark, 410 nm (10 μW/mm^2) and 505 nm (80 μW/mm^2) light illumination, n=3, error bar=SD. For B and C, activity was calculated from one oocyte cytosol expressing individual construct. Reactions started with ~82 μM cGMP.

D, Fluorescence emission values in both constructs, each measurement comes from extracted 3 oocytes cytosol after subtraction of control oocytes, n=3, error bar=SD.

3.4.4 Split PDE2A could recover enzyme activity

Theoretically, split fluorescent proteins or luciferase could recover their activities when co-expressing both fragments in living cell systems[102]. Split PDE2A in certain point could also be used as another strategy to engineer a light-regulated enzyme based on the available *Hs*PDE2A structure[109]. The first trial applied two groups of split PDE2A (**Figure 3.33 A**), one briefly named N1 (1-254 aa), C1(255-387 aa) and the other named N2 (1-259 aa), C2 (260-387 aa). It would be confirmed whether any arrangement of two cleaved fragments can recover the PDE enzyme activity, including N1+C1, N1+C2, N2+C1, N2+C2, etc. The linker region was chosen to locate at one of the loops between two α-helices which might reassemble to recover PDE activity after co-expression with a light-insensitive opsin mutant *Cr*2c-Cyclop K298A. After *in vivo* test, N2+C1 co-expression with K298A mutant was likely to reduce cGMP

concentration produced by *Cr*2c-Cyclop K298A mutant, as ~1/5 activity from only K298A expressed group (**Figure 3.33 B**). This gives a hint that N2 and C1 fragments could be fused with other light-induced modules (like Dronpa) to obtain a chimeric PDE enzyme with light regulation effect in the future.
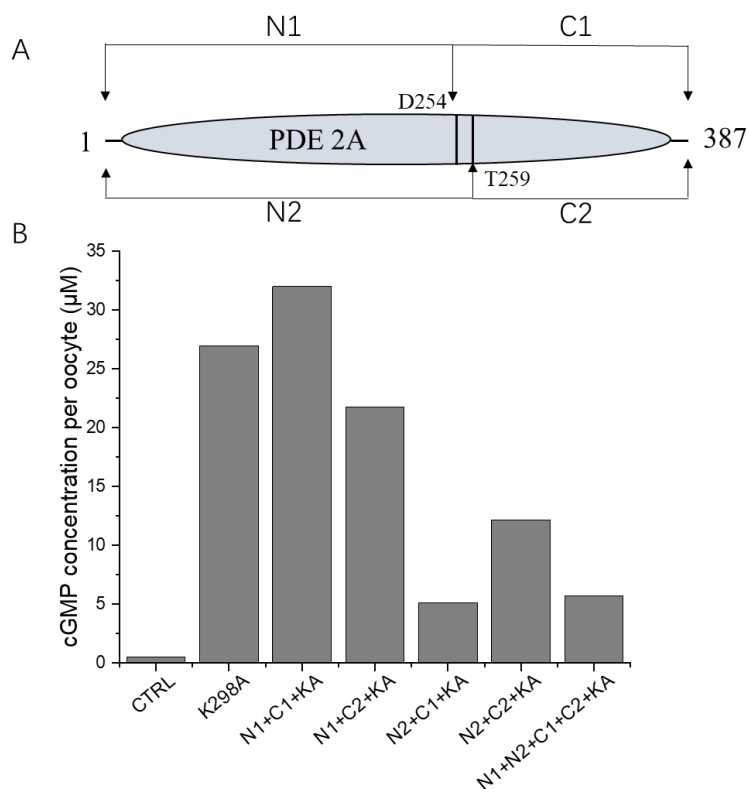
Figure 3.33 Co-expression of *Cr*2c-Cyclop K298A mutant with split PDE2A fragments.

A, Model of split peptide sequences, N1: 1-254 aa, C1: 255-387 aa, N2: 1-259 aa, C2: 260-387 aa.

PDE2A, 1-387 aa comes from C-terminal part of *Hs*PDE2A full-length (555-941aa).

B, cGMP concentration measurement from one oocyte *in vivo* test. For each group, 5 oocytes were homogenized in 500 μL stop solution (0.1 N HCl), 50 μL was taken for measurement. All cGMP concentrations were calculated back to one oocyte.

Expression conditions: K298A: ~5 ng cRNA; N1 (or N2) + C1 (or C2) + K298A: 3:3:1, total ~35 ng cRNA; N1+N2+C1+C2+K298A: 1.5:1.5:1.5:1.5:1, total ~35 ng cRNA, 3dpi.

3.4.5 Chimeras generation with separated rhodopsin and PDE domain

Based on strong light-regulated enzyme opsins, *Be*Cyclop and *Cr*2c-Cyclop, the rhodopsin domain and followed linker region could be fused with other PDE domains to obtain light-induced enzyme. Therefore, three chimeras were constructed as follows: Chimera1 was based on YFP::*Be*Cyclop construct, and the GC domain was replaced with PDE domain from

*Sr*RhoPDE. Chimera2 was based on *Cr*2c-Cyclop construct without YFP tag, and the GC domain was replaced by PDE domain from *Hs*PDE2A. Chimera3 was based on YFP::*Be*Cyclop construct, the GC domain was replaced with PDE domain from *Hs*PDE2A (**Figure 3.34 A**). The results showed that cGMP hydrolytic activity of chimera1 was ~5 times higher than that of the other two chimeras, but none of them has a light regulation effect (**Figure 3.34 B**). Chimera1 and 3 were fused with YFP in the N-terminal, and the fluorescence was determined with no obvious difference (Figure 3.34 C). It might access to light-regulated PDE enzyme based on more modified coiled-coil motif to connect their linker regions.
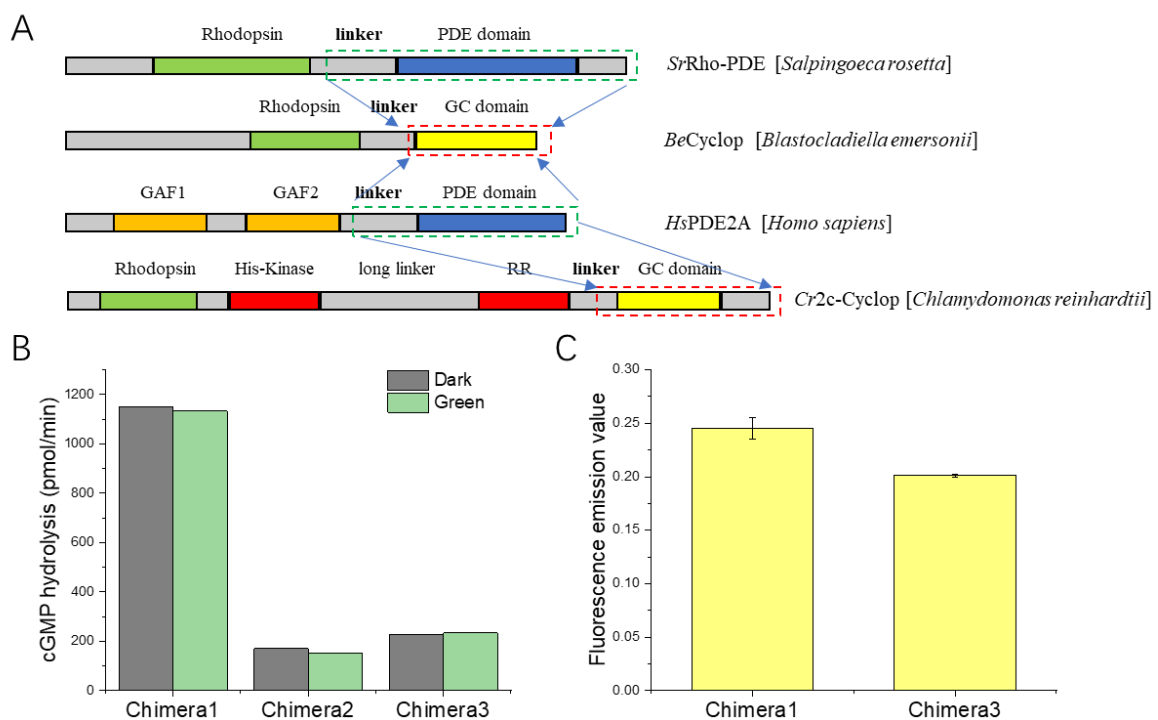
Figure 3.34 Three different chimeras with diverse domains of opsins or PDEs.

A, Models about chimera constructs. Three constructs were used YFP-*Be*Cyclop and *Cr*2c-Cyclop to replace GC domain (red dashed box) with other PDE domain (green dashed box). Arrows represent the replacement between two domains.

B, Comparison of activities under dark and green light illumination (50 μW/mm^2), which was calculated from one extracted oocyte membrane. Reactions started with ~91 μM cGMP.

C, Fluorescence emission values were measured from one oocyte membrane, n=3, error bar=SD.

3.5 Western blot and inclusion body purification

3.5.1 Visualization of *Cr*2c-Cyclop-YFP by western blot

To visualize *Cr*2c-Cyclop-YFP expressing in *Xenopus* oocytes, membrane was extracted from the same batch of oocytes. The control oocytes without cRNA injection and soluble YFP were used as a negative and positive control. *Cr*2c-Cyclop was detected at ~250 kDa by the anti-GFP antibody (**Figure 3.35**). ChR2-YFP was the other positive control recognized by antibody just above 55 kDa. Different treatments of membrane samples were operated. *Cr*2c-Cyclop-YFP showed stronger band with sonication (No.2) than without sonication (No.1). Obviously, targeted membrane proteins were strongly reduced or degraded after boiling treatment at 95 ℃ for 5 min (No. 3 for *Cr*2c-Cyclop-YFP, and No. 6 for ChR2-YFP).
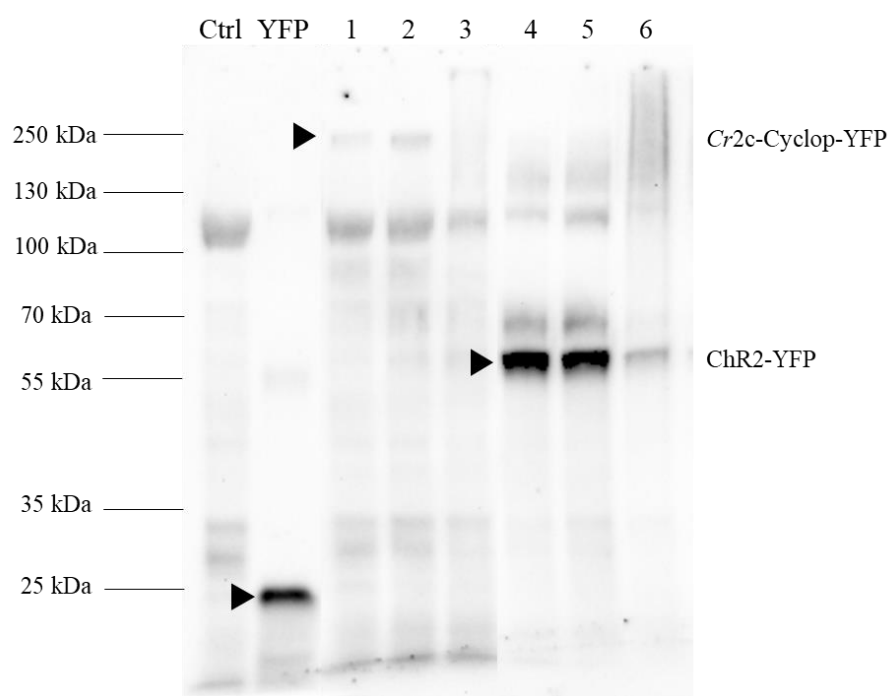
Figure 3.35 Detection of *Cr*2c-Cyclop-YFP by western blot.

