# Functional and Structural Studies of a Class III Photolyase and a Prokaryotic (6-4) Photolyase from *Agrobacterium tumefaciens*

Zur Erlangung des akademischen Grades eines
DOKTORS DER NATURWISSENSCHAFTEN

(Dr. rer. Nat.)

der Fakultät für Chemie und Biowissenschaften
des Karlsruher Instituts für Technologie (KIT)-Universitätsbereich

genehmigte

Dissertation

von

Fan Zhang

Aus Shaanxi, China

Dekan: Prof. Dr. Martin Bastmeyer

Referent: Prof. Dr. Tilman Lamparter

Korreferent: Prof. Dr. Holger Puchta

Tag der mündlichen Prüfung: 19. Juli. 2013

Die vorliegende Dissertation wurde am Botanischen Institut des Karlsruhe Instituts für Technologie (KIT), Lehrstuhl 1 für Molekulare Zellbiologie, im Zeitraum von September 2009 bis Juni 2013 angefertigt.

Hiermit erkläre ich, dass ich die vorliegende Dissertation, abgesehen von der Benutzung der angegebenen Hilfsmittel, selbständig verfasst habe.

Alle Stellen, die gemäß Wortlaut oder Inhalt aus anderen Arbeiten entnommen sind, wurden durch Angabe der Quelle als Entlehnungen kenntlich gemacht.

Diese Dissertation liegt in gleicher oder ähnlicher Form keiner anderen Prüfungsbehörde vor.

Karlsruhe, den 4. Mai 2013

Fan Zhang

## Zusammenfassung

Ultraviolette Strahlung kann DNA durch die Entstehung von Cyclobutan Pyrimidin Dimeren (CPD) und (6-4)-Photoprodukten schädigen. Solche Schäden können Mutationen induzieren, was den Zelltod und die Entstehung von Krebs nach sich ziehen kann. Um die Integrität des Genoms von Organismen zu gewährleisten, entwickelten sich Photolyasen zur Reparatur solcher UV-Schäden der DNA. Unter den Photolyasen lassen sich anhand ihrer Substratspezifität CPD Photolyasen und (6-4)-Photolyasen unterscheiden. Erstere wurden bisher aus allen drei Domänen des Lebens identifiziert und sind als die Vorfahren der Cryptochrom/Photolyase Familie (CPF) akzeptiert. Im Gegensatz dazu wurden (6-4)-Photolyasen nur in Eukaryoten gefunden, weshalb diese Proteine als eine spätere Entwicklung während der Evolution der Cryptochrom / Photolyase Familie angesehen wurden. Aktuellen phylogenetischen Studien zufolge können die Mitglieder der CPF in sieben Hauptklassen unterteilt werden. Gemäß dieser Untersuchungen existieren drei Klassen von CPD Photolyasen (Klasse I bis III), die Klasse der CRY-DASH Proteine, die Klasse der (6-4)-Photolyasen und tierischen Cryptochrome, die Klasse der pflanzlichen Cryptochrome sowie eine neue, als FeS-BCPs bezeichnete Klasse. Das Genom von *Agrobacterium tumefaciens* codiert für zwei photolyaseähnliche Proteine, PhrA und PhrB. Erstgenanntes gehört zur Klasse III der CPD Photolyasen, der Schwesterklasse der pflanzlichen Cryptochrome. PhrB hingegen gehört zur Fes-BCP-Gruppe, die entfernt mit den anderen CPFs verwandt ist.

Die vorliegende Dissertation umfasst eine funktionelle Charakterisierung und die Bestimmung der Kristallstrukturen von PhrA und PhrB. Mit einer Auflösung von 1,67 Å konnte die Kristallstruktur von PhrA als erste Struktur einer Klasse III Photolyase bestimmt werden. Diese Strukturauflösung zeigte eine neue Bindestelle des Antennenchromophors, in der 5,10-Methenyltetrahydrofolat (MTHF) durch π-Stapelung stabilisiert wird. Darüber hinaus konnte neben der klassischen eine zusätzliche Tryptophan-Triade identifiziert werden.

Diese beiden Ergebnisse wurden durch ortsspezifische Mutagenese spezifischer Aminosäuren vertiefend untersucht. Die Ergebnisse dieser strukturellen Untersuchungen von PhrA könnten durch dessen enge phylogenetische Homologie ein klareres Bild auf den evolutionären Übergang von Photolyasen zu Photorezeptoren aufzeigen.

Durch HPLC basierte Reparaturuntersuchungen *in vitro* konnte PhrB als (6-4)-Photolyase identifiziert werden. Darüber hinaus konnte die Kristallstruktur dieses Proteins mit einer Auflösung von 1.45 Å gewonnen werden. Im Gegensatz zu prototypischen CPFs enthält PhrB 6,7-Dimethyl-8-ribityllumazin (DMRL) als Antennenchromophor und ein an die katalytische Domäne gebundenes [4Fe-4S]-Cluster. Ein wesentlicher Teil dieser Eisen-Schwefel-Domäne ähnelt der großen Untereinheit von eukaryotischen und archaebakteriellen Primasen, was auf eine Dichotomie der PhrB-änlichen Proteine an der Basis der Entwicklung der CPF hindeutet. Darüber hinaus legt die Struktur von PhrB einen Mechanismus der DNA-Bindung nahe, der sich von dem der eukaryotischen Vertreter der (6-4)-Photolyasen unterscheidet. Die vorliegende Studie identifiziert PhrB als erste prokaryotische (6-4)-Photolyase und präsentiert die die prokaryotischen (6-4) Photolyasen als Vorfahren der CPF.

## Summary

Ultraviolet irradiation damages DNA by producing cyclobutane pyrimidine dimers (CPD) and (6-4) photoproducts, leading to mutagenesis, cell death and cancer. Photolyases have evolved for repair of UV damaged DNA to maintain organism genetic integrity. Substrate specificity distinguishes photolyases either as CPD photolyases or as (6-4) photolyases. CPD photolyases have been identified in all three domains of life and are generally accepted as the common ancestor of the cryptochrome/photolyase family (CPF). By contrast, (6-4) photolyases have only been found in eukaryotes and are presumed as a later invention during evolution. Recent phylogenetic studies divided the CPF members (CPFs) into seven major classes, three classes of CPD photolyases termed class I to III, CRY-DASH proteins, (6-4) photolyases together with animal cryptochromes, plant cryptochromes and a new class, designated FeS-BCPs. *Agrobacterium tumefaciens* bears two photolyase related proteins, termed PhrA and PhrB. PhrA belongs to the class III CPD photolyases, the sister class of plant cryptochromes. PhrB belongs to the FeS-BCP group which is distantly related to the other CPFs.

This dissertation presents the functional characterization and crystal structure determination of PhrA and PhrB. The crystal structure of PhrA, determined at 1.67 Å resolution, is the first one from a class III photolyase. The model revealed a novel 5,10-methenyltetrahydrofolate (MTHF) antenna chromophore binding site in which the chromophore is stabilized by π-stacking. Besides the classical Trp-triad, an additional Trp-triad was found. Both findings were further investigated by site-directed mutagenesis analyses. Due to the close phylogenetic homology, the structural studies on PhrA could give a clearer picture on the transition from photolyase to photoreceptor.

We identified PhrB as a (6-4) photolyase through an HPLC-based in vitro repair assay and determined the PhrB crystal structure at 1.45 Å resolution. Unlike prototypical CPFs, PhrB

contains 6,7-dimethyl-8-ribityllumazine (DMRL) as an antenna chromophore and a [4Fe-4S] cluster bound to the catalytic domain. A significant part of the Fe-S fold resembles that of the large subunit of eukaryotic and archaeal primases, suggesting that the PhrB-like photolyases branched at the base of the evolution of the CPF. PhrB structure also suggests a DNA binding mode different from that of the eukaryotic counterparts. This study presents the first prokaryotic (6-4) photolyase and proposes that the prokaryotic (6-4) photolyases are the ancestors of the CPF.

## Acknowledgements

First and foremost, I would like to express my utmost gratitude to my advisor, Prof. Dr. Tilman Lamparter. Tilman allowed me the great freedom to my projects. I appreciate our countless discussions in his office or email. With his supervision and encouragement, I could continuously improve my knowledge and skills on research. I am thankful that Tilman gave me strong support throughout doctoral processing.

I am deeply indebted to Dr. Inga Oberpichler for her guidance and support. Inga established the PhrA and PhrB protein expression and purification procedures, which laid the foundation for my projects. Inga imparted me invaluable education on photolyases protein biochemistry and function studies. I appreciate her help for my skills on making scientific presentations and posters and writing fellowship applications.

I am very grateful to Dr. Patrick Scheerer for his guidance and support on crystallography. Together with Patrick Scheerer, protein crystallization and model building of PhrA and PhrB were performed. Most of the high quality crystallographic figures were drawn by Patrick Scheerer in the final form. Patrick gave me several times training on crystal mounting and X-ray diffraction data collection at European Synchrotron Radiation Facility (ESRF, Grenoble) as well as introduction about their lab during my visiting in Berlin.

I would like to acknowledge Dr. Norbert Krauß for his support and education on protein crystallography and biochemistry. Together with Dr. Norbert Krauß, the PhrB paper was written in high quality. His academic rigour ensured the acceptance of PhrB paper with minor changes. I appreciate his supervision on me for crystallographic training in London.

I would like to thank Janine Wesslowski and Dominik Graf for their works on PhrB. Janine solved the long lasting insoluble problem of PhrB expression. Dominik together with Janine did tremendous work on the PhrB cofactors identification and in vitro DNA repair assay,

which provided me invaluable support for designing the optimal experimental strategies. I am thankful that they both offered me strong support during my teaching of the "Modul 8".

I would like to thank Sybille Wörner and Ernst Heene for their technical assistance. Sybille Wörner gave me training on protein purification and laboratory routine work since the very beginning. I am thankful that Sybille Wörner gave me strong support during my teaching of the "Modul 8". Ernst Heene offered me great help on HPLC analysis.

I would like to thank Prof. Dr. Peter Nick for creating such a wonderful studying and living environment in Botanic Institute. I appreciate his concern about the scientific progress during my toughest period. I am highly impressed by his humorous, sagacious and erudite style.

I would like to thank Dr. Gregor Rottwinkel for discussions, advices and help. I am thankful that he offered me strong support for my teaching of the "Modul 8".

I would like to thank XiaoLi Chang, Rui Yang, Qiong Liu and Fei Qiao for help on my living and studying. I would like to thank my officemates Thomas Horn, Katharina Svyatyna, Njimona Ibrahim and Juan Feng for discussions, advices and good humor. I would like to thank all the inhabitants of the Botanic Institute for the nice working atmosphere.

I would like to thank all the helps from other labs. A quick and detailed critical reading for the PhrB manuscript by Richard Pokorny (Philipps-Universität Marburg). A generous gift of (6-4) photoproducts from Korbinian Heil and Thomas Carell (Ludwig-Maximilians-Universität München) and their advices in photorepair assays. The synthetic DMRL from Markus Fischer and Adelbert Bacher (Technische Universität München). And the help from David von Stetten (ESRF) on data collection of PhrA and PhrB.

I would like to thank all the coauthor of the PNAS publication of PhrB from "Zhang et. al" for their work on the text and figures that used in this thesis.

I would like to thank China Scholarship Council (CSC) for funding support throughout the last four years and thank Karlsruhe House of Young Scientists (KHYS) for the 5 months travel scholarship that supported my crystallographic training in London.

Last, I would like to thank my parents, brother, my wife, and friends for their support and love.

# Curriculum Vitae

## Personal Details

**Gender:** Male
**Nationality:** Chinese
**Date of Birth:** August 1, 1982
**Private address:** Kirchgrund.4, 75045, Walzbachtal
**Email:** yyyyzhang2008@163.com

## Academic Education & Training

**Since 09.2009** **PhD study in Biochemistry and Protein Structural Biology**

Karlsruhe Institute of Technology (KIT), Germany, Botanical Institute with Prof. Tilman Lamparter

**Thesis**: Functional and structural studies of a class III photolyase and a prokaryotic (6-4) photolyase from *Agrobacterium tumefaciens*

**09.2006-06.2009** **Master study in Plant Breeding and Biotechnology**

Northwest Agriculture and Forestry University, Key Laboratory of Northwest Horticulture Plant Germplasm and Genetic Improvement of Ministry of Agriculture of P. R. China with Prof. YueJin Wang

**Thesis**: Construction of an SSH library of DANGSHANSU pear leaves inoculated with *Venturia nashicola* and ZAOSU pear NPR1 gene cloning and prokaryotic expression

**09.2002-06.2006** **Bachelor study in Landscape Designing and Horticulture Plant**

Northwest Agriculture and Forestry University, China

**Thesis**: Optimized arrangement of garden trees in Shaanxi Guanzhong areas

## Awards & Scholarships

| | |
|---|---|
| 2012 | Five months travel scholarship from Karlsruhe House of Young Scientists |
| 2009-2013 | China Scholarship Council scholarship |
| 2003-2005 | First-class University Scholarship in three academic years |
| 2003 | First-class Scholarship of Fund for Asian Agricultural Development Research |
| 2002 | Second-class National Scholarship |

## Publications

***Fan Zhang**, Patrick Scheerer, Inga Oberpichler, Tilman Lamparter and Norbert Krauß. Crystal structure of a prokaryotic (6-4) photolyase with an Fe-S cluster and a 6,7-dimethyl-8-ribityllumazine antenna chromophore. **Proceedings of the National Academy of Sciences. (**2013) 110: 7217-7222.

*****Patrick Scheerer, **Fan Zhang**, David von Stetten, Norbert Krauß, Inga Oberpichler and Tilman Lamparter. Crystal structure of a CPD class III photolyase reveals unique antenna chromophore binding site and alternative photoreduction pathways. **(In preparation).**

Inga Oberpichler, Antonio J. Pierik, Janine Wesslowski, Richard Pokorny, Ran Rosen, Michal Vugman, **Fan Zhang**, Olivia Neubauer, Eliora Z. Ron, Alfred Batschauer and Tilman Lamparter. A photolyase-like protein from *Agrobacterium Tumefaciens* with an iron-sulfur cluster. **PLos One.** ( 2011) 6, e26775.

**Zhang Fan**, Zhang Chaohong, Xu Yan and Wang Yuejin. Cloning and prokaryotic expression of ZAOSU pear nonexpressor of pathogenesis-related genes 1 gene (NPR1) in *Escherichia coli*. **Journal of Agricultural Biotechnology**. (2010) 18(1): 18-23 (In Chinese).

***** **Fan Zhang** and Patrick Scheerer contributed equally to the work.

## List of abbreviations

| | |
|---|---|
| Å | angstrom |
| Amp | ampicillin |
| bp | base pair |
| CPD | cyclobutane pyrimidine dimer |
| CRY | cryptochrome protein |
| DMRL | 6,7-dimethyl-8-ribityl lumazine |
| DTT | dithiothreitol |
| EDTA | ethylenediaminetetraacetic acid |
| 8-HDF | 8-hydroxy-5-deazariboflavin |
| FAD | flavin adenine dinucleotide |
| g | gram |
| h | hours |
| HPLC | high performance liquid chromatography |
| IPTG | isopropyl- β-D-1-thiogalaktopyranosid |
| kD | kilo Dalton |
| L | liter |
| LB | luria broth |
| m | meter |

| | |
|---|---|
| M | molar |
| mg | milligram |
| µg | microgram |
| µl | microliter |
| µM | micromolar |
| min | minutes |
| ml | milliliter |
| MTHF | 5,10-methenyltetrahydrofolate |
| ng | nanogram |
| nm | nanometer |
| nt | nucleotide |
| PCR | polymerase chain reaction |
| PEG | polyethylenglykol |
| s | seconds |
| (6-4) PPs | pyrimidine-pyrimidone-(6-4)-photoproducts |
| UV | ultraviolet |

# Contents

# 1 Introduction

## 1.1 Ultraviolet (UV) radiation damage to DNA

Sunlight contains three types of UV radiation termed as UV-A (320–400 nm), UV-B (295–320 nm) and UV-C(100–295 nm). The Earth's ozone layer absorbs all UV-C and approximately 90% of UV-B. Therefore, the UV radiation reaching the Earth's surface consists mainly UV-A and a small UV-B component. Most of the mutagenic and carcinogenic of sunlight are attributed to UV-B (Batista *et al.*, 2009; Perdiz *et al.*, 2000; van der Leun, 2004). The depletion of the atmospheric ozone layer results in intensified UV-B radiation on the Earth's surface. This effect will result in increasingly severe damage of the biosphere (Aucamp, 2007).

Exposure to UV radiation may lead to DNA photoproducts between adjacent pyrimidines within the same strand, mostly cyclobutane pyrimidine dimers (CPDs) and pyrimidine-pyrimidone-(6-4)-photoproducts ((6-4) PPs) (Batista *et al.*, 2009; Ravanat *et al.*, 2001). The CPDs are formed by a [2π + 2π] cycloaddition reaction between the two C5=C6 bonds of the adjacent pyrimidines. The (6–4) PPs are formed due to Paternó–Büchi reaction, a cycloaddition between the C5=C6 bond of the 5' nucleoside and the C4 carbonyl group of the 3' nucleoside, which proceeds via an oxetane or azetidine intermediate. These intermediates are instable above -80 °C and undergo rapid ring opening to the (6-4)PPs (Fig. 1.1) (Heil *et al.*, 2011). CPDs are the most abundant UV-induced DNA lesions and account for ~80% of the total amount of the photoproducts. They occur between all pyrimidine pairs, but with different frequency. The decreasing order of the formation frequency is: 5'-T<>T-3', 5'-T<>C-3', 5'-C<>T-3', 5'-C<>C-3'. The yield of T<>T is about three times higher than that of T<>C. The remaining ~20% of (6–4) PPs are formed more efficiently at TC rather than TT sites (Heil *et al.*, 2011).