Loading samples from left to right lanes: Ctrl, membranes were extracted from 10 control oocytes and loaded; YFP: the standard 40 ng soluble YFP was boiled and loaded.

No. 1, 10 oocytes membrane expressing *Cr*2c-Cyclop-YFP was extracted and incubated at room temperature with additional 3% SDS, 100 mM DTT; No. 2, 10 oocytes were treated as No. 1 but with additional sonication during membrane extraction processes; No. 3, 10 oocytes were treated as No. 1 but with additional boiling at 95 ℃ for 5 min after incubation;

No. 4, 5, 6 samples were treated as No.1, 2, 3 correspondingly, but each loading sample came from 5 oocytes membrane expressing ChR2-YFP.

YFP, *Cr*2c-Cyclop and ChR2-YFP bands are labeled with a black triangle symbol ▶.

3.5.2 Overexpression of histidine kinase fragment of *Cr*2c-Cyclop

To purify inclusion body as an epitope for antibody generation, histidine kinase fragment (251 aa) from *Cr*2c-Cyclop was selected to be overexpressed from *E. coli* BL21 (DE3). After transfection of recombinant PET-28b-*Cr*2c-Cyclop (his-kinase) DNA into *E. coli*, the targeted his-kinase peptide was induced to express with different concentrations of IPTG. The results established that 0.5 mM IPTG (No. 6) had the most efficient induction to overexpress the targeted protein as insoluble fractions compared with 1 mM (No. 7) and 2 mM IPTG (No. 8) induction, see **Figure 3.36**. Moreover, the his-kinase peptide can clearly form inclusion body detected in the pellet (No. 6). While in soluble fractions, the targeted band disappeared (No. 5). Therefore, 0.5 mM IPTG was further used to induce targeted protein overexpression.
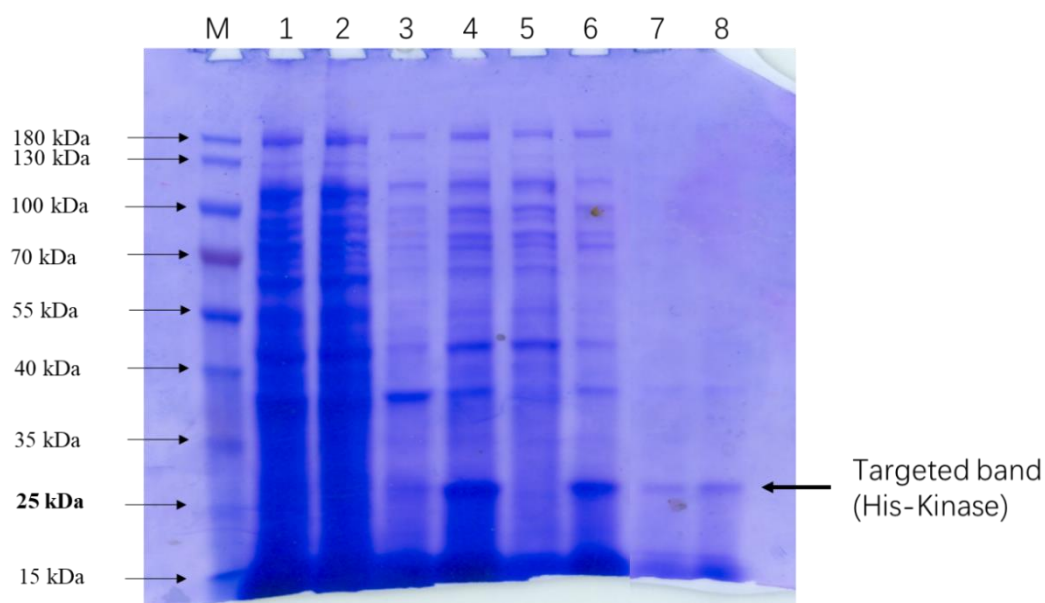
Figure 3.36 Overexpression of *Cr*2c-Cyclop his-kinase fragment.

Samples were loaded from left to right as follows: M, PageRuler™ Prestained Protein Ladder; 1, total cell lysate (-IPTG); 2, soluble fractions (-IPTG); 3, insoluble fractions (-IPTG); 4, total cell lysate (+0.5 mM IPTG induction); 5, soluble fractions (+0.5 mM IPTG induction); 6, insoluble fractions (+0.5 mM IPTG induction); 7, insoluble fractions (+1 mM IPTG induction); 8, insoluble fractions (+2 mM IPTG induction). Loading conditions: total 30 μL (10 μL samples+6 μL 5× SDS loading buffer, 3 μL 10% SDS, 11 μL Tris-buffer) heating at 90℃ for 5min before loading.

3.5.3 Inclusion body purification

After the first trial with a small volume (purification from 15 mL), large volume ~0.5 L recombinant *E. coli* was then applied for inclusion body purification. Strong sonicator (Q500 Sonicator system) was used to replace the previous "freeze and thaw" method to lyse *E. coli* cells more efficiently. It was clearly shown that the targeted his-kinase protein can form an inclusion body with 0.5 mM IPTG induction (**Figure 3.37 A**). However, two strongly contaminated bands appeared at 70 kDa and 15 kDa. To exclude influences from other contaminant proteins, it could be possibly solved by electro-elution of targeted gel slices. Therefore, 0.5 mL insoluble sample extracted from 100 mL transfected *E. coli* was then loaded in an SDS-PAGE gel with one small slot for the marker and one big slot for 0.5-1 mL sample loading (**Figure 3.37 B**). It was clear to see that the targeted strong band located between 25 kDa-35 kDa, which can be cut off as gel slices for further electro-elution processes.
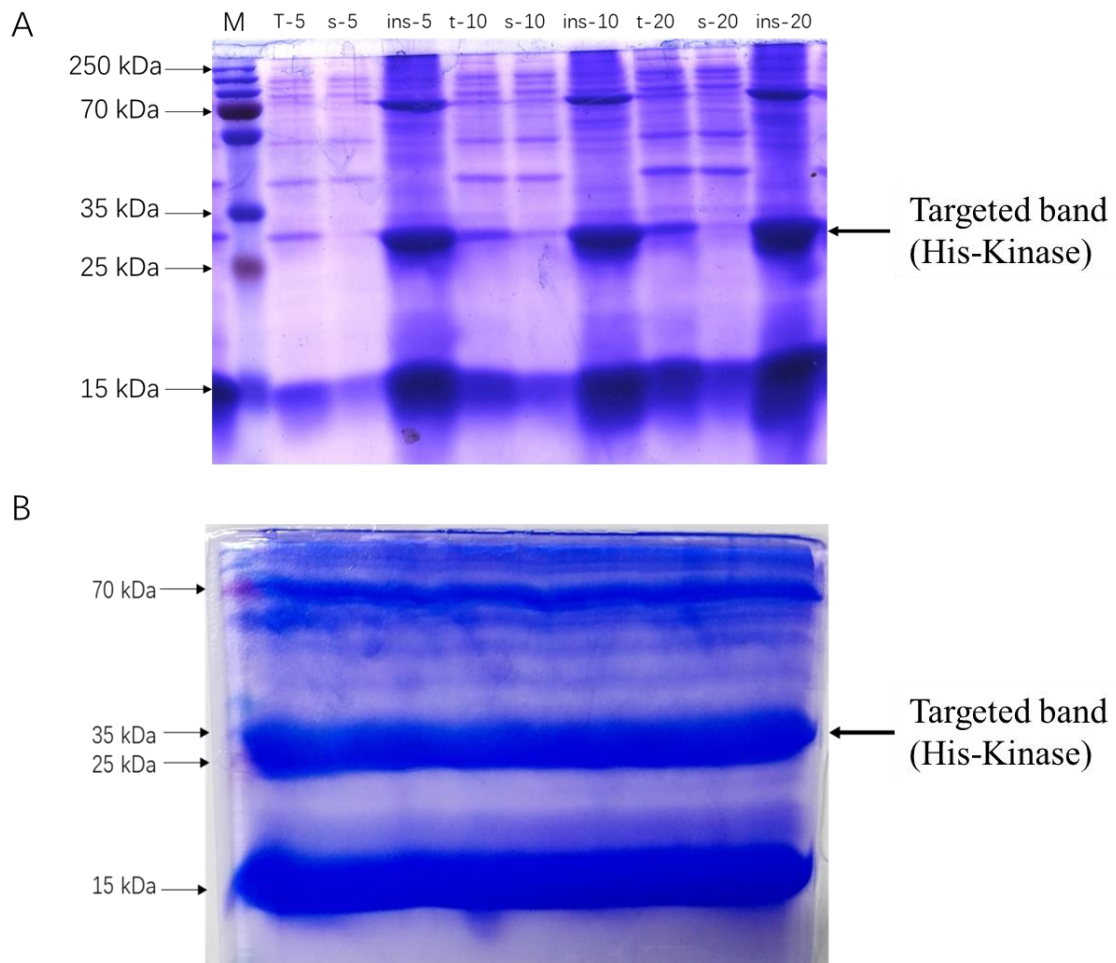
Figure 3.37 Inclusion body purification.

A, Overexpression from 0.5 L recombinant *E. coli* with 0.5 mM IPTG induction. From left to right lanes: M, PageRuler™ Plus Prestained Protein Ladder; t-5, total cell lysate 5 μL loaded; s-5, soluble fractions 5 μL loaded; ins-5, insoluble fractions 5 μL loaded. The remaining number 10, 20 are labeled as 10 μL or 20 μL loaded from above three types of samples.

B, Separation of a large amount of insoluble fractions by the SDS-PAGE gel. 0.5 mL insoluble sample was extracted from 100 mL recombinant *E. coli* and loaded in the big gel slot. The left small slot was loaded with the protein ladder.

4. Discussion

4.1 Mechanisms of a new class of light-inhibited two-component enzyme opsins

My first project was the characterization of two members in a new class of two-component enzyme opsin family. One opsin (abbreviated Cr2c-Cyclop by us) was identified from the genome of a unicellular green alga *C. reinhardtii*[62], and the other one (abbreviated Vc2c-Cyclops by us) was identified from the genome of a multicellular green alga *V. carteri*[110]. Both genome sequencing data are now available in the JGI database[111]. Homologous proteins of these two-component opsins have different loci on different chromosomes of the genomes.