UV-induced DNA lesions are responsible for much of the destructive effect of UV light since they can act as a physical blockage for replication and transcription, thereby drastically impeding metabolic processes in DNA. UV damage may subsequently lead to mutagenesis or

cell death .These DNA photoproducts triggered mutations are predominately found as C to T or CC to TT transitions. Mutations on proteins involved in cell cycle control, apoptosis, or DNA repair could result in carcinogenesis, for example, the p53 mutation that caused skin cancer (Batista *et al.*, 2009; Kanavy & Gerstenblith, 2011; Pfeifer *et al.*, 2005).

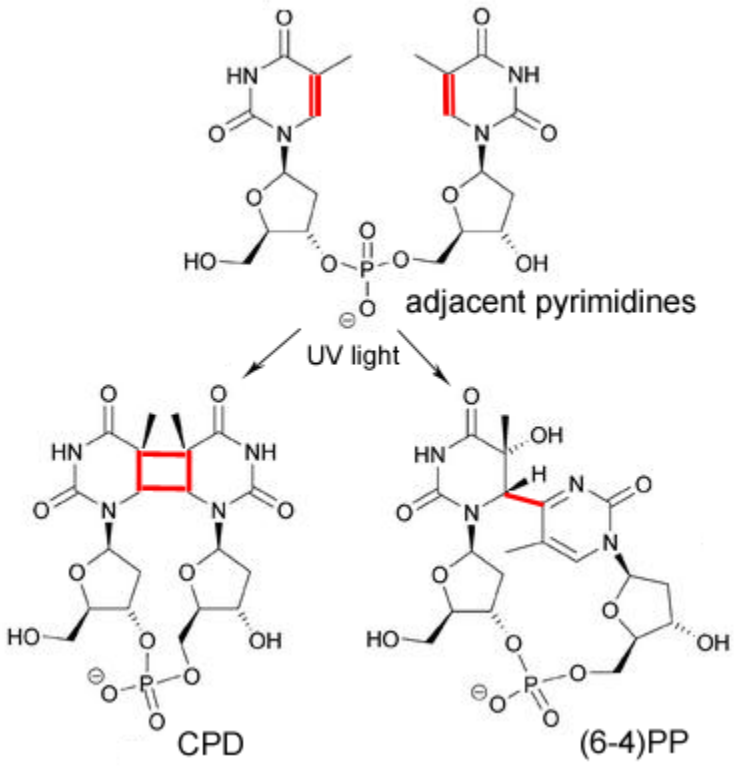


**Fig. 1.1. Depiction of the main pyrimidine–pyrimidine photolesions: cyclobutane pyrimidine dimer(CPD) and (6-4) photoproduct ((6-4)PP) (Heil *et al.*, 2011)**.

## 1.2 Repair of UV-damaged DNA

Several biological systems have been developed to repair or replace the harmful DNA lesions. Two most important mechanisms are nucleotide excision repair (NER) and photolyase-induced photoreactivation (Sancar, 1996; Sancar, 2003; Sinha & Hader, 2002).

### 1.2.1 Nucleotide excision repair (NER)

Nucleotide excision repair (NER) is a versatile and flexible DNA repair mechanism which is conserved from prokaryotes to eukaryotes (Hoeijmakers, 2001; Morita *et al.*, 2010). This mechanism can repair a broad range of structurally unrelated DNA lesions. The most relevant lesions subject to NER are UV-induced CPDs and (6-4)PPs. In addition, numerous other helix-disrupting (or "bulky") lesions are eliminated by this process such as benzo[a]pyrene-guanine adducts caused by smoking and guanine-cisplatin adducts formed during cancer chemotherapy (Sancar, 1994).

NER is initiated by the DNA lesion recognition and binding (Fig. 1.2). The DNA helix is subsequently unwound at the DNA lesion site, followed by the excision of a 25~30 oligonucleotides segment with lesion. Finally, the gap is filled by DNA repair synthesis and ligation of the remaining nick by ligase (Fuss & Cooper, 2006).

Depending on where the repair occurs, NER is divided into two subpathways. The NER occurring in DNA that undergoes transcription is called transcription-coupled repair (TCR), while the NER in non-transcribed parts of the genome, including the non-transcribed strand of transcribed genes, is called global genome repair (GGR). The two subpathways differ in the recognition of DNA damage but share the remaining repair process (Lehmann, 1995; Lindahl & Wood, 1999). Due to lacking of the photolyase-induced photoreactivation, NER plays a particular important role in the UV damaged DNA repair of placental mammals (Costa *et al.*, 2003).

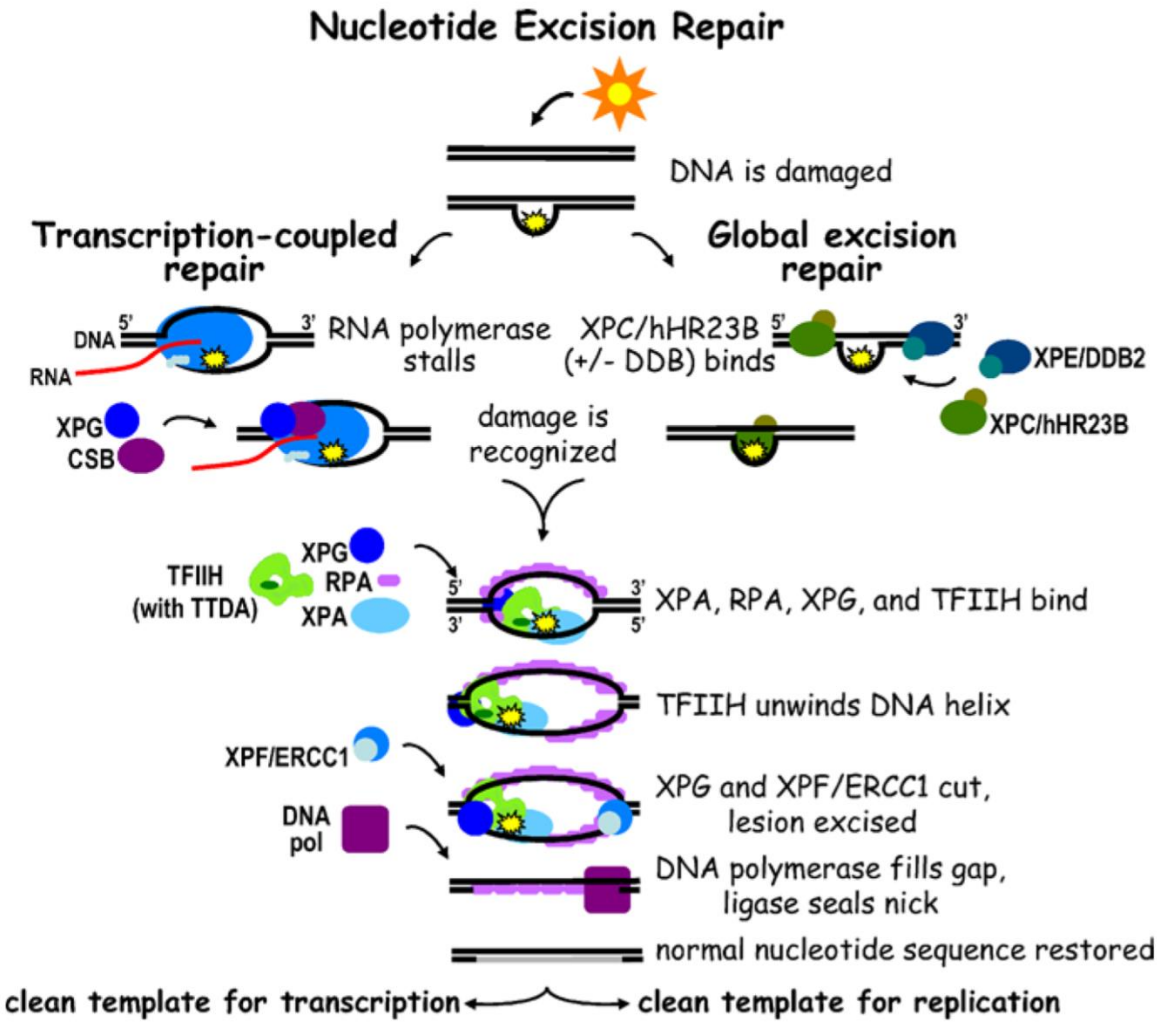


**Fig. 1.2. Nucleotide Excision Repair Schematic (Fuss & Cooper, 2006).**

### 1.2.2 Photolyase-induced photoreactivation

To defend against the deleterious effects of UV-induced mutagenic lesions, organisms have evolved photoreactivation as a highly specific and effective repair mechanism. This is believed to be one of the oldest and simplest DNA repair systems in nature (Eisen & Hanawalt, 1999).

The photoreactivation was independently discovered in two American laboratories. In 1949, Albert Kelner reported that UV-irradiated cells of *Streptomyces griseus*, *Escherichia coli*, *Penicillium notatum*, and *Saccharomyces cerevisiae* showed a 100–400,000-fold increase in survival if they were subsequently exposed to visible light. This phenomenon was termed photoreactivation (Kelner, 1949). Shortly thereafter, Renato Dulbecco reported an analogous phenomenon on the viability of UV-irradiated bacteriophage. However, it could only be detected when phage and sensitive *E. coli* were incubated together during the exposure to visible light (Dulbecco, 1949). With the discovery of DNA as the material that phage injected into cells (Hershey & Chase, 1952), Dulbecco reasoned that photoreactivation arises from the repair of UV-induced DNA photoproducts by a cellular factor through a light-dependent reaction (Dulbecco, 1950; Sancar, 2000).

In early studies cells survival was predominantly used as a biological characteristic. An in vitro photoreactivation system, DNA transformation assay, was taken into use in 1960. With this methodology, Rupert found that photoreactivation is an enzymatic process, mediated by an enzyme named photoreactivating enzyme (PRE) (Setlow & Carrier, 1966). It was later called photolyase. Rupert continued to study the repair reaction in some details and finally demonstrated that the photoreactivation follows Michaelis–Menten reaction kinetics with the notable exception that catalysis is absolutely dependent on light. Photolyase binds to UV-damaged DNA in the dark and is released from the repaired DNA upon illumination with visible light (Rupert, 1962; Sancar, 2000).

Photolyases from *E. coli* and budding yeast that were studied by Rupert, Dulbecco and others could repair only CPDs (Dulbecco, 1949; Rupert, 1962; Sancar, 2000). Forty four years after the discovery of CPD photolyase, the photolyase that repairs (6-4) photoproducts was discovered from *Drosophila melanogaster* by Takeshi Todo (Todo *et al.*, 1993).

## 1.3 DNA photolyases

DNA photolyases repair the UV-induced DNA lesions in single stranded or double stranded helical DNA through photoreactivation. Substrate specificity distinguishes two types, namely CPD photolyase that repair CPDs and (6-4) photolyase that repair (6-4) photoproducts. Although functionally distinct, both types are evolutionarily and structurally closely related (Essen & Klar, 2006; Sancar, 2003).

DNA photolyases are monomeric repair enzymes with a molecular weight of 50 to 65 kD and a length of 454 to 614 amino acids (Essen & Klar, 2006; Weber, 2005). They generally contain two noncovalently bound chromophoric cofactors. The first cofactor is a flavin adenine dinucleotide (FAD). Only the fully-reduced form of FAD ($FADH^-$) is enzymatically active. $FADH^-$ functions as the catalytic cofactor and electron donor. The second cofactor is a light-harvesting antenna chromophore, which varies depending on the protein. The presence of the antenna chromophore is not an absolute requirement for photolyases activity, since the photolyases lacking this chromophore are still biologically active. However, the antenna chromophore absorbs blue light and transfers excitation energy to the catalytic cofactor, consequently improving the DNA repair efficiency (Sancar, 2003; Selby & Sancar, 2012).

CPD photolyases have been reported in all three domains of life, while (6–4) photolyases have been identified in eukaryotes only (Muller & Carell, 2009). Photolyases seem to be absent in placental mammals including humans. Since DNA photolyases are found in a number of archaea, they are considered to be ancient repair enzymes, which were essential for organisms survival during early stages of evolution due to the intense UV irradiation passing through the anoxic atmosphere (Essen, 2006; Lucas-Lledo & Lynch, 2009).

Presently a vast number of photolyase and photolyase-like genes are known, their encoded proteins sharing a structurally highly conserved core domain and the active cofactor. These proteins together form a huge flavoprotein family, termed cryptochrome/photolyase family (CPF) (Chaves *et al.*, 2011). There are also CPF members that are incapable of repairing UV-induced DNA damage. These proteins are called cryptochromes. Cryptochromes regulate blue light responses in plants, the circadian rhythm in animals and possibly also function as magnetoreceptors in birds and insects. The first cryptochrome was found in *Arabidopsis thaliana* (Ahmad & Cashmore, 1993) and later discovered in humans (Hsu *et al.*, 1996).

### 1.3.1 Crystal structure of photolyase

In 1995, Hee-Won Park et al. presented the first crystal structure of a CPF member, the class I photolyase from *E. coli*. The globular structure is uncommon to most DNA binding proteins (Park *et al.*, 1995). It displays an architecture with an N-terminal α/β domain and a C-terminal α-helical domain. Two domains are connected via a long inter-domain loop that wraps around the α/β-domain. The α-helical domain harbors the FAD cofactor in an unusual U-shaped conformation with the adenine ring being stacked on top of the isoalloxazine ring. The α/β domain exhibits a typical Rossman nucleotide-binding fold (Rossmann *et al.*, 1974) with a five-stranded parallel β-sheet covered on both sides by α-helices (Fig. 1.3). This domain provides the binding site for the light-harvesting chromophore.

In 2004, Alexandra Mees et al. reported the co-crystal structure of a CPD photolyase from *Anacystis nidulans*, which is in complex with a DNA-duplex containing a synthetic CPD lesion analog (Mees *et al.*, 2004). The structure proved that CPD photolyase proteins completely open the DNA-duplex structure at the damaged site and flip the dinucleotide lesion out of the duplex into the catalytic center. The protein structure changes only marginally upon DNA binding.

In 2008, Melanie J. Maul et al. described the crystal structure of a (6-4) photolyase from *D. melanogaster* that bound to a DNA-duplex containing a synthetic (6-4) lesion (Maul *et al.*, 2008). The observed protein basic architecture was found to be same as that of CPD photolyases as well as the flipped-out DNA substrate binding mode. The crystal structures of cryptochromes revealed homologous overall fold topologies with photolyases. Cryptochromes have a C-terminal extension which is not however present in the structural models (Brautigam *et al.*, 2004; Brudler *et al.*, 2003; Huang *et al.*, 2006; Klar *et al.*, 2007; Muller & Carell, 2009).

Today, over a dozen crystal structures of CPF proteins are available from the protein data bank (PDB) (Table 1.1). The backbone structure of *E. coli* photolyase is preserved in the structurally characterized class I photolyases (Fujihashi *et al.*, 2007; Komori *et al.*, 2001; Tamada *et al.*, 1997), (6-4) photolyases (Maul *et al.*, 2008), plant cryptochromes (Brautigam *et al.*, 2004), DASH cryptochromes (Brudler *et al.*, 2003; Klar *et al.*, 2006) as well as Class II photolyases (Hitomi *et al.*, 2012; Kiontke *et al.*, 2011). The total rmsd (root mean square

deviation) for 13 structures with varying sequences is 2.362 Å as calculated with SSM (http://www.ebi.ac.uk/msd-srv/ssm) (Krissinel & Henrick, 2004).

All CPF proteins seems to utilize the same types of energy and electron transfer mechanism which, owing to the sensitivity to distance and orientation, force the proteins keeping a common fold to ensure an optimal arrangement of the essential cofactors and key amino acids (Epple & Carell, 1998; Epple & Carell, 1999; Muller & Carell, 2009). However, structural specificities are found for representatives of the different classes.

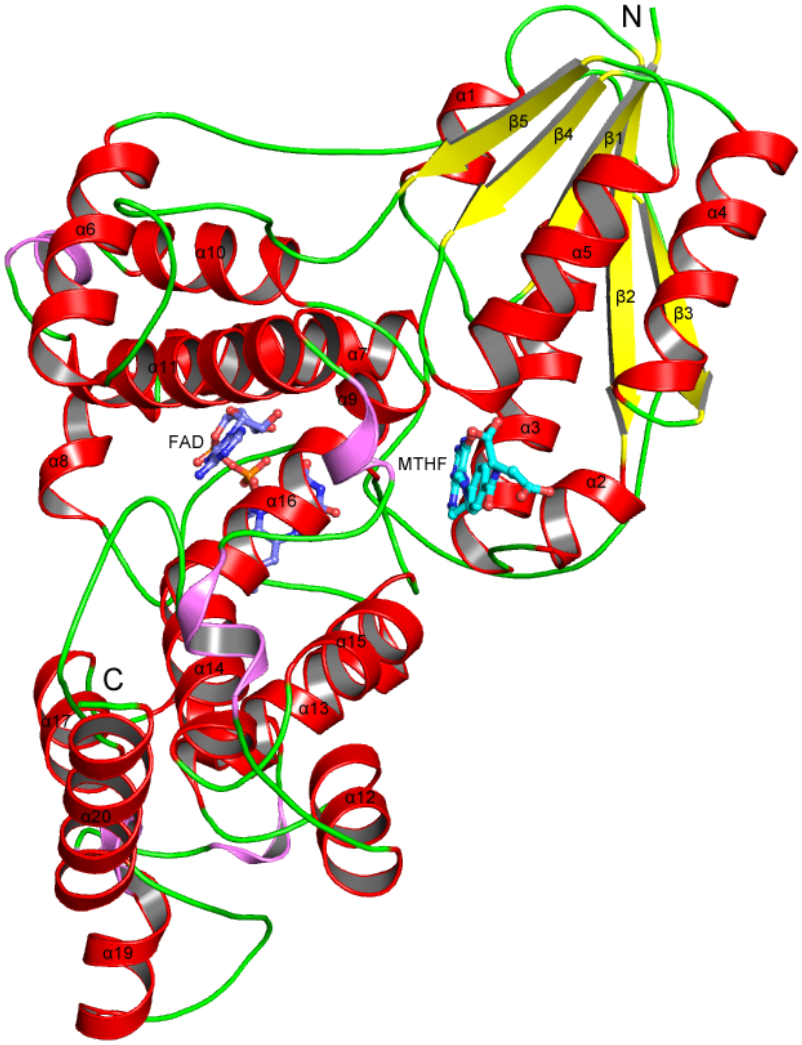


**Fig. 1.3. Overall structure of *E. coli* photolyase (PDB entry 1DNP).** Schematic drawing: yellow, β-strands; red, α-helixes; pink, 3(10) helices; green, no secondary structure; blue, FAD; cyan, MTHF (Park *et al.*, 1995)**.**

**Table 1.1. A list of all the photolyases and cryptochromes structures available from Protein Data Bank (PBD)** (Queried on February 4, 2013 ).

| Protein | Organism | PDB entry | Antenna chromophore | Substrate |
|---|---|---|---|---|
| CPD-PL I | *Escherichia coli* | 1DNP | MTHF | None |
| CPD- PL I | *Anacystis nidulans* | 1QNF<br>1TEZ<br>1OWL, 1OWM, 1OWN, 1OWO, 1OWP | 8-HDF<br>8-HDF<br>None | None<br>T<>T<br>None |
| CPD- PL I | *Thermus thermophilus* | 1IQR, 1IQU<br>2J07, 2J08, 2J09 | None<br>None<br>8-HDF, none | Thymine<br>None |
| CPD- PL I | *Sulfolobus tokodaii* | 2E0I | None | None |
| CPD- PL II | *Methanosarcina mazei* | 2XRY<br>2XRZ | None<br>None | None<br>T<>T |
| CPD- PL II | *Oryza sativa Japonica Group* | 3UMV | None | None |
| (6-4)-PL | *Drosophila melanogaster* | 3CVU, 3CVY<br>3CVV, 3CVW, 3CVX<br>2WQ6<br>2WQ7 | None<br>8-HDF, none<br>None<br>None | T(6-4)T<br>T(6-4)T<br>T(Dewar)C<br>T(6-4)C |
| (6-4)-PL | *Arabidopsis thaliana* | 3FY4 | None， | None， |
| Cry1 | *Arabidopsis thaliana* | 1U3C, 1U3D | None | AMP-PNP |
| Cry-DASH | *Synechocystis spp.* | 1NP7 | None | None |
| Cry3 | *Arabidopsis thaliana* | 2IJG<br>2J4D<br>2VTB | MTHF<br>MTHF<br>MTHF | None<br>None<br>T<>T |
| Cry1 | *Drosophila melanogaster* | 4GU5 | None | None |
| CryB | *Rhodobacter sphaeroides* | 3ZXS | DLZ<br>SF4 | None |

### 1.3.2 Reaction mechanism of photolyases

Based on previous studies, such as crystallography and ultrafast spectroscopy, the major part of the DNA photolyase repair mechanism has been resolved both structurally and dynamically (Li *et al.*, 2010; Liu *et al.*, 2011; Maul *et al.*, 2008; Mees *et al.*, 2004). The light-harvesting chromophore of photolyases absorbs UV-visible light (blue/UV light) to excite the fully-reduced $FADH^-$ from where an electron is transferred to either the CPD or the (6-4) lesion which drives the dimer into the monomers. The electron finally returns back to restore the active state of $FADH^-$ (Stuchebrukhov, 2011).

For CPD lesions (Fig. 1.4A), the single electron reduced cyclobutane ring undergoes a thermally forbidden (Woodward–Hoffmann) [2+2] cycloreversion reaction followed by back transfer of the 'enabling electron' to the semireduced $FADH^{\bullet}$ (Muller & Carell, 2009). The initial electron transfer to the dimer takes 250 ps, the ring opening occurs in two steps, the first in less than a few picoseconds and the second with a time constant of 90 ps. After the dimer is split, the electron returns back to FADH with a time constant of about 700 ps (Liu *et al.*, 2011).

In (6-4) photolyases (Figure 1.4B), upon excitation the FAD donates an electron to the (6-4) PPs to generate a charge-separated radical pair ($FADH^{\bullet}$+(6-4) $PP^{\bullet-}$), which subsequently induces proton transfer from an essential His residue of the photolyase to the (6-4) PPs. These successive steps naturally proceed to an intramolecular proton transfer from the –OH group on the C5 of the 5' base to the N3 at the 3' base to form a transient zwitterion. Then, the oxygen atom attacks the C4 position at the 3' base to form a transient oxetane-type structure. The transient oxetane formation facilitates the oxygen-atom transfer from the 5' to the 3' base followed by the splitting of the C6-C4 bond. After oxygen transfer and C-C bond cleavage the proton returns to the His residue and the electron returns to $FADH^{\bullet}$ to restore the active form of the enzyme and the two pyrimidine bases (Li *et al.*, 2010).

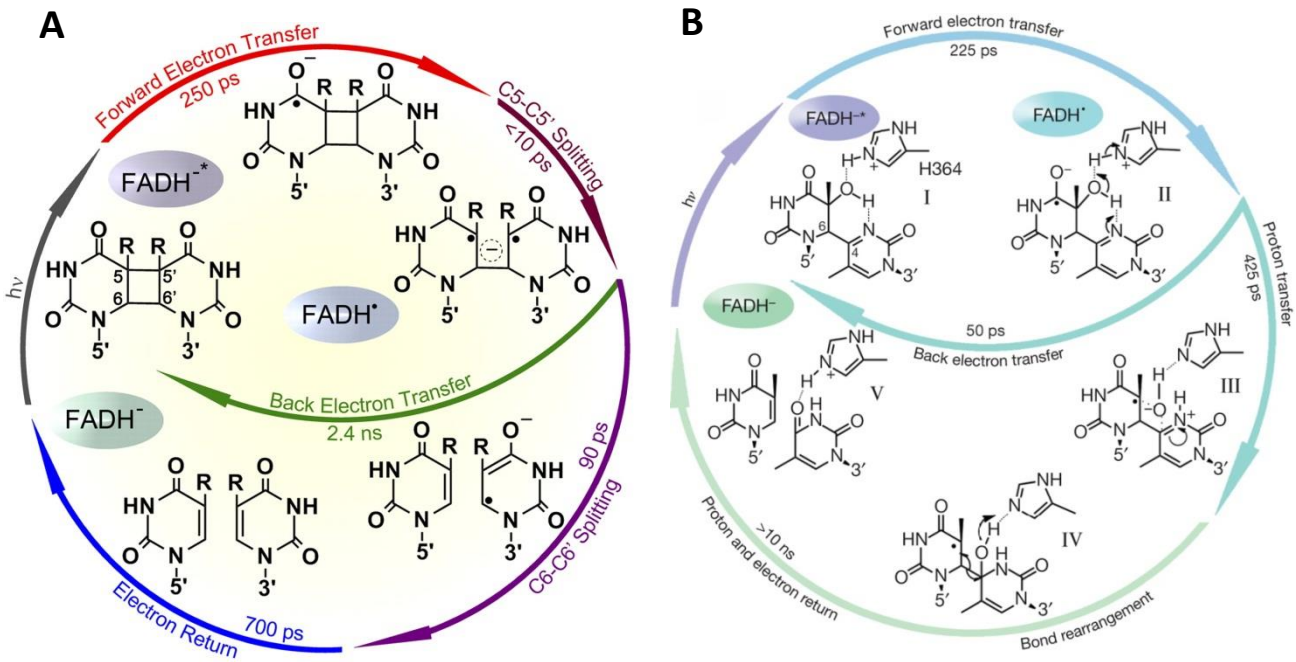


**Fig. 1.4. Complete photocycle of CPD (*A*) (Liu *et al.*, 2011) and (6-4)photoproducts (*B*) (Li *et al.*, 2010) repair by DNA photolyases with all time-resolved elementary steps of the elucidated molecular mechanism.**

### 1.3.3 Photoactivation of FAD

Flavin bound to proteins may exist in any one of the three redox states in five different forms: oxidized (e.g. FAD), semireduced semiquinones (e.g. neutral radical $FADH^{\bullet}$, anion radicals $FAD^{\bullet-}$) or fully reduced hydroquinones (e.g. $FADH^{-}$ or $FADH^{2}$) (Fig. 1.5) (Kao *et al.*, 2008b). Because of the different spectral properties (Kao *et al.*, 2008a), the flavin redox states of flavoproteins can be analyzed in vitro by monitoring protein absorption spectra (Liu *et al.*, 2010).

In photolyase, the catalytically active form of flavin is $FADH^{-}$, which is the predominant form in vivo. During the purification of photolyases under aerobic conditions the flavin cofactor is usually oxidized to the semireduced and eventually to the fully oxidized form (Payne *et al.*, 1987). Exposure of such inactive photolyases to light in the presence of a reducing agent, such as dithiothreitol (DTT), ethylenediaminetetraacetic acid (EDTA) or β-mercaptoethanol (BME), can convert the photoexcited flavin to the active $FADH^{-}$ form

(Heelis & Sancar, 1986; Sancar *et al.*, 1987). This functionally essential photoreduction process is termed photoactivation (Sancar, 2003).

Photoactivation in vitro requires photoexcitation of the flavin and an electron transfer pathway. The photoexcitation energy for flavin is obtained either via direct photon capture from blue light by itself or via energy transfer from the antenna chromophore (Moldt *et al.*, 2009; Partch & Sancar, 2005). Typically, the electron transfer pathway consists of three tryptophans (Trp-triad) in the protein moiety that is located within the C-terminal α-helical domain. The Trp-triad is conserved in most of the DNA photolyases subfamilies except class II photolyases, in which an alternative Trp-triad was identified at different position (Fig. 1.6) (Kiontke *et al.*, 2011; Maul *et al.*, 2008; Park *et al.*, 1995), and FeS-BCPs. The photoexited FAD extracts an electron from the proximal Trp in the active center of enzyme. The resulting Trp• radical subsequently abstracts an electron from the middle Trp leaving behind a Trp• radical which gets deactivated by obtaining an electron from the solvent-exposed distal Trp (Moldt *et al.*, 2009).

Whether the same mechanism is used by photolyase to photoreduce FAD *in vivo* is not clear (Kavakli & Sancar, 2004). However, point mutations of the Trp-triad that block the photoinduced electron transfer do not disturb photolyase enzymatic activity in vivo (Berndt *et al.*, 2007), supporting that the photoactivation are marginally responsible for physiological significance (Moldt *et al.*, 2009).

In vitro photoreduction of the flavin cofactor was also well studied in Cry1 from *A. thaliana*. Unlike the *E. coli* photolyase in which FAD is photoreduced to the fully reduced form, photoreduction results in the accumulation of semireduced FAD* in Arath-Cry1 (Lin *et al.*, 1995). The Trp-triad that function as electron transfer chain for FAD reduction in *E. coli* photolyase is also conserved in Arath-Cry1 as revealed by the crystal structure (Fig. 1.6) (Brautigam *et al.*, 2004). Interestingly, an involvement of tyrosine as the terminal intrinsic electron donor was observed earlier through transient absorption spectroscopy (Giovani *et al.*, 2003), but the exact tyrosine residue has not been identified so far.

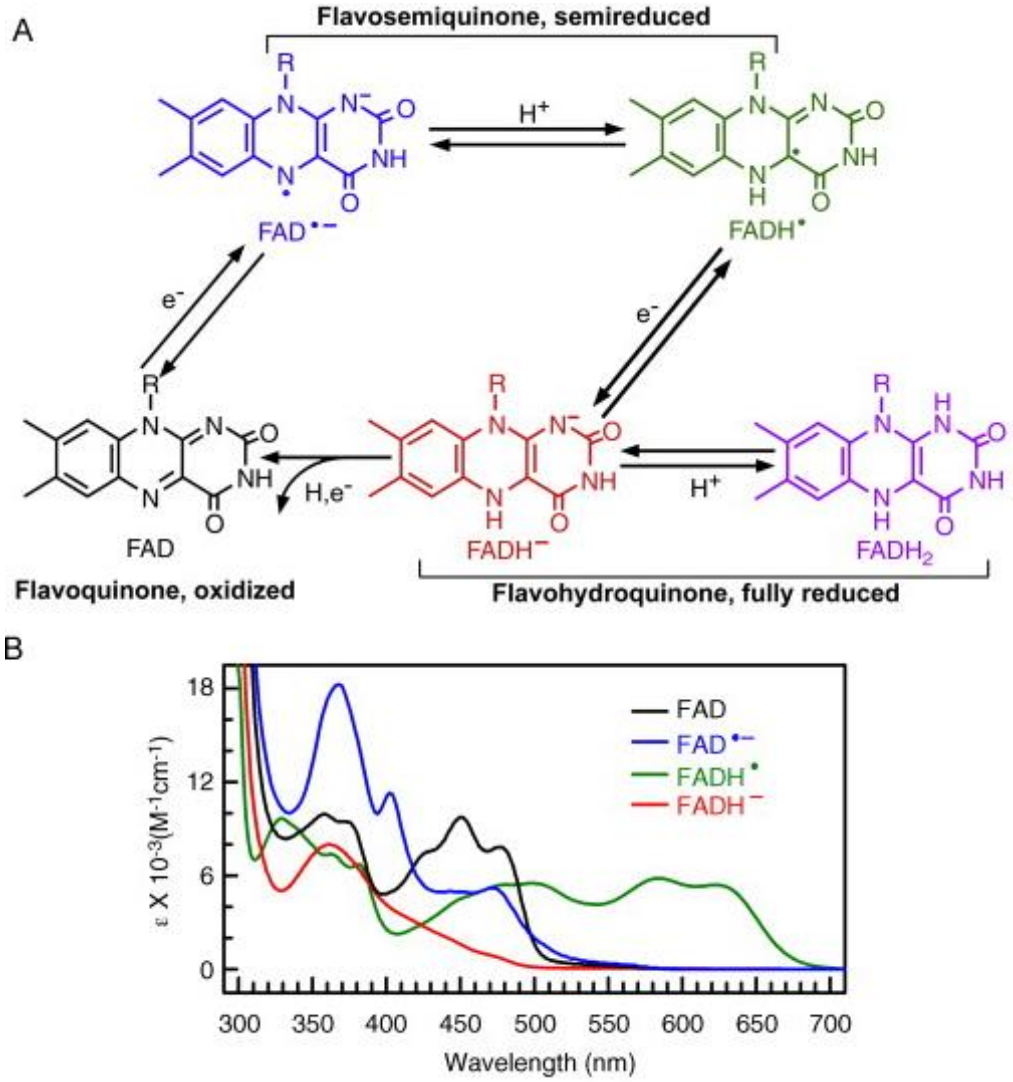


**Fig. 1.5. Oxidoreduction of flavins.** (*A*) Five possible redox forms of flavins are shown. R indicates different side groups in different flavins. The two different forms of semiquinone radicals: anion radical (e.g. $FAD^{\bullet-}$) and neutral radical (e.g. $FADH^{\bullet}$), and two forms of reduced flavins: protonated hydroquinone (e.g. $FADH_2$) and anionic hydroquinone (e.g. $FADH^{-}$) are shown. (*B*) Absorption spectra and extinction coefficients ($\varepsilon$) of different redox forms of selected cryptochrome/photolyase. Mosquito AgCRY1 (*Anopheles gambiae*) containing oxidized FAD (black line) or anion radical semiquinone ($FAD^{\bullet-}$, blue line), and *E. coli* photolyase containing neutral radical semiquinone ($FADH^{\bullet}$, green line) or fully reduced flavin ($FADH^{-}$, red line) (Liu *et al.*, 2010).