Two-component enzyme opsins share four conserved domains, including rhodopsin domain, histidine kinase, response regulator and guanylyl cyclase domain. Until now, the function or mechanism of the new enzyme opsins was not reported. Here I described two functional opsins, briefly named Cr2c-Cyclop and Vc2c-Cyclop identified from two green algae. Both 2c-Cylops were determined as a new subfamily of type-I rhodopsins with 8-TM helices and light-inhibited enzymatic activity. Obviously, the two enzyme opsins show reverse light-regulation effect compared with the other two classes of enzyme opsins, one is light-activated guanylyl cyclase opsin from fungi (e.g. BeCyclop)[7] and the other is a light-activated phosphodiesterase (RhoPDE) from a protist[8, 91]. However, all these enzyme opsins were confirmed by us to share an 8-TM topology in transmembrane regions. Moreover, they all manipulate cGMP levels via a light-dependent manner. Interestingly, these enzyme opsins exist in organisms which have eyespot organelles and flagellar. It was also suggested that the flagellar motion was depended on light-regulated cGMP signaling cascades in fungus, thus leading to phototaxis as well as sporulation[54, 112].

Regarding the expression system, it is difficult and time-consuming to express and purify such large two-component opsins (normally >200 kDa) or other enzyme opsins in *E. coli* system. This probably cannot obtain native targeted membrane-bound proteins for further biochemical analysis. Therefore, we chose to use *Xenopus* oocytes system to express our targeted enzyme opsins efficiently. After *in vitro* transcription from linearized DNA with a targeted fragment, a

certain amount of cRNA can then be microinjected into oocytes. After 2-3 days of incubation, expression of enzyme opsins can be observed in oocytes when fusing YFP tag to N- or C-terminal of targeted proteins in most cases. The crude membranes can be extracted efficiently for *in vitro* tests with a substrate in the reaction buffer system, as shown in **2.3.3** section. Based on fluorescence emission value and YFP standard curve in the same buffer, we could compare the expression levels of targeted proteins (~pmol range). This would be beneficial to quantify the turnover number of an enzyme as our previous studies[7], especially to compare mutants and wild type, thus uncovering properties and mechanisms of targeted opsins.

Considering the high molecular weight of two-component enzyme opsins (up to 200 kDa), it is difficult to investigate functions and properties of these opsins based on purification. In the last few years, a member of opsin named COP5 (HKR1) was analyzed photon-switchable properties by purification of only transmembrane rhodopsin domain in yeast *Pichia pastoris*. They suggested that HKR1 is more likely to be a UVA-absorbing rhodopsin, which can be switchable between the UV and blue light-absorbing isoform by illumination[64, 65]. Here I used an efficient *Xenopus* oocytes system to express *Cr*2c-Cyclop, and the cGMP production was then measured under dark and light illumination. It was determined by us that *Cr*2c-Cyclop functions as a guanylyl cyclase opsin and shows a green light-inhibition effect. However, there are some problems with characterization of the full-length opsin for its low expression in oocytes and low D/L activity ratio. Therefore, we generated another three different truncated constructs (deletion of either long C-terminal or short middle linker region between his-kinase and response regulator, or both fragments) to increase protein expression levels and enzymatic activity but still keep the dark to light ratio (**Figure 3.3**). Finally, we obtained a construct with deleted long C-terminal and short middle linker region, which showed the highest expression and D/L ratio in comparison with the other three constructs. In the following experiments, I used this truncated construct (briefly named *Cr*2c-Cyclop by us) to characterize its properties such as light sensitivity, action spectrum as well as mutagenesis analysis. This enables us to understand the working mechanisms of the enzyme opsin with a complex signal transduction.

It is necessary to optimize reaction system first, because many elements can influence guanylyl cyclase activity as well as light-regulation, such as ATP, concentrations of the substrate GTP,

and cation ions. With increasing concentrations of GTP as the substrate, the cyclase activity in dark also increased until it arrived at the highest reaction rate. After fitting into the Michaelis-Menten equation, the K_m value for GTP in dark was determined at ~0.13 mM (**Figure 3.4**). Interestingly, *Cr*2c-Cyclop activity was largely depended on ATP, and the dark activity decreased to similar activity in light without ATP supplement. Some cation ions play important roles in regulating cyclase activities. For instance, the structure of Cya2 from *Synechocystis PCC6803* was crystallized as a homologous guanylyl cyclase. This cyclase activity required Mg^{2+}, and further additional Mn^{2+} or Mn^{2+} alone dramatically increased cyclase activity[106]. For *Cr*2c-Cyclop, I used 5 mM Mg^{2+} in standard reaction buffer and demonstrated a clear dark/light ratio at ~10. When 2 mM Mn^{2+} was further added during the reaction, the enzyme opsin increased ~2 times higher cyclase activity in dark than standard buffer only with 5 mM Mg^{2+}, but activity in green light also increased ~4 times than standard buffer. This decreased the D/L ratio (**Figure 3.5**). Moreover, additional 2 mM Ca^{2+} destroyed light-regulated effect and decreased both activities in dark and light, probably by competing with Mg^{2+} binding site. Surprisingly, with 100 mM NaCl in the reaction system, the dark/light activity ratio increased to ~30 compared with D/L ratio at ~10 in less than 10 mM NaCl solution in previous tests. The following characterization processes were efficiently operated based on above-optimized reaction system.

Comparing with light sensitivity of other enzyme opsins, *Cr*2c-Cyclop showed the most sensitive to light with $K_{0.5}$ value at ~0.2 $\mu W/mm^2$ at both 532 nm and 596 nm light. However, some other photo-activated enzymes showed less sensitivity near the peak of their own action spectra, such as BeCyclop[7] ($K_{0.5}$=55 $\mu W/mm^2$), bPAC[53] ($K_{0.5}$=4 $\mu W/mm^2$), and mPAC[107] ($K_{0.5}$=6 $\mu W/mm^2$). Although *Vc*2c-Cyclop showed less sensitive to light than *Cr*2c-Cyclop, light intensity on the half-maximal inhibition was still ~1.3 $\mu W/mm^2$ at a wavelength of 556 nm light around its action spectrum peak, more sensitive than other enzyme opsins (**Figure 3.10**). Interestingly, after application of different wavelengths, it was demonstrated that action spectrum of *Cr*2c-Cyclop is peaking at ~541 nm while *Vc*2c-Cyclop shows a slight red shift, peaking at ~556 nm (**Figure 3.11**).

To unveil the mechanisms of *Cr*2c-Cyclop, *Cr*2c-Cyclop construct fusion with YFP was proven to be used for comparison in different conditions or mutagenesis analysis. All-*trans*-retinal (ATR), as an essential chromophore, is a necessity of microbial opsins to induce light regulation effects. Without supplement ATR, light-gated ChR2 showed weaker expression level thus decreasing photocurrent[4, 113]. Moreover, BeCyclop also decreased its light activation effect[7]. In my study, I confirmed that the expression of *Cr*2c-Cyclop::YFP showed no obvious difference with or without ATR, but the D/L activity ratio changed from ~30 with ATR to ~5 without ATR (**Figure 3.9 A**). We suggested that retinal-free *Cr*2c-Cyclop tends to decrease light inhibition effects without ATR supplement. More interestingly, one light-insensitive *Cr*2c-Cyclop::YFP K298A mutant was generated which cannot covalently bind to retinal by a Schiff base. This tends to keep the conformation in dark state because light cannot lead to a conformational change of rhodopsin domain without ATR bound. It was clearly demonstrated that the K298A mutant recovered the light activity similar with dark activity and no light-regulation was observed.

Two components (most cases as separated proteins in bacteria), comprising histidine kinase and response regulator, play essential roles in sensing many environmental stimuli and triggering many responses. These processes mainly depend on phosphoryl transfer from histidine kinases to response regulators[68]. It was clear that the dark activity of *Cr*2c-Cyclop dramatically decreased without supplement ATP from my experiment. This suggested phosphorylation processes were interrupted without ATP as phosphoryl group donor and then the output cyclase was inhibited. We further confirmed this mechanism by analyzing three important single mutants, including H352 in the H-box motif of the DHp domain for autophosphorylation, G533 in the G-box motif of the kinase for ATP binding, and D1092 as phosphoryl acceptor in response regulator. It was determined that all three key single mutants H352F, G533A and D1092T destroyed phosphorylation processes, which strongly reduced cyclase activity even lower than the activity of wild type under light illumination. Therefore, we anticipate that the phosphorylation processes can be continuous between two components to activate cyclase activity under dark, while the light-induced conformational change of rhodopsin tends to interrupt phosphoryl transfer and strongly inhibit cyclase activity. Hopefully, the detailed

working mechanisms will be solved by structure determination in the future.

In the kingdom of green algae, abundant photoreceptors are existing as light-gated channelrhodopsins, cryptochromes and enzyme opsins[114]. Apart from Channelrhodopsin-1 and 2 identified from *C. reinhardtii*, other channelrhodopsins were identified from different species of algae with different properties, which can be screened to obtain potential optogenetic tools based on special characters like photocurrent, light sensitivity, action spectrum and on-off kinetics[33]. Interestingly, many conserved two-component cyclase opsins (2c-Cyclops) exist not only in *C. reinhardtii* and *V. carteri* but also in green algae maintained during evolutionary processes (from *C. reinhardtii* to *V. carteri*)[115]. For example, few 2c-Cyclops are still existing in available genome databases of green algae *Gonium pectoral* and *Pleodorina starrii*[114]. Interestingly, I also demonstrated that the full-length of *Vc*2c-Cyclop from *V. carteri* also showed light-inhibited guanylyl cyclase activity, similar with *Cr*2c-Cyclop. This may give hints that 2c-Cyclops are conserved during evolutionary processes and keep their function with light-regulated cGMP (or cAMP not found yet) signaling cascades. To investigate the function of 2c-Cyclops in their natural host, it is necessary to require more knowledge of temporal and spatial expression, as well as efficient gene knockout methods. On the other hand, intracellular cAMP/cGMP is crucial to regulating phosphorylation of protein kinases A/G or activating transcription factor CREB (cAMP/cGMP response element binding), playing roles in many developmental processes of the multicellular organisms. Light-regulated guanylyl cyclase or light-stimulated phosphodiesterase could also be further engineered and applied to manipulate cAMP/cGMP levels *in vivo* as optogenetic tools.

4.2 Determination of a novel light-activated rhodopsin phosphodiesterase (RhoPDE)

My second project was the characterization of a novel rhodopsin phosphodiesterase (*Sr*RhoPDE) from *Salpingoeca rosetta*. Our results clearly demonstrated that this enzyme opsin belongs to a light-activated PDE, with an 8-TM topology in the N-terminal rhodopsin domain. In a low range of substrate concentrations (e.g. starting at ~10 μM for cGMP, ~100 μM for cAMP), illumination can cause 4-5 fold light activation effects to hydrolyze corresponding cGMP or cAMP. Then I generated two light-insensitive mutants (K296M and K296A), which cannot

covalently bind to all-*trans*-retinal (ATR) by a Schiff base. This can interrupt signal transmission from ATR isomerization to conformational change of rhodopsin, thus destroying light-regulated effects. It was determined that both light-insensitive mutants still remained strong PDE activity but no light regulation excluding artifactual light activation of RhoPDE by illumination conditions. Interestingly, the activity of K296A mutant is more close to wild type (wt) in light while K296M mutant is more similar with wt in dark. Yoshida et al. showed that RhoPDE had a 1.4-fold cGMP hydrolytic activity between light and dark treatments when starting the reaction at 100 μM cGMP[91]. This low light-activation was determined at pH 6.5. However, we determined a ratio (L/D) of ~2 at pH 7.3 when starting at 100 μM cGMP. This was further confirmed by us that different pH values influenced the RhoPDE activity, so as the initial substrate concentrations (**Figure 3.28**). Apart from reaction conditions, different expression systems still might have effects on L/D ratio of light-regulated enzymes. For example, in previous tests, we observed that the light-activated guanylyl cyclase BeCyclop from *B. emersonii* showed lower L/D ratio expressing in HEK293 cells than L/D ratio expressing in *Xenopus* oocytes[7]. Comparing with HEK293 cells, *Xenopus* oocytes can be homogenized faster by simply breaking oocytes with small-opening pipette tips during membrane extraction processes. This could cause less damage to membrane proteins and keep proteins in a more native state.