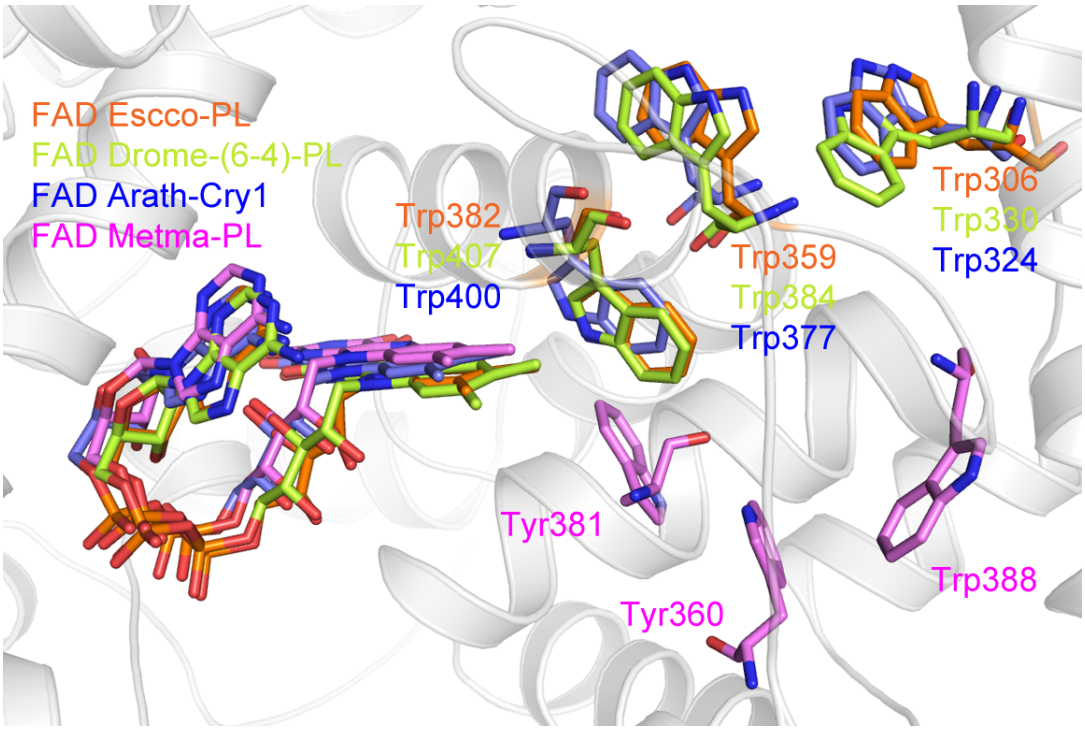


**Fig. 1.6. The electron transfer chain that reduces photoexcited $FAD_{ox}$ and $FADH^{\bullet}$ via a conserved tryptophan (W) triad (Trp-triad).** The Trp-triads of and FAD of *E. coli* CPD I photolyase (1DNP), *D. melanogaster* (6-4) photolyase (3CVV), *Methanosarcina mazei* CPD II photolyase (2XRY) and *A. thaliana* Cry1 (1U3C) are shown in sticks. *E. coli* photolyase is illustrated in ribbon representation.

### 1.3.4 Antenna chromophore

The second chromophore in photolyases serves as photoantenna by absorbing light and transferring the excitation energy to the catalytic cofactor. Although the antenna chromophores are not essential for enzyme function, they increase the absorption efficiency and broaden the activity spectra of photolyases. So far, 10-methenyltetrahydrofolate (MTHF, $\lambda_{max}$≈380 nm) (Johnson *et al.*, 1988) and several nucleotide-like compounds such as 8-hydroxy-5-deazariboflavin (8-HDF, F0, $\lambda_{max}$≈445 nm) (Eker *et al.*, 1990), flavin mononucleotide (FMN, $\lambda_{max}$≈446 nm) (Ueda *et al.*, 2005) and FAD (Fujihashi *et al.*, 2007) have been described as antenna chromophores in CPD photolyases (Fig. 1.7), while in (6-4) photolyase, only F0 has been described (Selby & Sancar, 2012). Plant cryptochromes have also been reported to contain an MTHF antenna chromophore, but the evidences are not definitive (Lin *et al.*, 1995; Malhotra *et al.*, 1995; Selby & Sancar, 2012).

The energy transfer from the antenna chromophores to FAD occurs via Förster resonance energy transfer. The efficiencies and rate constants depend on spectral properties, the orientations of each chromophore and the distances between both (Sancar, 2003). The close distances and the angles between the transition dipole moments of the two cofactors are favorable for efficient energy transfer. The energy transfer efficiencies are about 70% in *E. coli* photolyase (Payne & Sancar, 1990), 100% in *A. nidulans* photolyase (Kim *et al.*, 1992), and 78% (dark-adapted) and 87% (light-adapted) in *A. thaliana* cry3 (Song *et al.*, 2006). Crystal structure studies allow elucidating the detailed mechanism of the energy transfer between the antenna chromophore and FAD in photolyase.

All nucleotide-like chromophores are buried inside analogous clefts in the interior of the N-terminal subdomain and are approximately 17~18 Å away from FAD. In contrast, MTHF occupies a distinct shallow cleft between the N- and C-terminal subdomains and partially reaches out of the surface of the enzyme. The distance between MTHF and FAD is 15~17 Å, which is comparable with that of the nucleotide-like chromophores. Most of the amino acid residues that interact with nucleotide-like antennas are identical. By contrast, the binding sites for MTHF are less conserved. Only one of the 12 MTHF binding residues is conserved between *A. thaliana* Cry-DASH and *E. coli* photolyase (Chaves *et al.*, 2011; Sancar, 2003).

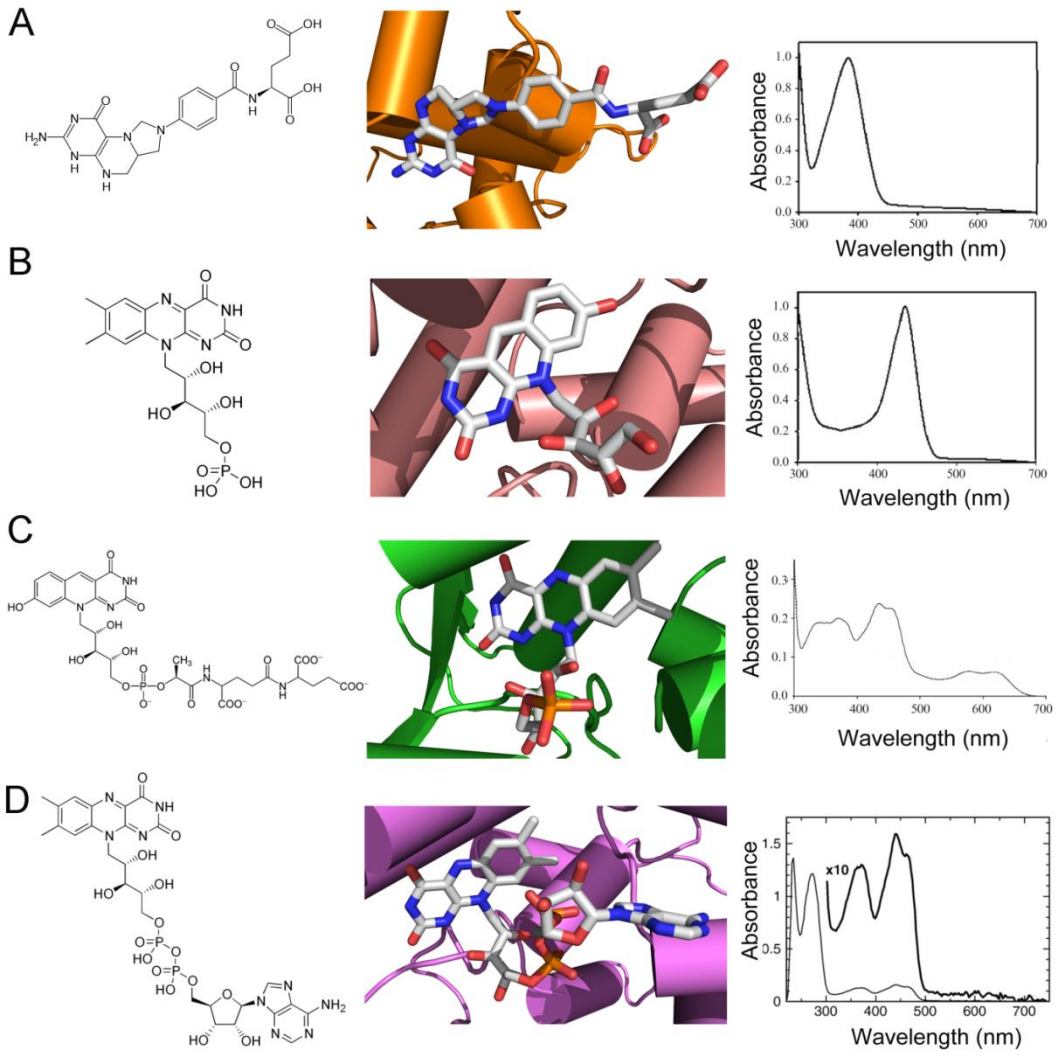


**Fig. 1.7. Antenna chromophores of photolyases.** The left drawings show the chemical structures of the antenna chromophores, the panels in middle show the 3D structure and the right panel shows the UV-Vis spectra of the protein, usually with FAD in the reduced $FADH^-$ form. (*A*) *E. coli* photolyase (PDB entry 1DNP) with 10-methenyltetrahydrofolate (MTHF) as antenna chromophore (Park *et al.*, 1995) (Sancar & Sancar, 2006). (*B*) *A. nidulans* photolyase with 8-hydroxy-5-deazaflavin (8-HDF, F0) as antenna chromophore (PDB entry 1QNF) (Fujihashi *et al.*, 2007) (Sancar & Sancar, 2006). (*C*) *T. thermophilus* photolyase with flavin mononucleotide (FMN) as antenna chromophore. (PDB entry 2E0I). The UV-Vis spectrum is taken after purification (FADH and FMN) (Klar *et al.*, 2006). (*D*) *S. tokodaii* photolyase with FAD as antenna chromophore (PDB entry 2E0I). The UV-Vis spectrum is taken immediately after purification (FAD and FAD) (Fujihashi *et al.*, 2007).

### 1.3.5 DNA binding

Ten to twenty photolyase molecules are believed to scan the genome for UV-lesions in every cell. They are structure-specific enzymes that recognize and bind UV-lesions with extremely high efficiency. The breakthrough in the understanding of the photolyase-DNA binding mechanism occurred in 1995, when Park and co-workers determined the first photolyase crystal structure from *E. coli* photolyase (Park *et al.*, 1995). It led to a proposal that the pyrimidine dimer is flipped out of the DNA double helix, stretching deeply into the binding cavity, and in forming a hydrogen-bonded contact between the carbonyl oxygens of the dimer and the adenine moiety of FAD.

The flipping-out hypothesis was subsequently supported and refined by biochemical data (Christine *et al.*, 2002; Vande Berg & Sancar, 1998), computer modeling (Antony *et al.*, 2000; Sanders & Wiest, 1999) and nuclear magnetic resonance (NMR) spectroscopy (Torizawa *et al.*, 2004). In 2004, the co-crystal structure of photolyase-DNA complex of the class I photolyase from *A. nidulans* was determined by Mees and co-workers. The flipping-out model was confirmed in more details, demonstrating that such a sequence-independent recognition of UV-lesions relies on salt bridges and hydrogen bonds mainly formed between the enzyme and the phosphates of the DNA strand comprising the lesion (Fig 1.8 A and B) (Mees *et al.*, 2004). The crystal structure of (6-4) photolyase from *D. melanogaster* with UV-damaged duplex DNA determined thereafter demonstrated the preservation of this DNA binding mechanism between CPD photolyases and (6-4) photolyase (Fig 1.8 C and D) (Maul *et al.*, 2008). An analogous binding mode was also revealed by the co-crystal structure enzyme-substrate complex of CPD II photolyase from *Methanosarcina mazei* (Fig 1.8 E and F) (Kiontke *et al.*, 2011).

Through the comparison of a dozen of photolyase and cryptochrome structures available from PDB, it is also clear that all photolyases share a positively charged DNA binding surface near the FAD accessing pocket which promotes interaction with DNA (Fig 1.8 B, D and F and Fig. 1.9), although the surface charge alone is not a sufficient determinant for DNA binding (Huang *et al.*, 2006).

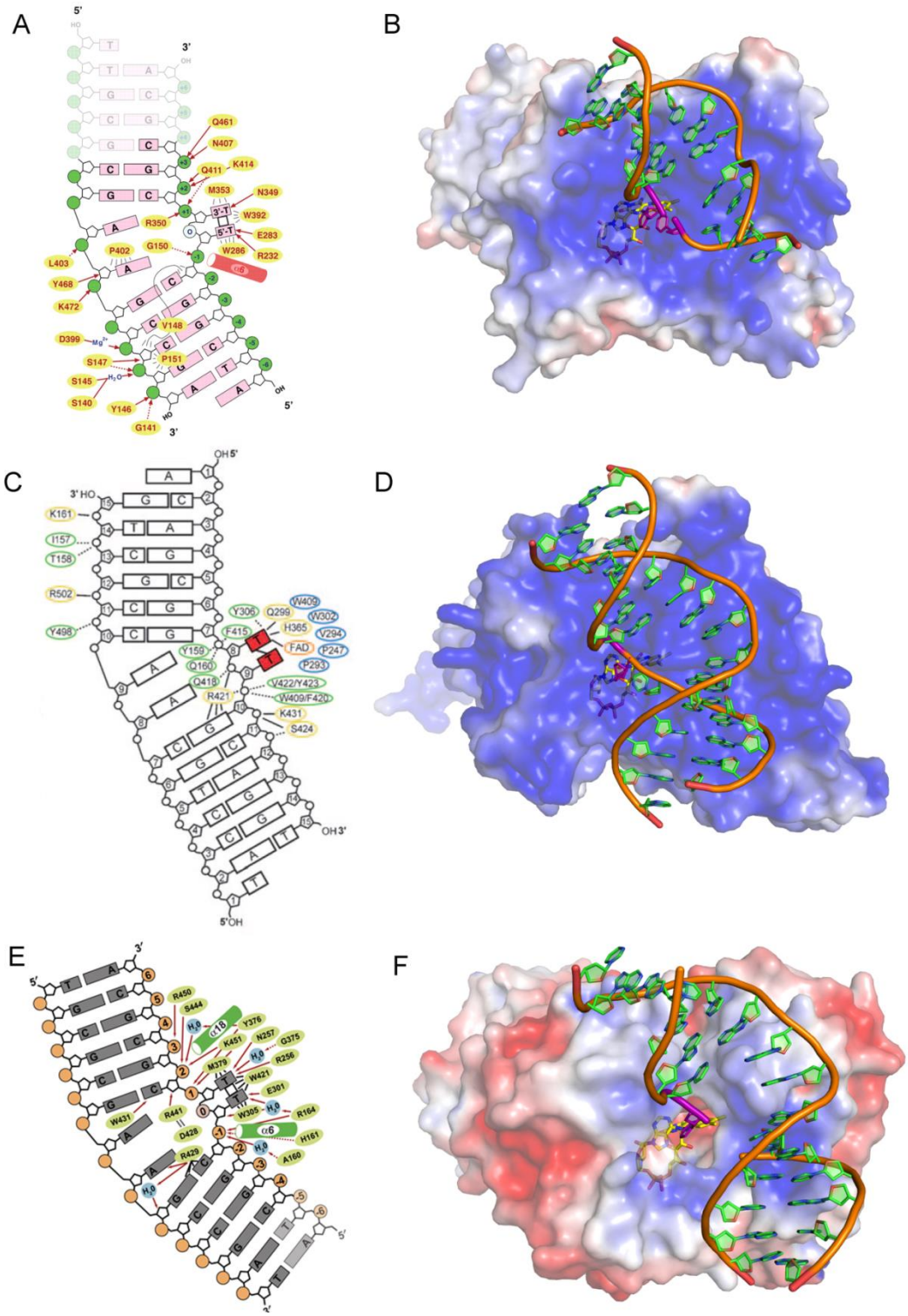


**Fig. 1.8. DNA substrate binding and surface charge representation of photolyase-DNA complex.** Diagrams on the left side are directly adapted from the original articles that describe the corresponding crystal structures. Electrostatic surface potentials were

calculated using the program APBS with the non-linear Poisson-Boltzmann equation and contoured at +/- 5kT/e. Negatively and positively charged surface areas are coloured in red and blue, respectively. The surfaces are shown in transparency mode. FAD is shown in yellow sticks. The DNA duplex is shown in cartoon ring mode, and DNA lesions are shown in purple sticks. (*A*) Diagram showing the interactions between duplex DNA and *A. nidulans* class I photolyase *(Anani-PL)*. Nucleotides not defined by electron density are shown faded. Dashed and arrows indicate interactions with the protein backbone and with side chains, respectively. Numbering of the phosphate groups starts from the intradimer formacetal group (0) (Mees *et al.*, 2004). (*B*) Surface charge representation of *Anani-PL* in complex with the CPD lesion containing duplex DNA (PDB entry 1TEZ). (*C*) Diagram of the protein-DNA contacts of *D. melanogaster* (6-4 ) photolyase (Drome(6-4)-PL) The flipped-out lesion is depicted in red, direct hydrogen-bond/electrostatic interactions are rendered in yellow, and water-mediated contacts in green and dashed lines. Residues forming the hydrophobic pocket around the pyrimidine substructure are shown in blue (Maul *et al.*, 2008). (*D*) Surface charge representation of Drome(6-4)-PL in complex to the (6-4) lesion containing duplex DNA (PDB entry 3CVV). (*E*) The diagram illustrates that the *M. mazei* class II photolyase (Metma-PL) is almost exclusively interacting with the CPD-comprising DNA strand. Dashed arrows indicate interactions with the protein backbone, whereas solid arrows characterize interactions with side chains. Black lines represent other mentioned interactions. Faded nucleotides are not defined by electron density (Kiontke *et al.*, 2011). (*F*) Surface charge representation of Metma-PL in complex to the CPD-comprising DNA (PDB entry 2XRZ).

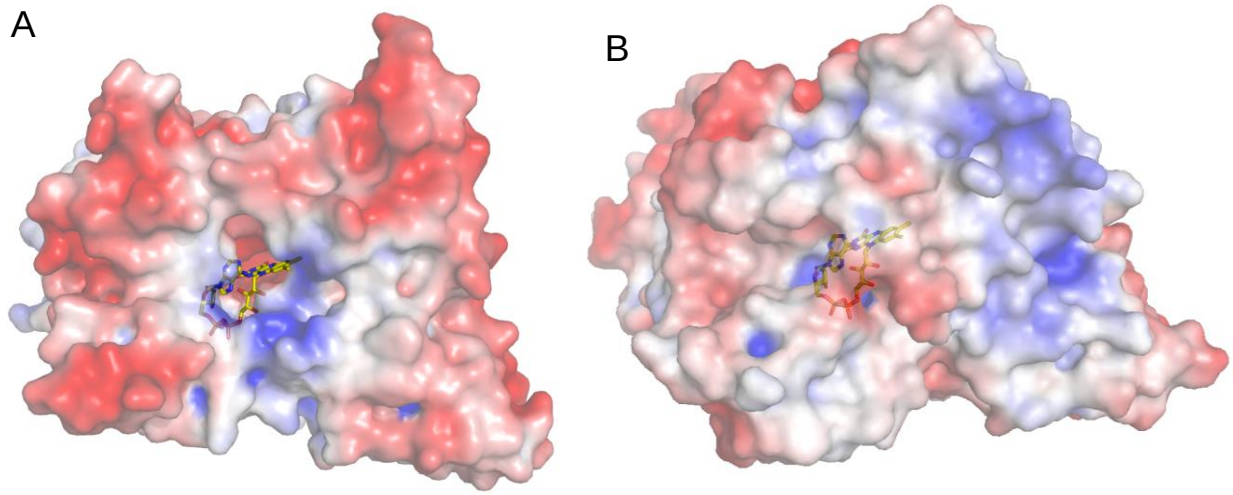


**Fig. 1.9. Surface charge representation of cryptochrome structures available from protein data bank (PDB).** Electrostatic surface potentials were calculated using the program APBS with the non-linear Poisson-Boltzmann equation and contoured at +/- 5kT/e. Negatively and positively charged surface areas are coloured in red and blue, respectively. (*A*) The surface charge representation of *A. thaliana* cryptochrome 1 (Arath-Cry1, PDB entry 1U3C). (*B*) The surface charge representation of *D. melanogaster* cryptochrome (Drome-Cry, PDB entry 4GU5).

## 1.4 Classification and evolution of the cryptochrome/photolyase family (CPF)

Members of the CPF are widely distributed in all three domains of life. Functionally, they are divided into three major classes: CPD photolyases, (6-4) photolyases and cryptochromes. Phylogenetically, the CPF is divided into the class I to class III CPD photolyases, plant CRYs, DASH CRYs, as wells as (6-4) photolyases and animal CRYs, respectively (Bayram *et al.*, 2010; Heijde *et al.*, 2010; Lin & Todo, 2005; Ozturk *et al.*, 2008). Our latest phylogenetic analysis designated a new class Fe-S bacterial cryptochromes and photolyases (FeS-BCPs) (Fig. 1.10) (Oberpichler *et al.*, 2011).

Class I and Class III CPD photolyases are found in microbial organisms, whereas class II CPD photolyases are mostly identified in higher organisms including animals and plants, but also in archaebacteria, eubacteria and single-celled algae (Okafuji *et al.*, 2010). Despite same substrate binding specificities, Class I and Class III CPD photolyases show lower similarity to class II CPD photolyases than to (6–4) photolyases (Okafuji *et al.*, 2010; Todo *et al.*, 1996; Todo, 1999). The (6-4) photolyases have only been observed in eukaryotes. Plant CRYs form a sister group of the class III CPD photolyases, while animal CRYs are closely related to (6-4) photolyases (Fig. 1.10) (Oberpichler *et al.*, 2011; Weber, 2005).

Today, it is widely accepted that the common ancestor of the CPF was likely a CPD-photolyase (Essen, 2006) and the plant cryptochromes have evolved from the Class III CPD photolyases while the animal cryptochromes are derived from (6-4) photolyases (Muller & Carell, 2009).

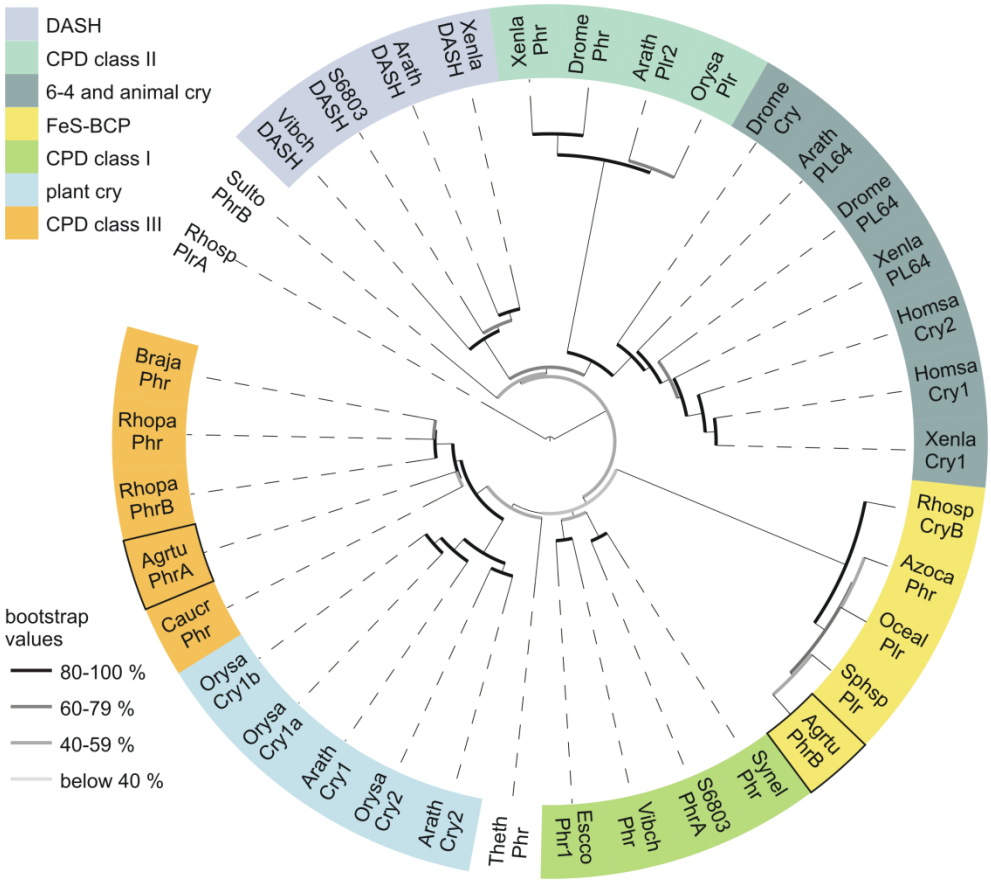


**Fig. 1.10. Phylogenetic tree of the cryptochrome/photolyase family.** The species names are abbreviated as follows: (*A. thaliana* (Arath), *A. tumefaciens* (Agrtu), *Azorhizobium caulinodans* ORS 571 (Azoca), *Bradyrhizobium japonicum* USDA 110 (Braja), *Caulobacter crescentus* (Caucr), *D. melanogaster* (Drome), *E. coli* (Escco), *Homo sapiens* (Homsa), *Oceanocaulix alexandrii* (Oceal), *Oryza sativa* (Orysa), *Rhodobacter sphaeroides* 2.4.1 (Rhosp), *Rhodopseudomonas palustris* (Rhopa), *Sphingomonas sp.* SKA58 (Sphsp), *Sulfolobus tokodaii* str. 7 (Sulto), *Synechococcus elongatus* PCC 6301 (Synel), *Synechocystis sp.* PCC 6803 (S6803), *Thermus thermophilus* HB8 (Theth), *Vibrio cholerae* (Vibch), *Xenopus laevis* (Xenla), Cry: cryptochrome, DASH: DASH cryptochrome, PL64: (6-4) photolyase, Phr: photolyase, Plr: photolyase-related protein, CPD: cyclobutan pyrimidine dimer, FeS-BCP: iron-sulfur cluster containing bacterial cryptochromes and photolyases (Oberpichler *et al.*, 2011).

## 1.5 Cryptochrome/photolyase-related proteins in *A. tumefaciens*

By screening the *A. tumefaciens* genome for sequences with homology to photoreceptors we identified a cryptochrome/photolyase-related protein named PhrA (gi: 17739623) (Oberpichler *et al.*, 2008). Phylogenetic analysis assigned it as a class III CPD photolyase. The specific repair activity was recently confirmed by in vitro studies (Oberpichler *et al.*, 2011).

The second cryptochrome/photolyase-related protein in *A. tumefaciens* has been found by transposon mutagenesis in a screen for light regulation of motility. Results of the mutant studies suggested that the protein (gi: 15158416) might serve as a photoreceptor for light-regulated motility. Sequence comparisons showed that this protein is homologous to known photolyases. We hence named this protein PhrB (Oberpichler *et al.*, 2011).

## 1.6 Aims of research

The goal of this work was to investigate the function and structure of the two cryptochrome/photolyase-related proteins in *A. tumefaciens*. PhrA has been experimentally shown to repair CPD photoproducts *in vivo* and *in vitro* (Oberpichler *et al.*, 2011). Phylogenetic studies assigned it as a class III photolyase, which is a newly characterized CPF class and is the sister group of plant cryptochrome. Several questions arose from our latest studies. Like *E. coli* photolyase, PhrA contains MTHF as an antenna chromophore but the MTHF in PhrA is more stable than that in *E. coli* photolyase as indicated by UV-Vis spectroscopy. Structural modeling revealed that none of the amino acids that coordinate MTHF of *E. coli* photolyase is conserved in PhrA. Hence the first question is whether PhrA shares the MTHF binding site and mode with *E. coli* photolyase. Class III photolyases show higher similarity to plant cryptochromes than to class I photolyases (Oberpichler *et al.*, 2011). The second question is how relatively small differences determine their functional difference, which consequently leads to the third question how a class III photolyase crystal structure looks like. There is so far no crystal structure of a class III CPD member available in the PDB. To this end, we conducted protein crystallography for PhrA. The crystal structure of PhrA revealed a novel site and mode of MTHF binding as well as an alternative electron pathway that could be responsible for FAD photoreduction. In order to verify the hypothesis, a series of point mutations were constructed by site-directed mutagenesis and subsequently analyzed with UV-Vis spectroscopy.

Another major concern of this research was to investigate the function and structure of PhrB. Previous spectroscopic studies on PhrB and one of its homologs, CryB from *Rhodobacter sphaeroides*, suggested that they both incorporate same yet unknown antenna (Hendrischk *et al.*, 2009; Oberpichler *et al.*, 2011). We thus firstly wish to characterize the antenna chromophore of PhrB. PhrB contains a [4Fe-4S] cluster and belongs to a recently denominated FeS-BCP class, which is distantly related to other CPFs. Since all the known CPF members do not contain an iron-sulfur cluster, we like to find out whether this newly designated class really functions as bona fide CPF members. However, the functional clues for PhrB from former studies are somehow paradoxical. PhrB has been shown playing a photoreceptor role that is involved in the cell motility regulation in vivo. The in vitro DNA CPD photoproduct repair assay showed a negative result. Whether PhrB has (6-4) photolyase

activity was not clear because of the immature methodology in our lab. Its homolog CryB was also reported as photoreceptor that regulates transcription of photosynthesis genes in vivo and is incapable of photoreactivation in vitro (Geisselbrecht *et al.*, 2012). These clues support a cryptochrome function for PhrB. On the other hand, mutant studies have shown both PhrB and PhrA are required for full photoreactivation in *A. tumefaciens* (Oberpichler *et al.*, 2011). CryB was also shown to be required for photoreactivation in *Rhodobacter sphaeroides* (Hendrischk *et al.*, 2009). We therefore reasoned that PhrB could function as a (6-4) photolyase. To investigate the function of PhrB, an HPLC-based (6-4) PPs repair assay was conducted. Protein crystallography was further performed with the expectation of the identification and localization of antenna chromophore as well as the binding mode and functional clue for iron-sulfur cluster. Based on the crystallographic studies the antenna chromophore identity could be proposed, which was subsequently confirmed by HPLC analysis.

# 2 Crystal structure of PhrA

## 2.1 Materials and Methods

### 2.1.1 PhrA protein purification and crystallization

Recombinant His-tagged PhrA protein was expressed in *E. coli* and purified via Ni-affinity chromatography and size-exclusion chromatography as described (Oberpichler *et al.*, 2011) and used for crystallization at concentrations of 11 mg/ml under dim light conditions or in darkness. Primary screening was performed with the sitting-drop vapour diffusion approach by using 432 crystallization conditions from commercial Kits (JBScreen Classic 1–10, JBScreen Basic 1–4 and JBScreen PEG/Salt 1–4, Jena Bioscience) at 277 K and 298 K in 96-well MRC plates. Promising conditions were screened further by changing pH and the concentration of precipitants or additives. After optimization, crystals were grown by sitting drop vapour diffusion at 277 K in 24-well Linbro plates. Each drop was prepared on a siliconized glass plate by mixing 4 µl PhrA (11 mg/ml) with equal volume of reservoir solution and equilibrated against 1 ml reservoir solution. The protein solution contained 12.5 mM tris (hydroxymethyl) amino methane, 1.25 mM ethylenediaminetetraacetic acid, 2.5 % glycerol, 75 mM sodium chloride, pH 7.8. The reservoir solution contained 3 M Ammonium sulfate and 10% (w/v) glycerol.

### 2.1.2 Data collection and structure analysis (Structure determination and refinement was done by Patrick Scheerer)

Diffraction data collection was performed at 100 K using synchrotron X-ray sources at BESSY II, Berlin, Germany, and ESRF, Grenoble, France (de Sanctis *et al.*, 2012). Best diffraction data were collected at beamline BL 14.1 (Mueller *et al.*, 2012) at BESSY II, at a wavelength of $\lambda = 0.91842$ Å. All images were indexed, integrated and scaled using the XDS program package (Kabsch, 2010) and the CCP4 program SCALA (Collaborative Computational Project, 1994; Evans, 2006). The PhrA crystals belonged to the trigonal space group $P3_221$ (a = b = 81.87 Å, c = 195.95 Å, $\alpha = \beta = 90°$, $\gamma = 120°$). Table 2.1 summarizes the statistics for crystallographic data collection and structural refinement. Initial phases for PhrA were obtained by the conventional molecular replacement protocol (rotation, translation, rigid-body fitting) using the *Arabidopsis thaliana* cryptochrome 1 (Arath-Cry1) crystal structure (PDB entry 1U3C) (Brautigam *et al.*, 2004) without ligands as the initial search model. Molecular replacement was achieved using the CCP4 program PHASER (McCoy *et al.*, 2007) by placing the Arath-Cry1 monomer (rotation function (RFZ): Z = 8.9; translation function (TFZ): Z = 17.9 for PhrA; RFZ and TFZ as defined by PHASER). An AutoBuild/RESOLVE (with the correct PhrA sequence) and a simulated annealing procedure with the resulting model were performed using a slow-cooling protocol and a maximum likelihood target function, energy minimization, and B-factor refinement by the program PHENIX (Adams *et al.*, 2010), which was carried out in the resolution range of 34.30 – 1.67 Å for the PhrA crystal structure. After the first rounds of refinement, the flavin adenine dinucleotide (FAD) and 5,10-methenyltetrahydrofolate (MTHF) in the ligand binding pockets were clearly visible in the electron density of both $\sigma_A$-weighted $2F_O$-$F_C$ maps, as well as in the $\sigma_A$-weighted simulated annealing omitted density maps. PhrA was modeled with TLS refinement (Winn *et al.*, 2001) using anisotropic temperature factors for all atoms. Restrained, individual B-factors were refined, and the crystal structure was finalized by the CCP4 program REFMAC5 (Vagin *et al.*, 2004) and other programs of the CCP4 suite (Cowtan *et al.*, 2011; Emsley *et al.*, 2010). The final model has agreement factors $R_{free}$ and $R_{work}$ of 18.0% and 15.6%, respectively. Manual rebuilding of the PhrA model and electron density interpretation were performed after each refinement cycle using the program COOT (Vagin *et al.*, 2004). The side chain atoms of Lys 42, 183, 193, 430, 440, 441, 444, 452, 477

(atoms: CG, CD, CE, NZ) and Asp 441 (atoms: CG, OD1, OD2) on the protein surface were removed from the PhrA structure due to not well-defined electron density. Structure validation was performed with the programs PHENIX (Adams *et al.*, 2002), SFCHECK (Vaguine *et al.*, 1999), PROCHECK (Laskowski *et al.*, 1993), WHAT_CHECK/WHAT IF (Hooft *et al.*, 1996; Vriend, 1990) and RAMPAGE (Lovell *et al.*, 2003). Potential hydrogen bonds and van der Waals contacts were analyzed using the programs HBPLUS (McDonald & Thornton, 1994) and LigPlot+ (Laskowski & Swindells, 2011). All crystal structure superpositions of backbone α-carbon traces were performed using the CCP4 program LSQKAB (Collaborative Computational Project, 1994). All molecular graphics representations in this work were created using PyMol (DeLano, 2002).