In addition, retinal binding to opsins is the prerequisite for light regulation. It was clearly shown that a lacking Schiff base bond to retinal yielded completely light-insensitive proteins as those two single K296A/M mutants. In fact, we suggest that retinal is saturate embedding in all wild type RhoPDE proteins after cRNA injection with additional 1 μM ATR in ND96 buffer. This retinal supplementation is able to enhance better expression for ChR2[113] and further increase photocurrents of most rhodopsins in our previous tests[3, 4, 116]. However, if retinal is not saturated to bind RhoPDE, the opsin-PDE without retinal would skew the measured K_m value for cGMP in light. Lamarche et al.[92] determined that the K_{cat} value was ranging from 19 to 28 s^{-1}, similar with a turnover in light determined by us (28 ± 5 s^{-1} at 100 μM cGMP is close to a maximal turnover of 32 ± 6 s^{-1}). This also suggested that RhoPDE is a robust phosphodiesterase in both dark or light conditions.

Interestingly, I confirmed that light illumination can enhance substrate affinity, and the K_m value changed from 80 μM in dark to 13 μM in light. But the maximal hydrolysis turnover was influenced slightly (increasing only by 30% in light). This can explain why Lamarche et al. did not observe light activation when using completely saturate 5 mM cGMP as a substrate[92]. In fact, I also did not yield a convincing conclusion about light activation when using 1 mM cGMP as the substrate in previous tests.

Moreover, I determined that the light-insensitive K296A mutant showed a K_m value for cGMP at 13 μM, the same value as RhoPDE wt in light, whereas K296M mutant was measured with K_m value for cGMP at 63 μM, close to K_m value of wt in dark condition. We anticipate that light stimulation tends to cause conformational changes of rhodopsin which further transduce to move PDE catalytic domain, thus leading to a higher affinity for cGMP. Hopefully, this question will be solved by structure determination in the future. It might be helpful to elucidate the working mechanisms by investigating those two mutants with quite different K_m values.

It is suggested that rhodopsin domain of RhoPDE includes 8-TM with both termini in the cytosolic side based on our BiFC (bimolecular fluorescence complementation) experiment as well as immunofluorescence result from Lamarche et al.[92]. Another enzyme rhodopsin BeCyclop was also demonstrated with an 8-TM topology first by us[7] and then confirmed by Trieu et al.[57]. We demonstrated that N-terminal YFP fused RhoPDE showed over 2 times higher expression level than C-terminal YFP fused construct. YFP::RhoPDE still remained the light/dark ratio as wt while RhoPDE::YFP strongly reduced light/dark ratio by increasing dark activity.

Until now, two classes of well-characterized enzyme rhodopsins were determined as 8-TM topology structure, such as Cyclops and RhoPDE. BeCyclop (BeGC1) was located in the external surface of the zoospores of *B. emersonii*, playing roles in fungal phototaxis[54]. This was then demonstrated that BeCyclop regulated zoospore motility by cGMP-activated K^+ channel probably participating in phototactic responses of the fungus[112]. *S. rosetta* belongs to choanoflagellates, which are considered as the closest living relatives of animals. We currently cannot explain the function of *Sr*RhoPDE in its natural host because *S. rosetta* is not studied in depth. However, *Sr*RhoPDE makes it possible to manipulate cGMP/cAMP as a potential

optogenetic tool. BeCyclop, a light-activated guanylyl-cyclase opsin, has already been applied as a powerful optogenetic tool to increase cGMP levels by illumination[7, 55]. Currently, it is unclear that whether some unknown factors tightly regulate *Sr*RhoPDE depending on light illumination in *S. rosetta*. In the future, it might be possible to improve the L/D activity ratio of RhoPDE, either by further studies in *S. rosetta* or by genetic engineering.

4.3 Combination or engineering of new light-regulated enzymes

Until now, optogenetic tools were mostly discovered from photoreceptors in many microbial organisms and most of them function as single protein scaffolds (e.g. light-gated channels or pumps, light-regulated enzymes), which are applied mainly in neuroscientific fields. In addition, optogenetic tools are expanded to regulate light-induced transcription by using two or more light-stimulated dimerization or oligomerization components. This could allow optogenetic tools to regulate a certain signaling cascade in subcellular level. Obviously, functional photoreceptors can be combined to enhance light-regulated efficiency, and they can also be applied with other components such as CNG (cyclic nuclear gated) channels to regulate membrane potentials. For example, in our previous study, a combination of BeCyclop and *C. elegans* CNG channel (with TAX-2 and TAX-4 subunits) can depolarize muscle cells by light-induced cGMP stimulation[7]. In my study, I characterized two cGMP related enzyme opsins, one light-inhibited 2c-Cyclop and the other light-activated phosphodiesterase (RhoPDE). Actually, RhoPDE has strong dark and light activity for cGMP hydrolysis. It would be more helpful to decrease hydrolytic activity but still keep light-regulated influences. Interestingly, I obtained two single mutants (L623F and E657Q) located in PDE catalytic domain, which decreased to ~1/40 activity of RhoPDE wild type (**Figure 3.30**). It could also be possible to combine *Cr*2c-Cyclop with RhoPDE L623F or E657Q mutant based on their cGMP production or hydrolysis activity under dark and light from one oocyte membrane. The co-expression system makes it possible to accumulate cGMP in dark due to *Cr*2c-Cyclop and hydrolyze cGMP in light due to RhoPDE L623F or E657Q mutant, thus generating a more efficient light-switchable enzyme. This tends to mimic type-II rhodopsins in vision system of vertebrates, which also show light-induced cGMP hydrolysis and dark-accumulated cGMP levels.

Therefore, the combination of *Cr*2c-Cyclop and RhoPDE L623F (or E657Q) mutant might be applied in a vision system to uncover more mechanisms about visual transmission processes in the future.

To engineer light-induced enzymes, an important question is about how to transmit signals from conformational changes of light-regulated modules to catalytic enzymes. The linker regions between light sensors and output effectors play key roles in signaling transmission. In some photoreceptors, it was demonstrated that coiled-coil linkers have effects on light regulation. For instance, as light-oxygen-voltage (LOV) photosensor domain was fused with histidine kinase domain by a modified Jα linker region, this chimeric his-kinase can be regulated efficiently (D/L kinase activity ~1000-fold *in vitro*) by light illumination[81]. Here I described a chimera which fused LOV domain with *Hs*PDE2A catalytic domain, and it showed slightly light activation (**Figure 3.32 B**). But it could be necessary to do more trials in the Jα linker region to enhance light-regulated efficiency. Interestingly, the N-terminal short α-helix of LOV can assemble into the coiled-coil region at the LOV-domain interface, but certain single mutants in the short α-helix decreased or inverted light regulation effects[82]. This site-directed mutant could also be used to generate two inverse light-induced enzymes by fusion LOV with other potential enzymes with similar linker characters.

Regarding light-associated or light-dissociated optogenetic tools, one typical example is Dronpa, which was engineered as a mutant GFP-like fluorescent protein, showing light-switchable properties[117]. This was further applied to generate single-chain photoswitchable kinases[108], and the newly engineered pdDronpa (photodissociable dimeric Dronpa) insertion could also be applied to generate other photo-stimulated enzymes. In my study, I tried to fuse pdDronpa in both N- and C- termini of PDE2A catalytic domain, but the light regulation was not observed under violet and cyan light illumination in the first trial (**Figure 3.32 C**). More trials could be possible to obtain a light-controlled PDE2A. Another possibility is fusing pdDronpa in split PDE2A. It was suggested that two split PDE2A fragments, especially N2+C1, could recover PDE activity (**Figure 3.33**). This makes it possible to obtain a light-induced enzyme activity by fusing pdDronpa (or LOV) domain with split PDE2A fragments respectively. In addition, as a strong green light-activated cyclase opsin (BeCyclop), the N-terminal

rhodopsin domain with the linker region (with predicted coiled-coil motif) could also be connected to the N-terminal of PDE catalytic domain either from *Hs*PDE2A or *Sr*RhoPDE with similar linker motifs, even though the light-regulated effect was not observed in the first trial (**Figure 3.34**). This might be solved by more knowledge of understanding how to efficiently design chimeric proteins in the future.

5. References

1.   Crosson, S., S. Rajagopal, and K. Moffat, *The LOV domain family: photoresponsive signaling modules coupled to diverse output domains.* Biochemistry, 2003. **42**(1): p. 2-10.

2.   Klar, T., et al., *Cryptochrome 3 from Arabidopsis thaliana: structural and functional analysis of its complex with a folate light antenna.* J Mol Biol, 2007. **366**(3): p. 954-64.

3.   Nagel, G., et al., *Channelrhodopsin-1: a light-gated proton channel in green algae.* Science, 2002. **296**(5577): p. 2395-8.

4.   Nagel, G., et al., *Channelrhodopsin-2, a directly light-gated cation-selective membrane channel.* Proc Natl Acad Sci U S A, 2003. **100**(24): p. 13940-5.

5.   Moglich, A., et al., *Structure and function of plant photoreceptors.* Annu Rev Plant Biol, 2010. **61**: p. 21-47.

6.   Spudich, J.L., et al., *Retinylidene proteins: structures and functions from archaea to humans.* Annu Rev Cell Dev Biol, 2000. **16**: p. 365-92.

7.   Gao, S., et al., *Optogenetic manipulation of cGMP in cells and animals by the tightly light-regulated guanylyl-cyclase opsin CyclOp.* Nature communications, 2015. **6**.

8.   Tian, Y.H., et al., *A novel rhodopsin phosphodiesterase from Salpingoeca rosetta shows light-enhanced substrate affinity.* Biochemical Journal, 2018. **475**: p. 1121-1128.

9.   Beavo, J.A. and L.L. Brunton, *Cyclic nucleotide research - still expanding after half a century.* Nature Reviews Molecular Cell Biology, 2002. **3**(9): p. 710-718.

10.   Spudich, J.L., *The multitalented microbial sensory rhodopsins.* Trends Microbiol, 2006. **14**(11): p. 480-7.

11.   Shichida, Y. and T. Matsuyama, *Evolution of opsins and phototransduction.* Philos Trans R Soc Lond B Biol Sci, 2009. **364**(1531): p. 2881-95.

12.   Ernst, O.P., et al., *Microbial and animal rhodopsins: structures, functions, and molecular mechanisms.* Chem Rev, 2014. **114**(1): p. 126-63.

13.   Schmidt, T.M., S.K. Chen, and S. Hattar, *Intrinsically photosensitive retinal ganglion cells: many subtypes, diverse functions.* Trends Neurosci, 2011. **34**(11): p. 572-80.

14.   Luecke, H., et al., *Structure of bacteriorhodopsin at 1.55 angstrom resolution.* Journal of Molecular Biology, 1999. **291**(4): p. 899-911.

15.   Palczewski, K., et al., *Crystal structure of rhodopsin: A G protein-coupled receptor.* Science, 2000. **289**(5480): p. 739-45.

16.   Man, D., et al., *Diversification and spectral tuning in marine proteorhodopsins.* EMBO J, 2003. **22**(8): p. 1725-31.

17.   Zhang, F., et al., *The microbial opsin family of optogenetic tools.* Cell, 2011. **147**(7): p. 1446-57.

18.   Shcherbakova, D.M., et al., *Natural Photoreceptors as a Source of Fluorescent Proteins, Biosensors, and Optogenetic Tools.* Annual Review of Biochemistry, Vol 84, 2015. **84**: p. 519-550.

19.   Yizhar, O., et al., *Microbial opsins: a family of single-component tools for optical control of neural activity.* Cold Spring Harb Protoc, 2011. **2011**(3): p. top102.

20.   Palczewski, K., *Chemistry and biology of vision.* J Biol Chem, 2012. **287**(3): p. 1612-9.

21.   Veleri, S., et al., *Biology and therapy of inherited retinal degenerative disease: insights from mouse models.* Dis Model Mech, 2015. **8**(2): p. 109-29.

22.   Govorunova, E.G., et al., *Microbial Rhodopsins: Diversity, Mechanisms, and Optogenetic*

Applications. Annu Rev Biochem, 2017. **86**: p. 845-872.

23. Oesterhelt, D. and W. Stoeckenius, *Rhodopsin-like protein from the purple membrane of Halobacterium halobium.* Nat New Biol, 1971. **233**(39): p. 149-52.

24. Oesterhelt, D. and W. Stoeckenius, *Functions of a new photoreceptor membrane.* Proc Natl Acad Sci U S A, 1973. **70**(10): p. 2853-7.

25. Kayushin Lp Fau - Skulachev, V.P. and V.P. Skulachev, *Bacteriorhodopsin as an electrogenic proton pump: reconstitution of bacteriorhodopsin proteoliposomes generating delta psi and delta pH.* 1974(0014-5793 (Print)).

26. Schobert, B. and J.K. Lanyi, *Halorhodopsin Is a Light-Driven Chloride Pump.* Journal of Biological Chemistry, 1982. **257**(17): p. 306-313.

27. Duschl, A., J.K. Lanyi, and L. Zimanyi, *Properties and photochemistry of a halorhodopsin from the haloalkalophile, Natronobacterium pharaonis.* J Biol Chem, 1990. **265**(3): p. 1261-7.

28. Varo, G., *Analogies between halorhodopsin and bacteriorhodopsin.* Biochimica Et Biophysica Acta-Bioenergetics, 2000. **1460**(1): p. 220-229.

29. Dawydow, A., et al., *Channelrhodopsin-2-XXL, a powerful optogenetic tool for low-light applications.* Proc Natl Acad Sci U S A, 2014. **111**(38): p. 13972-7.

30. Nagel, G., et al., *Light activation of channelrhodopsin-2 in excitable cells of Caenorhabditis elegans triggers rapid Behavioral responses.* Current Biology, 2005. **15**(24): p. 2279-2284.

31. Boyden, E.S., et al., *Millisecond-timescale, genetically targeted optical control of neural activity.* Nat Neurosci, 2005. **8**(9): p. 1263-8.

32. Han, X., *In vivo application of optogenetics for neural circuit analysis.* ACS Chem Neurosci, 2012. **3**(8): p. 577-84.

33. Klapoetke, N.C., et al., *Independent optical excitation of distinct neural populations.* Nat Methods, 2014. **11**(3): p. 338-46.

34. Govorunova, E.G., et al., *NEUROSCIENCE. Natural light-gated anion channels: A family of microbial rhodopsins for advanced optogenetics.* Science, 2015. **349**(6248): p. 647-50.

35. Pan, Z.H., et al., *Optogenetic Approaches to Restoring Vision.* Annu Rev Vis Sci, 2015. **1**: p. 185-210.

36. Gradinaru, V., et al., *Optical deconstruction of parkinsonian neural circuitry.* Science, 2009. **324**(5925): p. 354-9.

37. Shimano, T., et al., *Assessment of the AAV-mediated expression of channelrhodopsin-2 and halorhodopsin in brainstem neurons mediating auditory signaling.* Brain Res, 2013. **1511**: p. 138-52.

38. Vuong, T.M., M. Chabre, and L. Stryer, *Millisecond activation of transducin in the cyclic nucleotide cascade of vision.* Nature, 1984. **311**(5987): p. 659-61.

39. Hardman, J.G., *Cyclic nucleotides and regulation of vascular smooth muscle.* J Cardiovasc Pharmacol, 1984. **6 Suppl 4**: p. S639-45.

40. Dodge-Kafka, K.L., L. Langeberg, and J.D. Scott, *Compartmentation of cyclic nucleotide signaling in the heart: the role of A-kinase anchoring proteins.* Circ Res, 2006. **98**(8): p. 993-1001.

41. Matsumoto, Y., et al., *Cyclic nucleotide-gated channels, calmodulin, adenylyl cyclase, and calcium/calmodulin-dependent protein kinase II are required for late, but not early, long-term memory formation in the honeybee.* Learn Mem, 2014. **21**(5): p. 272-86.

42. Punta, M., et al., *The Pfam protein families database.* Nucleic Acids Res, 2012. **40**(Database issue): p. D290-301.

43. Sunahara, R.K., et al., *Exchange of substrate and inhibitor specificities between adenylyl and guanylyl cyclases.* Journal of Biological Chemistry, 1998. **273**(26): p. 16332-16338.

44. Gomelsky, M. and M.Y. Galperin, *Bacterial second messengers, cGMP and c-di-GMP, in a quest for regulatory dominance.* EMBO J, 2013. **32**(18): p. 2421-3.

45. Marden, J.N., et al., *Cyclic GMP controls Rhodospirillum centenum cyst development.* Mol Microbiol, 2011. **79**(3): p. 600-15.

46. An, S.Q., et al., *A cyclic GMP-dependent signalling pathway regulates bacterial phytopathogenesis.* EMBO J, 2013. **32**(18): p. 2430-8.

47. Romling, U., M.Y. Galperin, and M. Gomelsky, *Cyclic di-GMP: the first 25 years of a universal bacterial second messenger.* Microbiol Mol Biol Rev, 2013. **77**(1): p. 1-52.

48. Walseth, T.F., et al., *The Fate of O-18 in Guanosine-Monophosphate during Enzymic Transformations Leading to Guanosine 3',5'-Monophosphate Generation.* Journal of Biological Chemistry, 1981. **256**(5): p. 2176-2179.

49. Goldberg, N.D., et al., *Cyclic-Amp Metabolism in Intact Platelets Determined by O-18 Incorporation into Adenine-Nucleotide Alpha-Phosphoryls.* Advances in Cyclic Nucleotide and Protein Phosphorylation Research, 1984. **16**: p. 363-379.

50. Iseki, M., et al., *A blue-light-activated adenylyl cyclase mediates photoavoidance in Euglena gracilis.* Nature, 2002. **415**(6875): p. 1047-51.

51. Schroder-Lang, S., et al., *Fast manipulation of cellular cAMP level by light in vivo.* Nat Methods, 2007. **4**(1): p. 39-42.

52. Ryu, M.-H., et al., *Natural and engineered photoactivated nucleotidyl cyclases for optogenetic applications.* Journal of Biological Chemistry, 2010. **285**(53): p. 41501-41508.

53. Stierl, M., et al., *Light modulation of cellular cAMP by a small bacterial photoactivated adenylyl cyclase, bPAC, of the soil bacterium Beggiatoa.* Journal of Biological Chemistry, 2011. **286**(2): p. 1181-1188.

54. Avelar, G.M., et al., *A rhodopsin-guanylyl cyclase gene fusion functions in visual perception in a fungus.* Curr Biol, 2014. **24**(11): p. 1234-40.

55. Scheib, U., et al., *The rhodopsin-guanylyl cyclase of the aquatic fungus Blastocladiella emersonii enables fast optical control of cGMP signaling.* Sci Signal, 2015. **8**(389): p. rs8.

56. Kumar, R.P., et al., *Structure and monomer/dimer equilibrium for the guanylyl cyclase domain of the optogenetics protein RhoGC.* J Biol Chem, 2017. **292**(52): p. 21578-21589.

57. Trieu, M.M., et al., *Expression, purification, and spectral tuning of RhoGC, a retinylidene/guanylyl cyclase fusion protein and optogenetics tool from the aquatic fungus Blastocladiella emersonii.* Journal of Biological Chemistry, 2017. **292**(25): p. 10379-10389.

58. Hegemann, P., *Algal sensory photoreceptors.* Annu. Rev. Plant Biol., 2008. **59**: p. 167-189.

59. Kateriya, S., et al., *"Vision" in Single-Celled Algae.* Physiology, 2004. **19**(3): p. 133-137.

60. Deininger, W., et al., *Chlamyrhodopsin represents a new type of sensory photoreceptor.* EMBO J, 1995. **14**(23): p. 5849-58.

61. Fuhrmann, M., et al., *The abundant retinal protein of the Chlamydomonas eye is not the photoreceptor for phototaxis and photophobic responses.* 2001(0021-9533 (Print)).

62. Merchant, S.S., et al., *The Chlamydomonas genome reveals the evolution of key animal and plant functions.* Science, 2007. **318**(5848): p. 245-250.

63. Greiner, A., et al., *Targeting of Photoreceptor Genes in Chlamydomonas reinhardtii via Zinc-finger Nucleases and CRISPR/Cas9.* The Plant Cell, 2017.

64. Luck, M., et al., *A photochromic histidine kinase rhodopsin (HKR1) that is bimodally switched by ultraviolet and blue light.* Journal of Biological Chemistry, 2012. **287**(47): p. 40083-40090.

65. Penzkofer, A., et al., *Bistable retinal schiff base photodynamics of histidine kinase rhodopsin HKR1 from Chlamydomonas reinhardtii.* Photochem Photobiol, 2014. **90**(4): p. 773-85.

66. Luck, M., et al., *Photochemical chromophore isomerization in histidine kinase rhodopsin HKR1.* FEBS Lett, 2015. **589**(10): p. 1067-71.

67. Luck, M. and P. Hegemann, *The two parallel photocycles of the Chlamydomonas sensory photoreceptor histidine kinase rhodopsin 1.* J Plant Physiol, 2017. **217**: p. 77-84.

68. Zschiedrich, C.P., V. Keidel, and H. Szurmant, *Molecular Mechanisms of Two-Component Signal Transduction.* Journal of Molecular Biology, 2016. **428**(19): p. 3752-3775.