**Table 2.1**. Data collection and refinement statistics

| | **PhrA – class III photolyase** (PDB entry 4U63) |
|---|---|
| **Data collection**[a] | BESSY II, BL14.1 |
| (wavelength) | λ = 0.91842 Å |
| Space group | P$3_221$ |
| Cell dimensions | |
| *a*, *b*, *c* (Å) | 81.87 81.87 195.95 |
| α, β, γ (°) | 90.0, 90.0, 120.0 |
| Resolution (Å) | 34.40 – 1.67 |
| | (1.76 – 1.67)[b] |
| $R_{merge}$ | 0.085 (0.932) |
| $R_{pim}$ | 0.030 (0.327) |
| $\langle I/\sigma(I)\rangle$ | 13.7 (2.3) |
| Completeness (%) | 100.0 (100.0) |
| Redundancy | 9.2 (9.0) |
| Wilson B factor (Å$^2$) | 21.5 |
| **Refinement** | |
| Resolution (Å) | 34.40 – 1.67 |
| No. Reflections | 84703 |
| $R_{work}$/ $R_{free}$ (%) | 15.6 / 18.0 |
| Overall B Factor (Å$^2$) | 30.1 |
| No. atoms / residues | 4650 / 482 |
| (1 monomer per asymmetric unit) | |
| Protein (PhrA) | 3913 / 478 |
| Expression Tag | 25 / 4 |

| | |
|---|---|
| Others: | |
| Water | 548 / 547 |
| Flavin-adenine dinucleotide (FAD) | 53 / 1 |
| 5,10-Methenyl-6,7,8-trihydrofolic acid (MTHF) | 33 / 1 |
| Sulfate ion ($SO_4$) | |
| 2-Amino-2-hydroxymethyl-propane-1,3-diol (Tris-buffer) | 70 / 14<br>8 /1 |
| R.m.s[c] deviations | |
| Bond lengths (Å) | 0.008 |
| Bond angles (°) | 1.23 |
| Ramachandran plot[d] | |
| % favoured | 97.5 |
| allowed | 2.5 |
| outlier | 0.0 |

[a] One crystal was used; [b] highest resolution shell is shown in parenthesis; [c] R.m.s, root mean square; [d] Ramachandran plot created by RAMPAGE (Lovell *et al.*, 2003) using the Richardsons' data.

### 2.1.3 Sequence analysis.

All the representative sequences for the involved sub-families of CPF in this study were selected according to the former phylogenetic study, based on which CPD III photolyase was discovered (Partch, 2006). All the selected sequences were retrieved from public databases via the National Center for Biotechnology Information web site (www.ncbi.nlm.nih.gov). Multiple sequence alignments were performed with ClustalX 2.0.12 (Thompson *et al.*, 2002) and ClustalW2 (http://www.ebi.ac.uk/Tools/clustalw2/index.html).

### 2.1.4 Site directed mutagenesis.

Site-directed mutagenesis of were carried out according to the QuikChange protocol (Stratagene, La Jolla, CA) as described by the supplier. The primers (Table 2.2) for mutation experiments were synthesised from GATC *Biotech* AG (Konstanz, Germany). The correctness of mutant sequences was confirmed by DNA sequencing. The expression and purification protocols are same as for the wild type (WT) PhrA.

### 2.1.5 UV-Vis spectroscopy analysis and protein photoreduction.

UV-Vis spectral measurements of recombinant proteins were performed with a Jasco V550 photometer. The overall spectra were measured immediately after purification. 24 h incubation in darkness at 4 °C were performed to oxidise the FAD chromophore. Thereafter, 10 mM DTT was added and photoreduction was conducted at 10 °C. The light source for photoreduction consisted of blue light emitting diodes HLMP-HB57-LMC (Avago Technologies, Boeblingen, Germany) with a maximum emission wavelength of 470 nm and a light intensity of 55 $\mu$mol $m^{-2}$ $s^{-1}$. Spectra were measured 1, 2, 3, 5, 7, 10, 15, 20, 25, 30, 35, 40, 45, 50, 55, 60, 70, 80, 90, 100, 110 and 120 min after onset of the light.

**Table 2.2 Mutants generated by site-directed mutagenesis.** Primer sequences are given in the second column; the mutated sites are underlined.

| Mutant | Primer | soluble or insoluble (14°C expression) |
|---|---|---|
| WT | | soluble |
| W384A | fw: 5′-GCGTCCAATGCCGCCAACGCGCAATGGGTGGCAGG-3′<br>rew: 5′-CCTGCCACCCATTGCGCGTTGGCGGCATTGGACGC-3′ | insoluble |
| W384F | fw: 5′-GCGTCCAATGCCGCCAACTTCCAATGGGTGGCAGG-3′<br>rew: 5′-CCTGCCACCCATTGGAAGTTGGCGGCATTGGACGC-3′ | soluble |
| W318A | fw: 5′-CCTTCAAGGCCGCGACGCGCGGCATGACGGGTTACCC-3′<br>rew: 5′-GGGTAACCCGTCATGCCGCGCGTCGCGGCCTTGAAGG-3′ | insoluble |
| W318F | fw: 5′-CCTTCAAGGCCTTCACGCGCGGCATGACGGGTTACCC-3′<br>rew: 5′-GGGTAACCCGTCATGCCGCGCGTGAAGGCCTTGAAGG-3′ | soluble |
| W361A | fw: 5′-CTGCTGATCGACGCGCGGAAGGGCGAGAAGTGG-3′<br>rew: 5′-CCACTTCTCGCCCTTCCGCGCGTCGATCAGCAG-3′ | soluble |
| W367A | fw: 5′-GGAAGGGCGAGAAGGCGTTTCGCGATACGCTCGTCGATGC-3′<br>rew: 5′-GCATCGACGAGCGTATCGCGAAACGCCTTCTCGCCCTTCC-3′ | soluble |
| W196A | fw: 5′-CCCACCAAGCCGGACGCGGCAAAGGATTTCAGCG-3′<br>rew: 5′-CGCTGAAATCCTTTGCCGCGTCCGGCTTGGTGGG-3′ | soluble |
| W336A | fw: 5′-CGGCATGCGGCAATTGGCGCAACATGGCACCATGC-3′<br>rew: 5′-GCATGGTGCCATGTTGCGCCAATTGCCGCATGCCG-3′ | soluble |
| W308A | fw: 5′-CGACAGTTTCGACGCCTTTTCCGCGCGGGACGATG-3′<br>rew:5′-CATCGTCCCGCGCGGAAAAGGCGTCGAAACTGTCG-3′ | soluble |
| W308F/ W361F | | soluble |
| W308A /W367A | | insoluble |
| W308F/W367F | | soluble |

## 2.2 Results and Discussion

### 2.2.1 Crystallization of PhrA

Primary crystallization screening for PhrA yielded 2 hits within 1 week at 277 K, while no hit found at 298 K. Under the first condition, 2 M ammonium sulfate was used as precipitant and 2 M sodium chloride as additive. This combination yielded needle-shaped yellow crystals (Fig. 2.1A) which stopped growing after 1 week (Fig. 2.1B). The second combination contains 3 M ammonium sulfate and 10% glycerol. The rectangular-like yellow crystal had obvious defects as indicated by cracks on the surfaces (Fig. 2.1C). Increasing the time of growth showed that the crystals consist of numerous bundled thin strips (Fig. 2.1D), which are not suitable for X-ray diffraction data collection.

Subsequent optimization was performed based on the clues obtained from the primary screen. Firstly, ammonium sulfate is the most promising precipitant as it is contained in both solutions where crystals grew. Secondly, sodium chloride tends to produce needle crystals since the crystals that grew without sodium chloride are obviously bigger. Thirdly, PhrA protein seems more stable at low temperature as no hit was found at 298 K. Finally, glycerol also seems good for PhrA stabilization and crystallization.

As for the first condition (2 M ammonium sulfate 2 M sodium chloride), the drop size, the concentration of precipitant, additive or protein and the pH of solution were varied. These variations either changed the size of needle crystals or led to the disappearance of crystals but no improvement for the quality of crystals shape. For the second condition (3 M ammonium sulfate and 10% glycerol), the concentration of precipitant and the pH of solution were varied. In this case, there was also no improvement of the crystal quality. Interestingly, the best shaped crystals were obtained from exactly the original conditions after increasing the drop size 2 times (Fig. 2.2). Under these conditions, the yellow rectangular PhrB crystals appeared within one week and grew to their final size of about $0.08 \times 0.1 \times 0.2$ $mm^3$ after 3 weeks. Such PhrA crystals were directly flash frozen in liquid nitrogen for subsequently X-ray diffraction data collection.

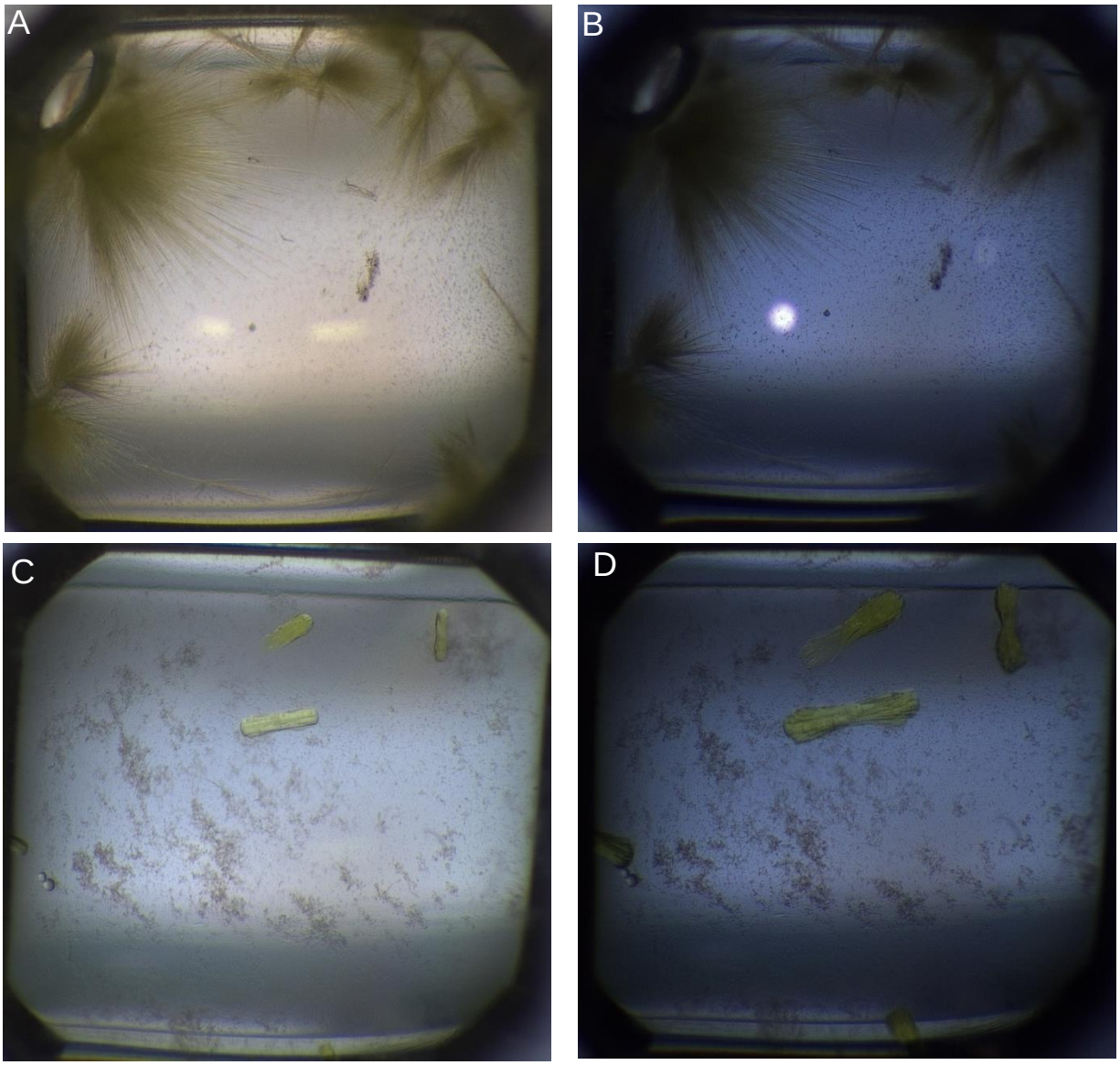


**Fig. 2.1. Crystals from 2 hits in the primary screen.** The protein concentration was always 11 mg/mL and the temperature 277 K. Drops were set after mixing 2 µl protein solution and 2 µl reservoir solution. (*A*) and (*B*) Crystals were grown in 2 M ammonium sulfate as precipitant and 2 M sodium chloride, for 1 week and 2 weeks incubation, respectively. (*C*) and (*D*) Crystals were grown in 3 M ammonium sulfate and 10% glycerol for 1 week and 4 weeks respectively.

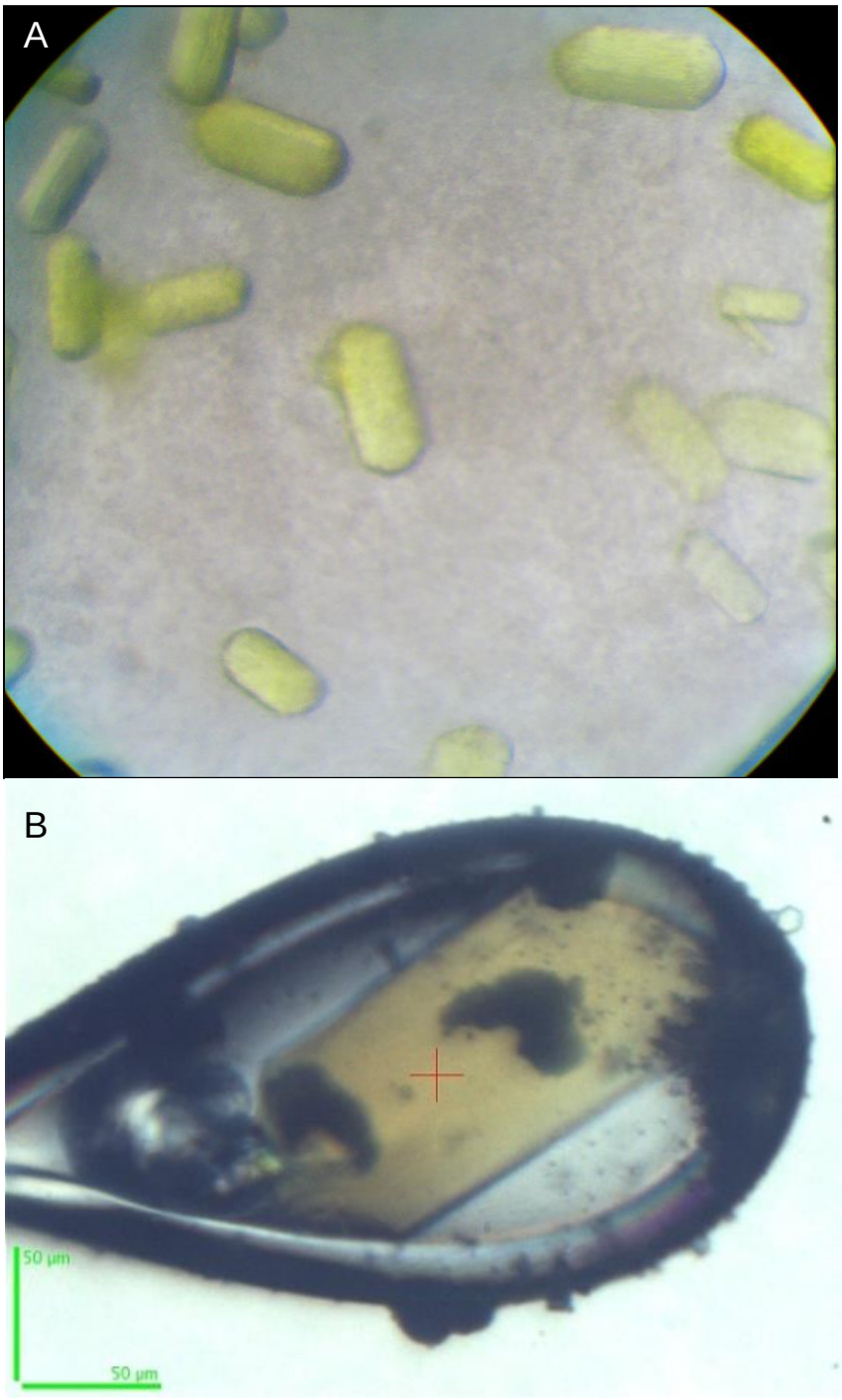


**Fig. 2.2. Crystals that grown from the 3 M ammonium sulfate and 10% glycerol condition.** Protein concentration is 11 mg/ml. Temperature is 277 K. Drops are set by mixing of 4 µl protein solution and 4 µl reservoir solution. (*A*) Crystals in the mother liquid. (*B*) Crystal mounted and frozen in a 0.3 mm loop at 100 K.

### 2.2.2 Overall structure of PhrA

The best crystals diffracted with a resolution of 1.67 Å and the structure of PhrA was solved by molecular replacement using the crystal structure of *A. thaliana* Cry1 (Arath-Cry1, PDB entry 1U3C) as initial search model. The atomic model contains 482 residues (aa 2-478 plus 5 residues of the N-terminal tag) that reveal a typical overall fold of CPF proteins (Fig. 2.3). The structure of the plant cryptochrome Arath-Cry1 matches best with PhrA (rmsd = 1.8 Å), in accordance with phylogenetic studies in which both proteins belong to sister groups. Structures with the second and third best matches are from *A. nidulans* photolyase (Anani-PL, PDB entry 1TEZ, rmsd = 1.9 Å) and *E. coli* photolyase (Escco-PL, PDB entry 1DNP, rmsd = 2.2 Å), Structures of other CPF members yielded rmsd values between 2.2 and 3.8 Å. The N-terminal α/β domain (residues 2-130) and the C-terminal helical domain (residues 208-478) are connected via a long loop of 77 amino acids (residues 131-207). The U-shaped catalytic chromophore FAD is deeply buried in the center of the protein within the helical domain, as in other CPF proteins. Structure determination showed that the antenna cofactor of PhrA is MTHF, as proposed by UV-Vis spectrometry (Oberpichler *et al.*, 2011).