69. Kim, D. and S. Forst, *Genomic analysis of the histidine kinase family in bacteria and archaea.* Microbiology, 2001. **147**(Pt 5): p. 1197-212.

70. Wuichet, K., B.J. Cantwell, and I.B. Zhulin, *Evolution and phyletic distribution of two-component signal transduction systems.* Curr Opin Microbiol, 2010. **13**(2): p. 219-25.

71. Capra, E.J. and M.T. Laub, *Evolution of two-component signal transduction systems.* Annu Rev Microbiol, 2012. **66**: p. 325-47.

72. Alvarez, A.F., et al., *Organization and mode of action of two component system signaling circuits from the various kingdoms of life.* Environmental Microbiology, 2016. **18**(10): p. 3210-3226.

73. Yamamoto, K., et al., *Functional characterization in vitro of all two-component signal transduction systems from Escherichia coli.* Journal of Biological Chemistry, 2005. **280**(2): p. 1448-1456.

74. Gerken, H., et al., *MzrA: a novel modulator of the EnvZ/OmpR two-component regulon.* Molecular microbiology, 2009. **72**(6): p. 1408-1422.

75. Casino, P., V. Rubio, and A. Marina, *Structural insight into partner specificity and phosphoryl transfer in two-component signal transduction.* Cell, 2009. **139**(2): p. 325-336.

76. Willett, J.W. and J.R. Kirby, *Genetic and biochemical dissection of a HisKA domain identifies residues required exclusively for kinase and phosphatase activities.* PLoS Genet, 2012. **8**(11): p. e1003084.

77. Liu, Y., et al., *A pH-gated conformational switch regulates the phosphatase activity of bifunctional HisKA-family histidine kinases.* Nat Commun, 2017. **8**(1): p. 2104.

78. Schmidt, N.W., G. Grigoryan, and W.F. DeGrado, *The accommodation index measures the perturbation associated with insertions and deletions in coiled-coils: Application to understand signaling in histidine kinases.* Protein Sci, 2017. **26**(3): p. 414-435.

79. Moglich, A. and K. Moffat, *Structural basis for light-dependent signaling in the dimeric LOV domain of the photosensor YtvA.* J Mol Biol, 2007. **373**(1): p. 112-26.

80. Gilles-Gonzalez, M.A., G.S. Ditta, and D.R. Helinski, *A haemoprotein with kinase activity encoded by the oxygen sensor of Rhizobium meliloti.* Nature, 1991. **350**(6314): p. 170-2.

81. Moglich, A., R.A. Ayers, and K. Moffat, *Design and signaling mechanism of light-regulated histidine kinases.* J Mol Biol, 2009. **385**(5): p. 1433-44.

82. Diensthuber, R.P., et al., *Full-Length Structure of a Sensor Histidine Kinase Pinpoints Coaxial Coiled Coils as Signal Transducers and Modulators.* Structure, 2013. **21**(7): p. 1127-1136.

83. Berntsson, O., et al., *Sequential conformational transitions and alpha-helical supercoiling regulate a sensor histidine kinase.* Nature Communications, 2017. **8**.

84. Circolone, F., et al., *Structural basis for the slow dark recovery of a full-length LOV protein from Pseudomonas putida.* J Mol Biol, 2012. **417**(4): p. 362-74.

85. Ohki, M., et al., *Structural insight into photoactivation of an adenylate cyclase from a*

photosynthetic cyanobacterium. Proceedings of the National Academy of Sciences of the United States of America, 2016. **113**(24): p. 6659-6664.

86. Lindner, R., et al., *Photoactivation Mechanism of a Bacterial Light-Regulated Adenylyl Cyclase.* J Mol Biol, 2017. **429**(9): p. 1336-1351.

87. Scheib, U., et al., *Rhodopsin-cyclases for photocontrol of cGMP/cAMP and 2.3 angstrom structure of the adenylyl cyclase domain.* Nature Communications, 2018. **9**.

88. Gasser, C., et al., *Engineering of a red-light–activated human cAMP/cGMP-specific phosphodiesterase.* Proceedings of the National Academy of Sciences, 2014. **111**(24): p. 8803-8808.

89. Conti, M. and J. Beavo, *Biochemistry and physiology of cyclic nucleotide phosphodiesterases: essential components in cyclic nucleotide signaling.* Annu Rev Biochem, 2007. **76**: p. 481-511.

90. Francis, S.H., M.A. Blount, and J.D. Corbin, *Mammalian cyclic nucleotide phosphodiesterases: molecular mechanisms and physiological functions.* Physiol Rev, 2011. **91**(2): p. 651-90.

91. Yoshida, K., et al., *A unique choanoflagellate enzyme rhodopsin exhibits light-dependent cyclic nucleotide phosphodiesterase activity.* J Biol Chem, 2017. **292**(18): p. 7531-7541.

92. Lamarche, L.B., et al., *Purification and Characterization of RhoPDE, a Retinylidene/Phosphodiesterase Fusion Protein and Potential Optogenetic Tool from the Choanoflagellate Salpingoeca rosetta.* Biochemistry, 2017. **56**(43): p. 5812-5822.

93. Tischer, D. and O.D. Weiner, *Illuminating cell signalling with optogenetic tools.* Nat Rev Mol Cell Biol, 2014. **15**(8): p. 551-8.

94. Kennedy, M.J., et al., *Rapid blue-light-mediated induction of protein interactions in living cells.* Nat Methods, 2010. **7**(12): p. 973-5.

95. Rivera-Cancel, G., L.B. Motta-Mena, and K.H. Gardner, *Identification of natural and artificial DNA substrates for light-activated LOV-HTH transcription factor EL222.* Biochemistry, 2012. **51**(50): p. 10024-34.

96. Motta-Mena, L.B., et al., *An optogenetic gene expression system with rapid activation and deactivation kinetics.* Nat Chem Biol, 2014. **10**(3): p. 196-202.

97. Zhao, E.M., et al., *Optogenetic regulation of engineered cellular metabolism for microbial chemical production.* Nature, 2018. **555**(7698): p. 683-687.

98. Shu, X., et al., *Mammalian expression of infrared fluorescent proteins engineered from a bacterial phytochrome.* Science, 2009. **324**(5928): p. 804-7.

99. Kaberniuk, A.A., A.A. Shemetov, and V.V. Verkhusha, *A bacterial phytochrome-based optogenetic system controllable with near-infrared light.* Nature Methods, 2016. **13**(7): p. 591-+.

100. Redchuk, T.A., et al., *Near-infrared optogenetic pair for protein regulation and spectral multiplexing.* Nature Chemical Biology, 2017. **13**(6): p. 633-+.

101. Heckman, K.L. and L.R. Pease, *Gene splicing and mutagenesis by PCR-driven overlap extension.* Nat Protoc, 2007. **2**(4): p. 924-32.

102. Kerppola, T.K., *Bimolecular fluorescence complementation (BiFC) analysis as a probe of protein interactions in living cells.* Annu Rev Biophys, 2008. **37**: p. 465-87.

103. Dobson, L., I. Remenyi, and G.E. Tusnady, *CCTOP: a Consensus Constrained TOPology prediction web server.* Nucleic Acids Res, 2015. **43**(W1): p. W408-12.

104. McDonnell, A.V., et al., *Paircoil2: improved prediction of coiled coils from sequence.* Bioinformatics, 2006. **22**(3): p. 356-8.

105. Marchler-Bauer, A., et al., *CDD: NCBI's conserved domain database.* Nucleic acids research, 2014:

p. gku1221.

106. Rauch, A., et al., *Crystal structure of the guanylyl cyclase Cya2.* Proceedings of the National Academy of Sciences, 2008. **105**(41): p. 15720-15725.

107. Raffelberg, S., et al., *A LOV-domain-mediated blue-light-activated adenylate (adenylyl) cyclase from the cyanobacterium Microcoleus chthonoplastes PCC 7420.* Biochemical Journal, 2013. **455**(3): p. 359-365.

108. Zhou, X.X., et al., *Optical control of cell signaling by single-chain photoswitchable kinases.* Science, 2017. **355**(6327): p. 836-842.

109. Pandit, J., et al., *Mechanism for the allosteric regulation of phosphodiesterase 2A deduced from the X-ray structure of a near full-length construct.* Proc Natl Acad Sci U S A, 2009. **106**(43): p. 18225-30.

110. Prochnik, S.E., et al., *Genomic analysis of organismal complexity in the multicellular green alga Volvox carteri.* Science, 2010. **329**(5988): p. 223-6.

111. Nordberg, H., et al., *The genome portal of the Department of Energy Joint Genome Institute: 2014 updates.* Nucleic Acids Res, 2014. **42**(Database issue): p. D26-31.

112. Avelar, G.M., et al., *A Cyclic GMP-Dependent K+ Channel in the Blastocladiomycete Fungus Blastocladiella emersonii.* Eukaryot Cell, 2015. **14**(9): p. 958-63.

113. Ullrich, S., R. Gueta, and G. Nagel, *Degradation of channelopsin-2 in the absence of retinal and degradation resistance in certain mutants.* Biol Chem, 2013. **394**(2): p. 271-80.

114. Kianianmomeni, A. and A. Hallmann, *Algal photoreceptors: in vivo functions and potential applications.* Planta, 2014. **239**(1): p. 1-26.

115. Nishii, I. and S.M. Miller, *Volvox: simple steps to developmental complexity?* Curr Opin Plant Biol, 2010. **13**(6): p. 646-53.

116. Nagel, G., et al., *Functional expression of bacteriorhodopsin in oocytes allows direct measurement of voltage dependence of light induced H+ pumping.* FEBS Lett, 1995. **377**(2): p. 263-6.

117. Habuchi, S., et al., *Reversible single-molecule photoswitching in the GFP-like fluorescent protein Dronpa.* Proc Natl Acad Sci U S A, 2005. **102**(27): p. 9511-6.

6. Appendix

6.1 Supplementary Table

Table S1 List of primers used in this study.

Name	Primers 5'→3'	Remarks
Cr2c-Cyclop, Vc2c-Cyclop, Cop5 (YFP tag fused as indicated in Figure legends)		
BamHI 5' start F	CGGGATCCGCCACCATGAAGCT	**For K298A, H352F, G533A mutant**
1080 SacI H352F R	GGTTCTGAGCTCGAAGCTCATCAGGGACATGAACT	**For H352F mutant**
1060 SacI H352F F	CTGATGAGCCACGAGCTCAG	**For H352F mutant**
1800 NcoI R	CCTCCGCTGCACCATGGTAT	**For K298A, H352F, G533A mutant**
Rho BiFC KpnI F	GGGGTACCATGAAGCTGAGACAGAGA	**For BiFC construct**
Rho BiFC XhoI R	CCGCTCGAGCACGCTGTCTCTCAGGTC	**For BiFC construct**
Cr2c His-K NcoI F	CATGCCATGGCCAAGGTGCTGTTCAG	**For his-kinase purification**
Cr2c His-K XhoI F	CCGCTCGAGTTAGCCCTCGTGGGCTTCCAC	**For his-kinase purification**
Cop5Rho BiFC KpnI F	CGGGGTACCAAGGCACCCACCGGCAGCCT	**For BiFC construct**
Cop5Rho BiFC XhoI R	CCGCTCGAGCTGGATCACTTGGATGCGGT	**For BiFC construct**
Vc2cRho BiFC KpnI F	CGGGGTACCAAGCGCACCGGACTGCCA	**For BiFC construct**
Vc2cRho BiFC SalI R	ACGCGTCGACAGCCAACAACCGCCGCT	**For BiFC construct**
900 AclI K298A F	CCTGAACGTTTTCGCCAACTTCATGGCCGCGGTGCTGTTCAGCAGCAGCA	**For K298A mutant**
890 AclI R	GAAGTTGGCGAAAACGTTCAGGT	**For K298A mutant**
1590 KpnI G533A F	GGGGTACCGGCCTGGCACTGAACATCGTGAAGCAGC	**For G533A mutant**
1590 KpnI R	GGGGTACCGCCGTACTTCCGGGTCACGC	**For G533A mutant**
1800 NcoI F	CCAGGGATACCATGGTGCA	**For D1092T mutant**
3300 BspHI D1092T R	TCCGCTCATGACTGGCATCATGCATGTGAGGAGGATCAAATCTGG	**For D1092T mutant**
3300 BspHI F	GCATGATGCCAGTCATGAG	**For D1092T mutant**
3950 EcoRI R	CTCCAATGCGAATTCGGAGT	**For D1092T mutant**
RC BamHI start F	CGGGATCCGCCACCATGCTCTCGGTCGACGATGAGGA	**For truncated RC construct**
RC XhoI R	CAGTCCGCTCGAGTGCGGCCGCTGGCA	**For truncated RC construct**
SrRhoPDE (YFP tag fused as indicated in Figure legends)		
RP BiFC KpnI F	CGGGGTACCGGCCGTAAAAATGCCGCCAA	**For BiFC construct**
RP BiFC XhoI R	CCGCTCGAGCATGCCGGTCAGGGCA	**For BiFC construct**
RP no ATG XhoI F	CCGCTCGAGAAGGGCCGTAAAAATGCCGCCAA	**RhoPDE fusing YFP in N-terminal**
RP HindIII R	GCCCAAGCTTATCAGTGCATCTGGCGA	**RhoPDE fusing YFP in N-terminal**

RhoPDE K296A qcF	GCCGCCGCTGTTGGCCTGGCCA	For RhoPDE K296A mutant
RhoPDE K296A qcR	GCCAACAGCGGCGGCGTAATC	For RhoPDE K296A mutant
RhoPDE K296M qcF	GCCGCCATGGTTGGCCTGGCCA	For RhoPDE K296M mutant
RhoPDE K296M qcR	GCCAACCATGGCGGCGTAATC	For RhoPDE K296M mutant
RhoPDE M459D qcF	GCCGTGGACGTGCTGCATACCACA	For RhoPDE M459D mutant
RhoPDE M459D qcR	GCAGCACGTCCACGGCGTGGC	For RhoPDE M459D mutant
RhoPDE L623F qcF	CCGTTTCCAAAAAGAGTTCAACAACCAG	For RhoPDE L623F mutant
RhoPDE L623F qcR	CTCTTTTTGGAAACGGGCGCTCCAA	For RhoPDE L623F mutant
RhoPDE E657Q qcF	AAGGGCCAGATCGGCTTCATTAACT	For RhoPDE E657Q mutant
RhoPDE E657Q qcR	GCCGATCTGGCCCTTGGCCAGA	For RhoPDE E657Q mutant
RhoPDE D189C qcF	GTGTCGTGTATCGCCACAGCCGAT	For RhoPDE D189C mutant
RhoPDED189C qcR	GGCGATACACGACACGCTACTACG	For RhoPDE D189C mutant
RhoPDE T192C qcF	ATCGCCTGTGCCGATGCCTTGTTGC	For RhoPDE T192C mutant
RhoPDE T192C qcR	ATCGGCACAGGCGATATCCGACAC	For RhoPDE T192C mutant
RhoPDE C169T qcF	ATCACGACTCCCACGCTGTTGTA	For RhoPDE C169T mutant
RhoPDE C169T qcR	CGTGGGAGTCGTGATGATCCACTC	For RhoPDE C169T mutant
RhoPDE F451W qcF	GAACCCATGGCACAATTGGCGCCA	For RhoPDE F451W mutant
RhoPDE F451W qcR	CAATTGTGCCATGGGTTCTTGTTGTAC	For RhoPDE F451W mutant
RhoPDE M564F qcF	CACCGATTTCGCCACGCACCACA	For RhoPDE M564F mutant
RhoPDE M564F qcR	GTGGCGAAATCGGTGTTCAGGAT	For RhoPDE M564F mutant
RhoPDE F627W qcF	CAAAAAGAGTGGAACAACCAGGTCGA	For RhoPDE F627W mutant
RhoPDE F627W qcR	GGTTGTTCCACTCTTTTTGCAGACG	For RhoPDE F627W mutant
RhoPDE F627Y qcF	CAAAAAGAGTACAACAACCAGGTCGA	For RhoPDE F627Y mutant
RhoPDE F627Y qcR	GGTTGTTGTACTCTTTTTGCAGACG	For RhoPDE F627Y mutant
RhoPDE S1 qcF	CATGCTGAAAGAGCTGCATGTGATCGAGC	For RhoPDE S1 mutant
RhoPDE S1 qcR	CATGCAGCTCTTTCAGCATGCCGGTCAGG	For RhoPDE S1 mutant
RhoPDE S2 qcF	CATGCTGAAATTGGCTGCCAGCAGCCT	For RhoPDE S2 mutant
RhoPDE S2 qcR	GGCAGCCAATTTCAGCATGCCGGTCA	For RhoPDE S2 mutant
RhoPDE S3 qcF	CATGCTGAAAGGACATTCCGCCGCTAT	For RhoPDE S3 mutant
RhoPDE S3 qcR	GGAATGTCCTTTCAGCATGCCGGTCA	For RhoPDE S3 mutant
RhoPDE S4 qcF	GCTGGATGGTTTGGCTGCCAGCAGCCT	For RhoPDE S4 mutant
RhoPDE S4 qcR	GGCAGCCAAACCATCCAGCTGGGCCAT	For RhoPDE S4 mutant
RhoPDE S5 qcF	GCTGGATGGTGGACATTCCGCCGCTAT	For RhoPDE S5 mutant
RhoPDE S5 qcR	CGGAATGTCCACCATCCAGCTGGGCCAT	For RhoPDE S5 mutant
RhoPDE S6 qcF	GATGATTTGGCTGCCAGCAGCCT	For RhoPDE S6 mutant
RhoPDE S6 qcR	CAGCCAAATCATCATCGACACCATCCA	For RhoPDE S6 mutant

RhoPDE S7 qcF	GATGATAAGTTGGCTGCCAGCAG	For RhoPDE S7 mutant
RhoPDE S7 qcR	CCAACTTATCATCATCGACACCATCCA	For RhoPDE S7 mutant
RhoPDE S8 qcF	GATGATCGTGCCCGTGGACGCAAAG	For RhoPDE S8 mutant
RhoPDE S8 qcR	GGCACGATCATCATCGACACCATCCA	For RhoPDE S8 mutant

*Hs*PDE2A (catalytic domain), LOV, Dronpa domain (YFP tag fused as indicated in Figure legends)

PDE2a XbaI F	GCTCTAGAGCCACCATGAAGAAAGTTAACGAGGCTCAGTATCGCAGC	For PDE2a N1, N2 cloning
PDE2a N1 HindIII R	GCCCAAGCTTAGTCACAGGAGGTCATGAG	For PDE2a N1 cloning
PDE2a N2 HindIII R	GCCCAAGCTTAGGTCTGGTCAGAGAGGTC	For PDE2a N2 cloning
PDE2a C1 BamHI F	CGCGGATCCGCCACCATGCTCTCTGACCAGACCAAG	For PDE2a C1 cloning
PDE2a C2 BamHI F	CGCGGATCCGCCACCATGAAGGGCTGGAAGACTACG	For PDE2a C2 cloning
PDE2a HindIII R	GCCCAAGCTTACTCAGCATCAAGGCTGCA	For PDE2a C1, C2 cloning
LOV BamHI-NcoI F	CGGGATCCGCCACCATGGGATGTACAGTCTCTGCA	For LOV-PDE2a-YFP construct
LOV XbaI noStop R	GCTCTAGATTCTGCAGTTTTCTTAATCAGCATG	For LOV-PDE2a-YFP construct
PDE2a XbaI noATG F	GCTCTAGAAAGAAAGTTAACGAGGCTCAGTATCGCAGC	For LOV-PDE2a-YFP construct
PDE2a SalI noStop R	ACGCGTCGACCTCAGCATCAAGGCTGCA	For LOV-PDE2a-YFP construct
Dronpa BamHI F	CGGGATCCGCCACCATGAGCGTGATCAAGCCCGAC	For Dronpa-PDE2a-Dronpa clone
Dronpa XbaI noStop R	GCTCTAGAGCCCTTGGCCTGCCTG	For Dronpa-PDE2a-Dronpa clone
Dronpa SalI noATG F	ACGCGTCGACAGCGTGATCAAGCCCGAC	For Dronpa-PDE2a-Dronpa clone
Dronpa Stop HindIII R	GCCCAAGCTTAGCCCTTGGCCTGCCTG	For Dronpa-PDE2a-Dronpa clone
pGEM- 5'F	GCGAATTAATTCGAGCTCGGT	For Chimera1,3 cloning
BeGC-Op-PDE R	ACTATGCATTTGCTGCACCAGCTGTTCCAGCACGCC	For Chimera1 cloning
BeGC-Op-PDE F	GGCGTGCTGGAACAGCTGGTGCAGCAAATGCATAGT	For Chimera1 cloning
pGEM-3'R	GAGCAGATACGAATGGCTACAT	For Chimera1 cloning
Cr2c-Op 3146SalI F	CGGTCGACGATGAGGATATC	For Chimera2 cloning
Cr2c-PDE2a R	CTGAGCCTCATTCACTTTGAGCTTCATCGACTCTGT	For Chimera2 cloning
Cr2c-PDE2a F	ACAGAGTCGATGAAGCTCAAAGTGAATGAGGCTCAG	For Chimera2 cloning
PDE2a HindIII R	GCCCAAGCTTACTCAGCATCAAGGCTGCA	For Chimera2,3 cloning
BeGC-Op-PDE2a F	GGCGTGCTGGAACAGCTGAAAAAAGTGAATGAGGCT	For Chimera3 cloning
BeGC-Op-PDE2a R	AGCCTCATTCACTTTTTTCAGCTGTTCCAGCACGCC	For Chimera3 cloning

6.2 Supplementary Figures

Figure S1 Alignment between *Cr*2c-Cyclop and *Vc*2c-Cyclop.