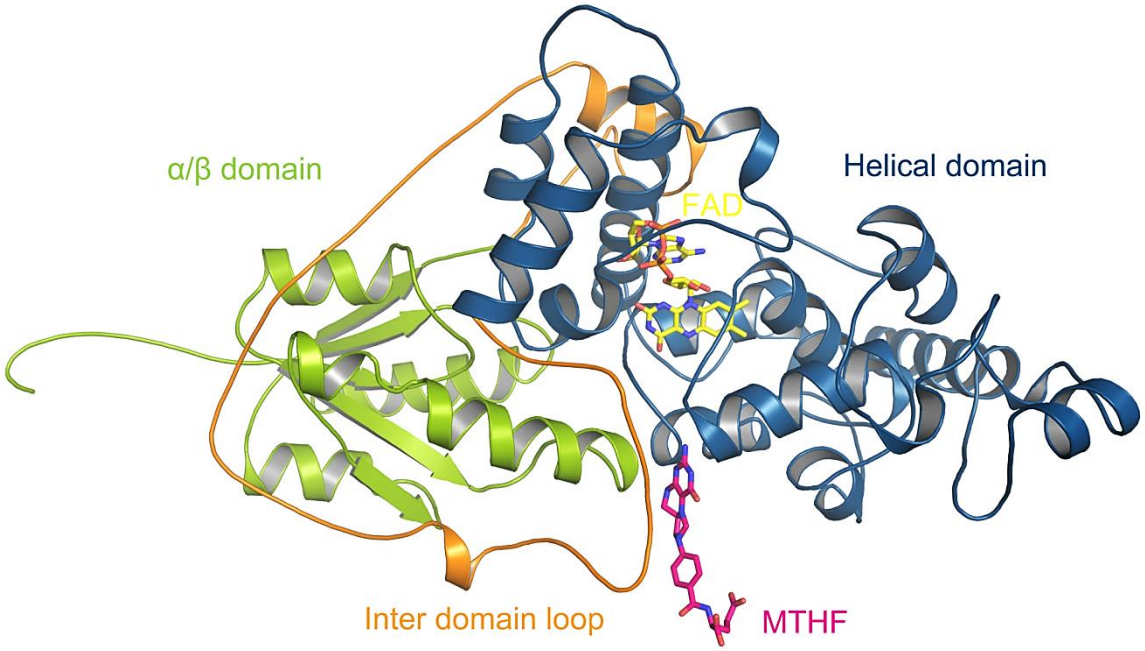


**Fig. 2.3. Overall structure and cofactor arrangement of PhrA.** The ribbon representation shows the α/β subdomain (green) and the helical subdomain (blue) connected by a long inter domain linker (orange). The cofactors MTHF (magenta) and FAD (yellow) are illustrated in ball-and-stick representation.

### 2.2.3 MTHF binding in PhrA

In Escco-PL and Arath-Cry3, MTHF is bound to the same site within a groove at the surface of the protein, around 17 Å away from FAD. In those cases, MTHF is coordinated by amino acids from helices α2, α5, and α14 (Huang *et al.*, 2006; Klar *et al.*, 2007; Park *et al.*, 1995). The PhrA crystal structure reveals a different MTHF binding site in which MTHF is 19 Å away from FAD (Fig. 2.4A**)** and located on the opposed side relative to the isoalloxazine ring of FAD. The angle that is formed between the conjugated ring system of MTHF and the isoalloxazine rings of FAD is similar to that of Escco-PL. Thus, MTHF of PhrA serves equally well as antenna chromophore for light capture and Förster energy transfer to FAD(H). The binding pocket is formed by helices α2, α7, α15, α17 and one loop that connects helix α6 and α7. Nine amino acids, Ala50, Pro194, Trp196, Phe200, Leu 335, Trp336, Gly339, Asp370 and Thr371, interact with MTHF (Fig. 2.4B). The backbone carbonyl groups of Leu335, Trp336, Asp370 and Thr371 form hydrogen bonds with nitrogen atoms of the pterin ring system. Trp196 and Trp336 side chains interact by π-stacking with MTHF: their ring systems are located at 3.5 Å distances on both sides of the pterin-imidazole ring system. The conjugated systems of both Trp residues are aligned parallel to that of MTHF. Ala50, Phe200 and Gly339 form additional hydrophobic interactions with the ring system of MTHF. The benzyl/glutamic acid end of MTHF reaches out of the pocket, as the MTHF cofactor of other photolyases, and is stabilized from one side by a hydrophobic interaction with Pro194.

The π-stacking interaction is exceptional for chromophores in the CPF family. We therefore tested the role of either Trp by site-directed mutagenesis. The spectra of purified W196A and W336A mutants are devoid of the strong 380 nm absorption characteristic for MTHF which is found for the wild type protein or other mutants and resemble those of free oxidized FAD (Fig. 2.4C). Thus, both Trp residues are required for efficient interaction. The W196A and W336A mutants do either not bind MTHF or bind MTHF weakly so that it is lost during purification. As expected, the PhrA MTHF binding pocket could not be found in the structure of Escco-PL (Park *et al.*, 1995) or other photolyase structures.

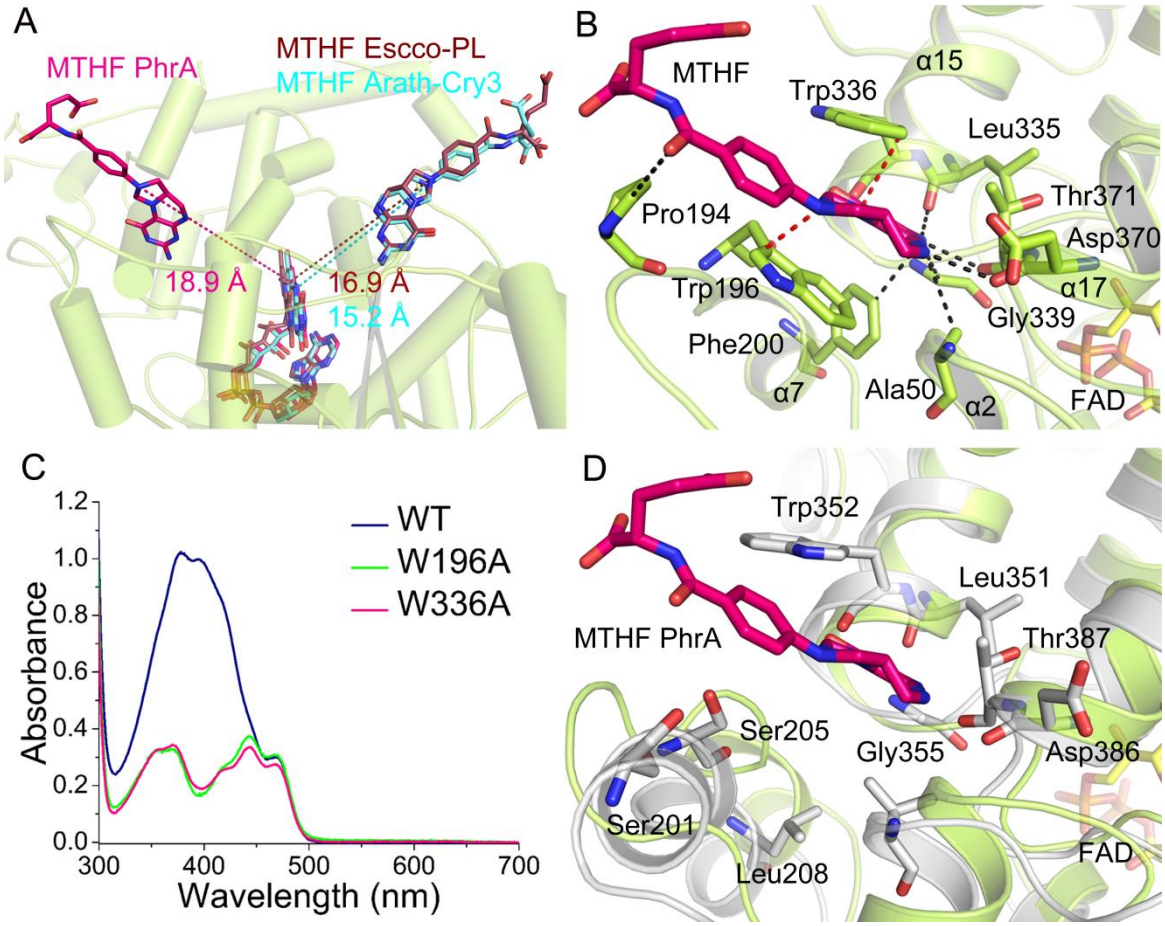


**Fig. 2.4. MTHF binding sites of PhrA and other MTHF binding CPFs.** (*A*) MTHF chromophores of MTHF binding CPFs after superpositioning of the proteins. PhrA is illustrated in ribbon representation. FAD and MTHF of PhrA, *E. coli* CPD photolyase (Escco-PL, PDB entry 1DNP) and *A. thaliana* Cry3 (Arath-Cry3, PDB entry 2VTB) are drawn as sticks. (*B*) MTHF binding in PhrA. Potential hydrogen bonds or van der Waals contacts are shown in dashed lines. The MTHF interacting amino acids are drawn as sticks. (*C*) UV-Vis spectra of wild type PhrA and the mutants W196A and W336A. (*D*) Superposition of PhrA and *A. thaliana* Cry1, (Arath-Cry1, PDB entry 1U3C), MTHF cavity. Protein backbones of PhrA and Arath-Cry1 are drawn as ribbon in green and gray, respectively. MTHF of PhrA and the potential MTHF interacting amino acids of Arath-Cry1 are drawn in sticks.

### 2.2.4 MTHF binding in CPD III Photolyase

To investigate the conservation of MTHF binding site in CPD III photolyase, we aligned 35 members, based on which the phylogenetic studies led to the definition of class III photolyases (Partch, 2006). Except Phe200 and Pro194, the MTHF coordinating residues are conserved to 80%-97% (Table 2.3). The two π-stacking residues, Trp196 and Trp336 are conserved in 33 and 32 of the 35 sequences, respectively. In photolyases gb ABA78296 from *Rhodobacter sphaeroides* and gb AAV95194 from *Ruegeria pomeroyi*, Trp196 is replaced by Arg and Trp336 by Phe and Tyr, respectively; in photolyase gi 53686955 from *Nostoc punctiforme*, Trp336 is replaced by Asn. Three class III representatives thus have only one π-stacking amino acid (Trp, Tyr or Phe). The results of our W196A or W336A mutants of PhrA described above suggests that these proteins have either lost their capacity to attach MTHF or that they bind MTHF with a lower affinity. To this end, we propose that the other 31 CPD III photolyases incorporate MTHF in the same tight manner as PhrA.

**Table 2. 3. The conservation of the MTHF interacted residues of PhrA in CPD III, CryDASH and CPD I subfamilies.** 35 CPD III members, 22 PlantCry members, 85 CPD I and 15 CryDASH members were used for conservation analysis. "-" indicates that there is no residue at the homologous position. Percentages show the degree of conservation. Amino acids that can also appear at the same position are given in brackets.

| PhrA | A50 | P194 | W196 | F200 | L335 | W336 | G339 | D370 | T371 |
|---|---|---|---|---|---|---|---|---|---|
| CPDIII | A97%(V) | P71%(G,N,I) | W94%(R) | L51% | L94%(M) | W91%(N,F,Y) | G100% | D85%(E,N,K) | T80%(C,A,H) |
| PlantCry | V72% | K77% | S100% | L100% | 90%(I,A) | W95%(L) | G100% | D100% | T80%(A,V) |
| CPDI | - | - | F29% | - | L81% | N47% | - | Q22% | H24% |
| CryDASH | - | - | V20% | - | L100% | A26% | - | T46% | L20% |

### 2.2.5 Potential MTHF binding site in plant cryptochrome

The question whether plant cryptochromes have an antenna chromophore is not yet settled. Indirect evidence suggested that purified recombinant *Arabidopsis* Cry1 (=HY4) contains MTHF (Malhotra *et al.*, 1995), but in other preparations only FAD was present as chromophore (Brautigam *et al.*, 2004). It therefore seems that plant cryptochromes do not incorporate an antenna chromophore or that a chromophore binds weakly in vivo which is then lost during purification. As the sister group of class III photolyases, plant cryptochromes might use the same MTHF binding site as PhrA. Indeed, the Escco-PL MTHF binding cavity is also filled in Arath-Cry1 (Brautigam *et al.*, 2004; Park *et al.*, 1995), whereas a groove of proper size is found at the position of the PhrA MTHF binding pocket. All four hydrogen bond forming residues including Trp336 are highly conserved in plant cryptochromes and the hydrogen bond forming atoms are at the same positions in Arath-Cry1. The ring system of the Trp336 homolog of Arath-Cry1 (Trp352) is rotated out of the prospective binding pocket but a rotation around a C-atom could bring it to the same position as in PhrA. (Fig. 2.4D). The other π-stacking Trp of PhrA, Trp196, is replaced by a Ser in plant cryptochromes (Ser205 in Arath-Cry1). We conclude that MTHF would fit into the pocket of plant cryptochromes and interact with the protein but that the interaction is weaker than in PhrA. Whether or not the affinity is high enough to ensure MTHF binding in vivo requires further investigations.

### 2.2.6 Two Trp-triads in PhrA

In PhrA, the transition of FAD from the oxidized to the enzymatically active reduced form has been confirmed as a blue light dependent process in vitro (Oberpichler *et al.*, 2011). In the "classical" way of photoreduction, electrons are tunneled from the surface via a triad of three conserved Trp residues to FAD (Oberpichler *et al.*, 2011). PhrA has the same three Trp residues (Trp308, Trp361 and Trp384), which are located at edge-to-edge distances of 4.4, 4.8 and 4 Å to each other and to the isoalloxazine ring of FAD, respectively. The PhrA crystal structure revealed also a second Trp-triad as a putative electron pathway (Fig. 2.5A), Trp367-Trp318-Trp384, with edge-to-edge distances of 3.8, 3.8 and 4 Å, respectively. Trp384 is shared by both triads and is at the same time the nearest one to FAD. To

investigate whether both Trp-triads function as electron pathways for FAD photoreduction, we mutated each of the involved Trp to Ala or Phe and analyzed photoreduction properties through UV-Vis spectroscopy.

The wild type and mutant proteins showed bright-yellow color after experiencing the same purification procedure under aerobic conditions except W384F that had a greenish color. This color is due to a strong absorption in the 510-560 nm range, which is related to semi-reduced FAD. All other mutants and the wild type protein showed insignificant absorption in this range and have spectral characteristic of fully-oxidized FAD (Fig. 2.6A). Apparently, FAD oxidation in W384F is arrested in the semi-reduced form. For photoreduction assays, all purified proteins are incubated overnight under oxidative conditions. After this time, the W384F mutant could also convert into the oxidized state (Fig. 2.6A). Hence, the Trp384 is required for efficient oxidation.

Photoreduction of the wild type protein as characterized by the loss of absorbance at 450 nm follows a typical pattern and is almost complete at 50 min after onset of light. Like the wild type protein, all proteins with a single mutation are capable of photoreduction by blue light in the presence of DTT (Fig. 2.6B), but the reduction is slower than in the wild type whereas the initial decrease was very slow in W384F, a mutant in which according to our hypothesis both pathways are affected by replacement of the shared Trp. The double mutant W308F/W361F (both amino acids within the classical pathway) is also characterized by an initial delay, which lasts however only a few minutes. Thereafter, a very rapid reduction was observed. The double mutant W308F/W367F in which the surface Trps of both triads were exchanged showed the slowest photoreduction. However, the rate was still not zero and the curve was again complex. In PhrA mutants, the second component could result from a light induced protein conformational change that facilitates electron flow. Although these complex photoreduction pattern are not clearly understood, our results show that both Trp-triads are required for normal photoreduction. The results also suggest that possibly other pathway exists, since the W384F mutant and the W308F/W367F double mutant still undergo photoreduction.