Alignment of the full-length of *Cr*2c-Cyclop (Cre11.g467678) and *Vc*2c-Cyclop (Vocar.0009s0380.1) protein sequences from JGI genome database. Four conserved domains are labeled in colored boxes, key residues are depicted with the red color background. Sequences with yellow background indicate short middle linker and long C-terminal regions used or not in individual experiments.
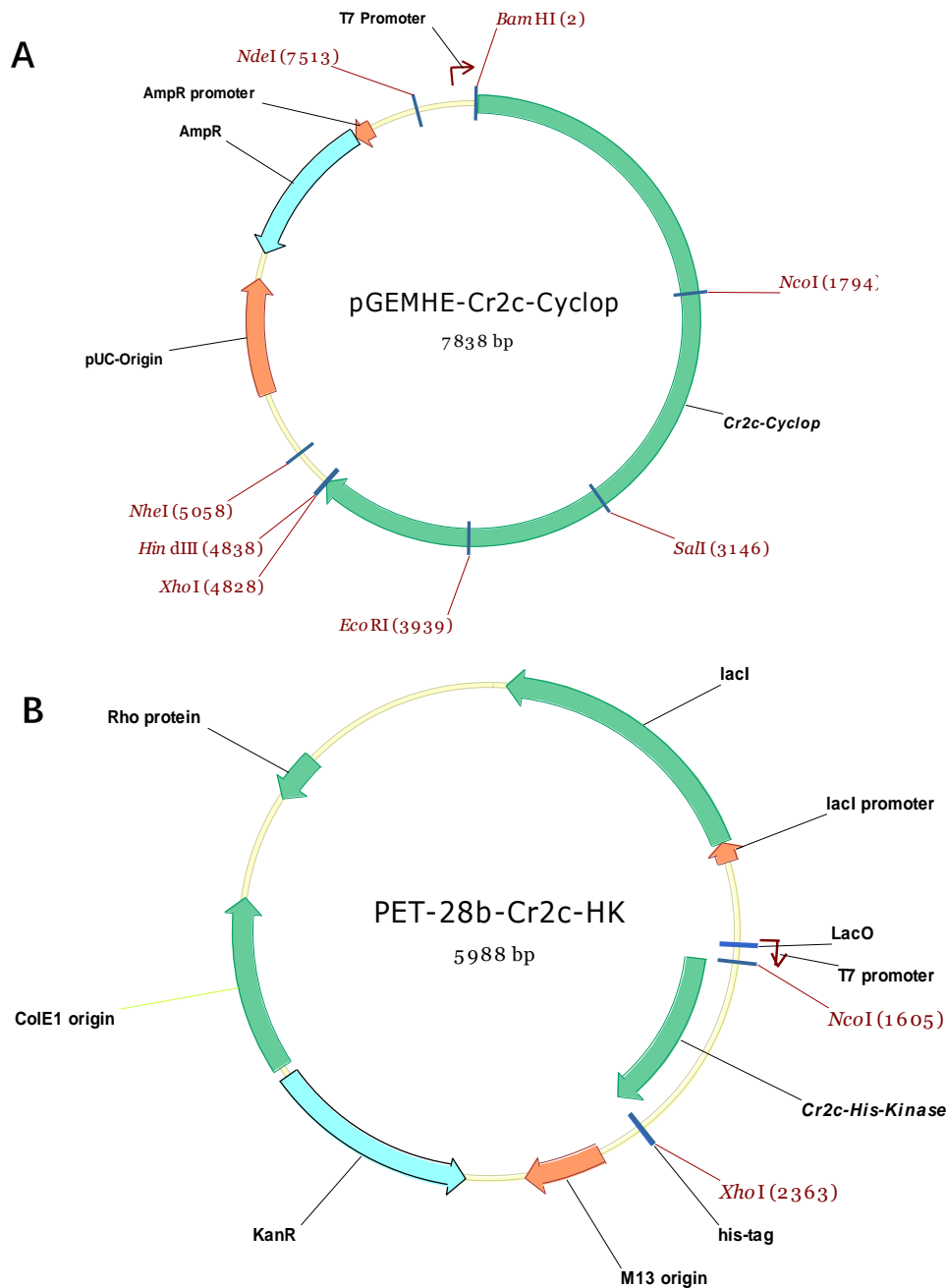
A, pGEMHE vector was applied to generate most constructs. Here is an example fusing *Cr2c-Cyclop* in multiple cloning sites (MCS) region with 5'-BamHI and 3'-XhoI-HindIII with a stop codon. 5' XhoI-YFP tag-3' HindIII was then fused in 3' cloning site without stop codon if necessary. The plasmid was extracted from transfected *E. coli* MRF strain and linearized by NheI. Linearized DNA was then prepared for *in vitro* transcription to obtain cRNA for *Xenopus* oocytes injection.

B, PET-28b vector was applied to overexpress the targeted protein. The his-kinase DNA fragment (251aa) was fused in multiple cloning sites (MCS) region with 5'-NcoI and 3'-XhoI with stop codon before his-tag. Transfected *E. coli* BL21(DE3) strain was employed to induce the recombinant protein expression and inclusion body purification without his-tag.

Both vectors were depicted by Vector NTI software.

6.3 Abbreviation

2c-Cyclop	two-**c**omponent **Cycl**ase **op**sin
AC	Adenylyl cyclase
AMP-PNP	Adenylyl-imidodiphosphate
APSSP	Advanced Protein Secondary Structure Prediction
ATP	Adenosine triphosphate
ATR	All-*trans*-retinal
BiFC	Bimolecular fluorescence complementation
BLUF	Blue-light using FAD
BR	Bacteriorhodopsin
BSA	Bovine serum albumin
CA	**C**atalytic and **A**TP-binding
cAMP	cyclic adenosine 3'-5'-monophosphate
CC	**C**oiled-**C**oils
CCTOP	Constrained Consensus TOPology prediction
c-di-GMP	cyclic diguanylate
cDNA	complementary DNA
CFTR	Cystic fibrosis transmembrane conductance regulator
cGMP	cyclic guanosine 3'-5'-monophosphate
ChIP-Seq	Chromatin immunoprecipitation and high-throughput sequencing
ChR1	Channelrhodopsin-1
ChR2	Channelrhodopsin-2
cN	cyclic Nucleotide
CNG	Cyclic nucleotide gated
CREB	cAMP/cGMP response element binding
cRNA	complimentary ribonucleic acid
CyclOp	**Cycl**ase **Op**sin
DHp	**D**imerization and **H**istidine-containing **p**hosphotransfer
dpi	days post injection

DTT	Dithiothreitol
ELISA	Enzyme-linked immunosorbent assay
EST	Expressed sequence tag
EDTA	Ethylenediaminetetraacetic acid
EGTA	Ethylene glycol tetraacetic acid
FAD	Flavin adenine dinucleotide
FMN	Flavin mononucleotide
GAF	domains from c**G**MP-specific phosphodiesterases, **A**denylyl cyclases and **F**hlA
GC	Guanylyl Cyclase
GCAP	Guanylyl cyclase activating protein
GDP	Guanosine diphosphate
GPCR	G protein-coupled receptor
GRK	G protein receptor kinase
GTP	Guanosine-5'-triphosphate
HK	**H**istidine **K**inase
HR	Halorhodopsin
IFPs	Infrared-fluorescent proteins
IPTG	Isopropyl β-D-Thiogalactoside
K_{cat}	turnover rate
kDa	kilodalton
K_m	Michaelis-Menten constant
LAPD	Light-activated phosphodiesterase
LOV	Light-oxygen-voltage
MCS	Multiple cloning sites
NIR	Near-infrared
PAC	Photoactivated adenylyl cyclase
PAS	Per-Arnt-Sim
pdDronpa	photodissociable dimeric Dronpa
PDE	Phosphodiesterase
RC	**R**esponse regulator and **C**yclase domain

RhoPDE	Rhodopsin phosphodiesterase
OD	Optical density
PKA	Protein kinase A
PKG	Protein kinase G
RR	**R**esponse **R**egulator
RT-PCR	Reverse transcription PCR
SDS	Sodium dodecyl sulfate
SDS-PAGE	SDS-polyacrylamide gel electrophoresis
TCS	Two-component system
TM	Transmembrane helix
UCR	Upstream conserved regulatory
UV	Ultra violet
WT	Wild type
YFP	Yellow fluorescent protein

Acknowledgement

Firstly, I would like to show my deep appreciation to Prof. Dr. Georg Nagel as my primary supervisor, who gives me many instructions to research on intriguing optogenetic fields during my PhD period. As one of the creators of "optogenetics", Prof. Nagel is specialized in the discovery and development of a fantastic story about ChR2 from his solid scientific work. He has built efficient set-ups in our lab and developed a very good scientific relationship with other collaborators. This allows me to obtain relative knowledge more efficiently. In addition, I appreciate my secondary supervisor Prof. Dr. Robert Kittel who is an expert in the neuroscientific field by using *Drosophila* as a model organism. This enables me to understand how to use optogenetic tools by *in vivo* experiments. I also appreciate my third supervisor Prof. Dr. Thomas Müller, who is an expert in structural biology and protein chemistry. This allows me to put efforts to know more about molecular mechanisms in structure views.

Secondly, I would like to deeply appreciate Dr. Gao Shiqiang, who is the Postdoc in our group from the beginning of my PhD period. He is an expert to organize many scientific projects and put them forward step by step. With his instructions, I learned many detailed experiments to find out more interesting and meaningful points in my projects. This makes me consider more detailed experimental designs to solve scientific questions. I also thank our lab members Elfriede Reisberg, Yang Shang, Yu Jing, Duan Xiaodong, Sebastian Beck, Zhou Yang, Tang Ruijing, Julia Köber, who give me immediate help in my experiments. Thanks to other lab members in Botany I institute, giving me help and a very good experimental environment. Thanks to other friends Zhu Mo, Huang Hua, Cheng Cheng from Botany II, giving me many instructions about GSLS procedures.

Many thanks to all friends and teachers in China and Germany for the support and help.

Last but not least, I deeply appreciate my family members and all my relatives. My parents put much effort to support and help me all the time and encourage me to overcome any difficulties in my life. I also deeply appreciate my wife Liu Ruiqi, who continuously gives me excellent assistance in my work and life. This allows me to possess positive attitudes in my life thanks to big harmonious family members and relatives.

Affidavit

I hereby confirm that my thesis entitled "Characterization of novel rhodopsins with light-regulated cGMP production or cGMP degradation" is the result of my own work. I did not receive any help or support from commercial consultants. All sources and / or materials applied are listed and specified in the thesis.

Furthermore, I confirm that this thesis has not yet been submitted as part of another examination process neither in identical nor in similar form.

Wuerzburg, 23.07.2018

Place, Date   Signature

Eidesstattliche Erklärung

Hiermit erkläre ich an Eides statt, die Dissertation „Charakterisierung neuartiger Rhodopsine mit Licht-regulierter cGMP-Produktion oder cGMP-Degradation" eigenständig, d.h. insbesondere selbständig und ohne Hilfe eines kommerziellen Promotionsberaters, angefertigt und keine anderen als die von mir angegebenen Quellen und Hilfsmittel verwendet zu haben.

Ich erkläre außerdem, dass die Dissertation weder in gleicher noch in ähnlicher Form bereits in einem anderen Prüfungsverfahren vorgelegen hat.

Würzburg, 23.07.2018

Ort, Datum   Unterschrift

Curriculum Vitae