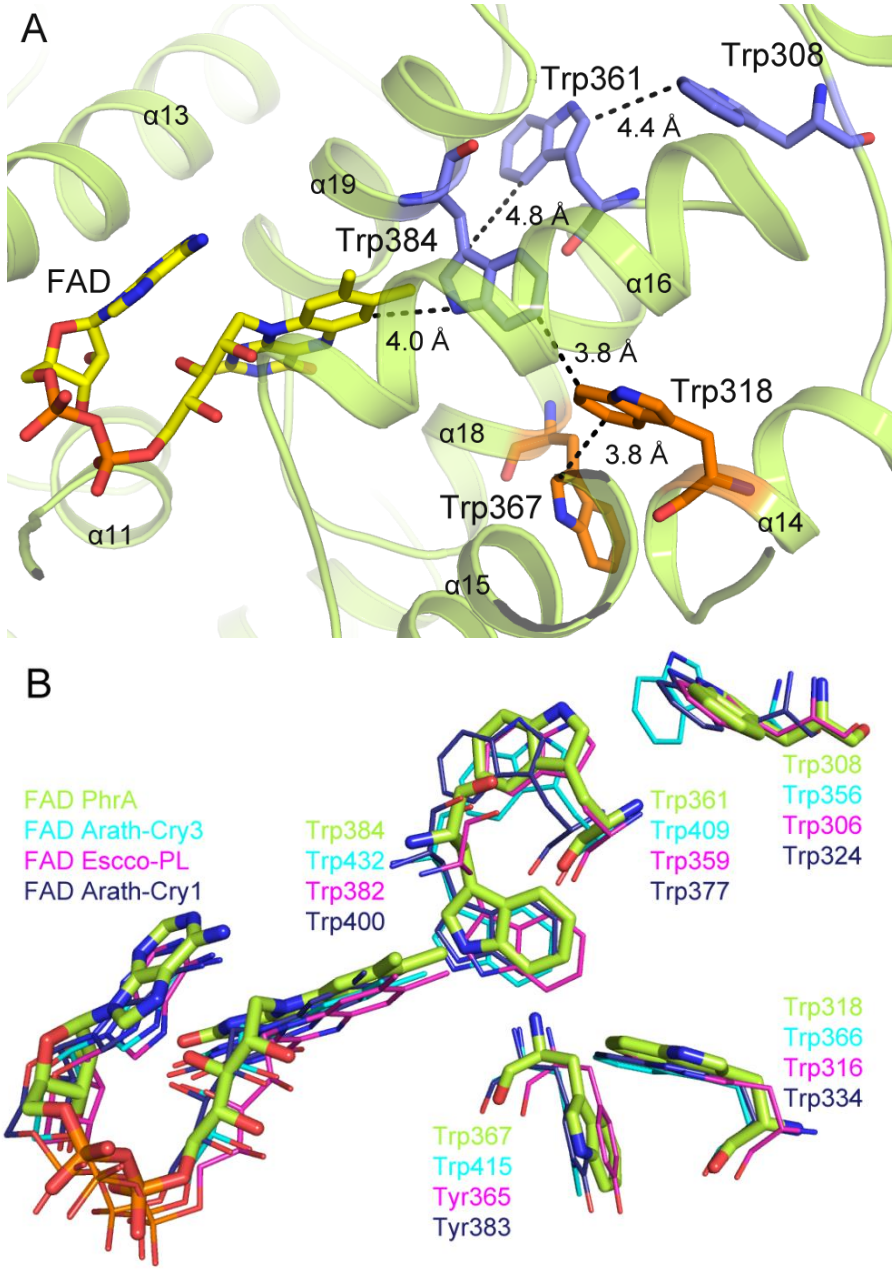


**Fig. 2.5. Trp-triads of PhrA and comparison with other PCF structures.** (*A*) Two Trp-triads in PhrA. The classical Trp-triad is shown in blue sticks, while the new one is shown in orange sticks. (*B*) Superposition of the residues of PhrA Trp-triads with *E. coli* photolyase (Escco-PL), *A. thaliana* cryptochrome 1 (Arath-Cry1) and *A. thaliana* cryptochrome 3 (Arath-Cry3).

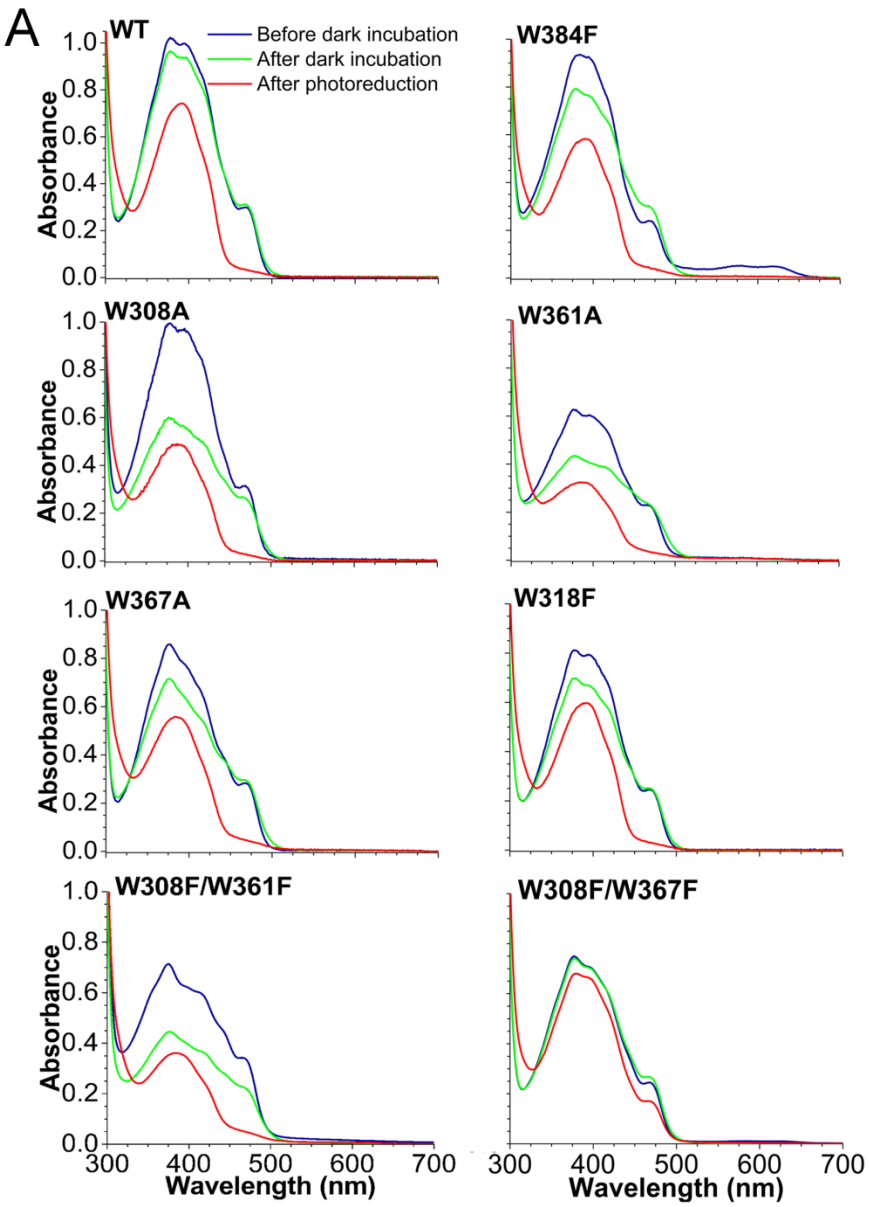
A
WT
Before dark incubation
After dark incubation
After photoreduction
W384F
W308A
W361A
W367A
W318F
W308F/W361F
W308F/W367F
Absorbance
Wavelength (nm)
B
WT
W308A
W361A
W384F
W367F
W318A
W308F/W367F
W308F/W361F
A450 nm
Time (sec)

**Fig. 2.6. Photoreduction of PhrA.** UV-vis spectral properties (*A*) and photoreduction of PhrA wild type (WT) and mutated proteins (*B*). Absorption values at 450 nm were taken from UV-vis spectra measured at indicated time points upon onset of blue-light irradiation. For each protein, these values were normalized against the value measured at t= 0 min.

.

### 2.2.7 The unique Trp-triad in other classes of CPF

The classical photoreduction pathway is conserved in all class I photolyases, Cry-DASH proteins, cryptochromes, eukaryotic (6-4) photolyases and class III photolyases. We found Trp367 and Trp318 of the second electron pathway of PhrA conserved in 34 out of 35 investigated class III photolyases. In the one remaining sequence, the surface located Trp367 is substituted by a Tyr. The new Trp-triad is also conserved in 60% of 15 analyzed Cry-DASH members, including Arath-Cry3 (Fig. 2.5B) (PDB entry 2VTB, Trp415-Trp366-Trp432). In Escco-PL and Arath-Cry1, Trp318 of PhrA is conserved (Trp316 and Trp366, respectively), whereas Trp367 of PhrA is replaced by a Tyr residue (Tyr365 and Tyr383, respectively). Of the 85 and 22 investigated class I and plant Cry members, 57% and 95% follow the same pattern, respectively. Our above results with PhrA suggest that other CPF proteins that share the Trp residues of the alternative triad might also have two electron pathways. Since besides Trp also Tyr can serve as electron transmitter in photolyases (Aubert *et al.*, 1999; Aubert *et al.*, 2000), also those CPF members in which one of the Trp is replaced by Tyr might have a second functional electron pathway. Photolyase mutants in which Trp residues of the classical pathway are replaced are usually characterized by a reduced or completely blocked photoreduction. These results speak for one pathway only. However, the alternative pathway could be active in vivo. The Escco-PL W306F mutant in which the classical pathway is blocked can repair DNA in vivo (Kavakli & Sancar, 2004). Based on these experiments it has been proposed that photoreduction is not required for in vivo function and that FAD of photolyases is always in the reduced state, due to the redox properties inside the cell. However, if a second pathway is present, these data can be interpreted in a different way. It could be that during DNA repair (which involves an electron cycle between $FADH^-$ and DNA) electrons are leaking out and that FADH must be reduced with electrons from the surrounding medium via one of the triads.

### 2.2.8 Darkness bleach of MTHF in PhrA mutants

The mutated proteins reveal bleaching of MTHF absorption bands in the 370-420 nm range during 24 hours dark incubation, while the wild type protein shows only negligible UV-Vis spectral changes (Fig. 2.6). Except W367, which is located 6.2 Å away from the MTHF, all other mutated tryptophans are more than 10 Å away from MTHF, excluding a direct impact on the chromophore (Fig. 2.7). Interestingly, the W308 and W381 as the two farthest located ones showed the most significant darkness bleach among all tryptophan mutants (Fig. 2.6). The fact that W318A and W384A are insoluble while W318F and W384F are soluble clearly hints to a structural importance of these residues. In contrasts to the MTHF bleach of photolyases under strong light condition (Hammalvarez *et al.*, 1989; Heelis *et al.*, 1987; Moldt *et al.*, 2009), the MTHF bleach of PhrA mutants occurred in darkness which is very likely arisen from the degradation of MTHF. The stability of MTHF in mutants is obviously decreased due to the mutation caused structural changes.

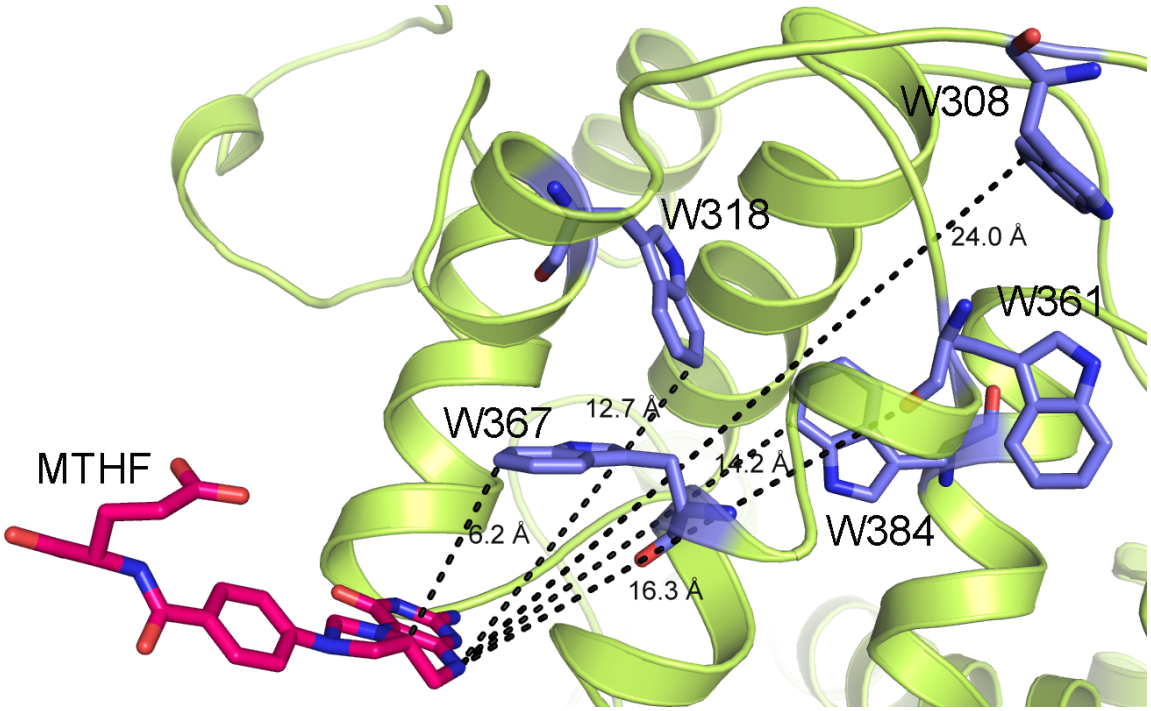


**Fig. 2.7. The edge-to-edge distances between the Trp-Triads consisting tryptophans and MTHF.**

### 2.2.9 Surface charge of PhrA

Consistent with other photolyases, PhrA possesses a positive surface charge at the FAD-accessing cavity region (Fig. 2.8). The predominantly negatively charged surface of plant cryptochromes at this region is one of the main causes for their missing DNA binding and repair capacity. The surface property of the cavity is defined by twelve residues , which differ between Arath Cry1 and photolyases (Brautigam *et al.*, 2004). Also, the FAD-accessing cavity of Arath-Cry1 is more commodious than that of photolyases or CRY-DASH (Huang *et al.*, 2006; Klar *et al.*, 2007; Mees *et al.*, 2004; Park *et al.*, 1995). Although PhrA is closely related to plant cryptochromes, the surface of the FAD-accessing cavity is positively charged, in line with that of Escco-PL, Anani-PL and Arath-CryDASH. Six out of the twelve residues are strictly conserved in PhrA, Escco-PL and Anani-PL but different in Arath-Cry1 (Fig. 2.8D). The positive charge in PhrA is related to Asn343 and Gln406. At these positions Arath-Cry has negatively charged Asp393 and Glu422 residues. Thus, despite the close relationship between PhrA and Arath-Cry1, the surface in the FAD cavity of PhrA is more similar to that of other photolyases, with which it shares the common function. The change of surface charge properties must be closely related to the functional change from photorepair to photoreceptor.

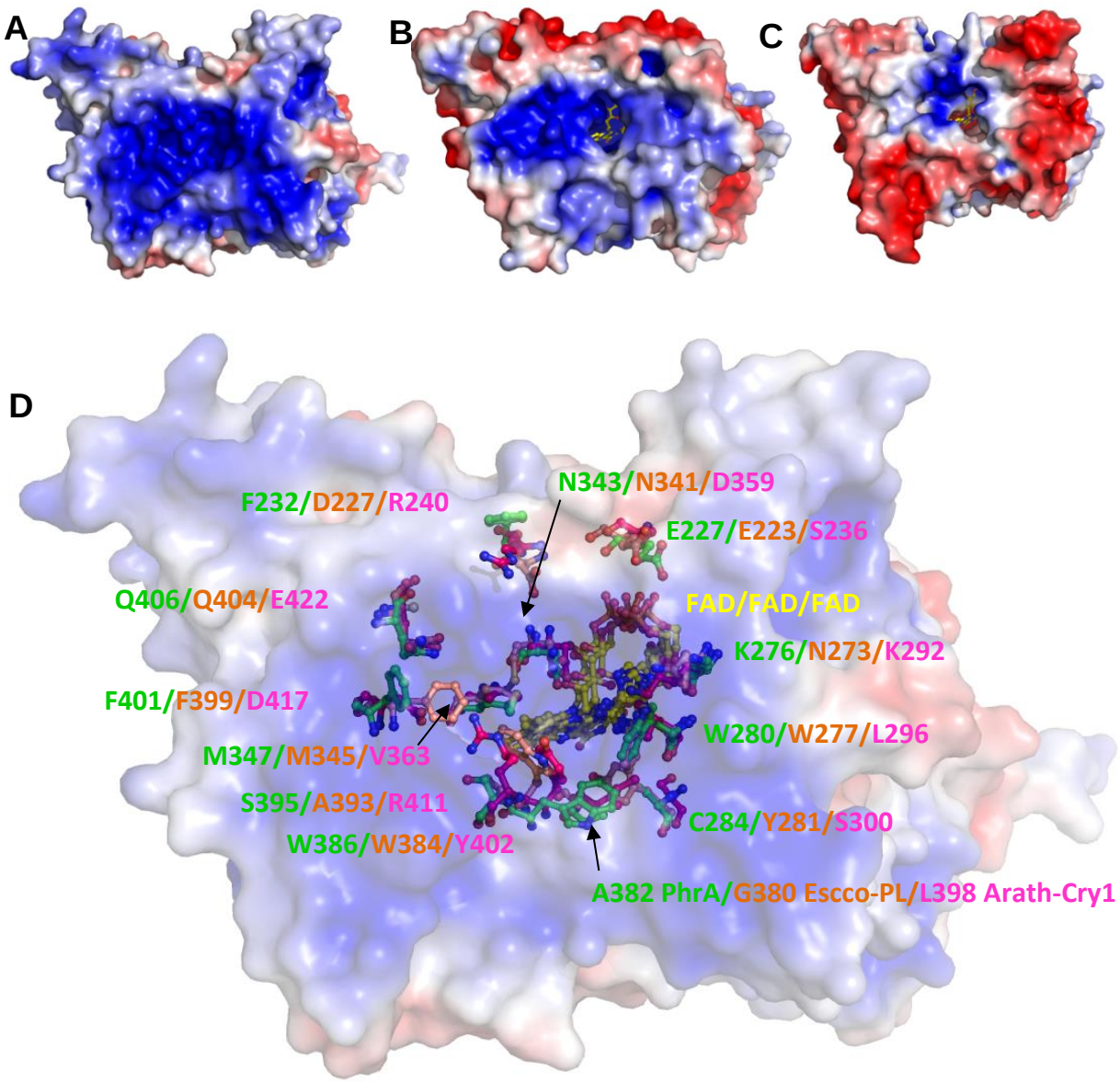


**Fig. 2.8. Surface charge of PhrA and comparison with *E. coli* photolyase and *A. thaliana* cryptochrome 1.** Electrostatic surface potentials were calculated using the program APBS with the non-linear Poisson-Boltzmann equation and contoured at +/- 5kT/e. Negatively and positively charged surface areas are colored in red and blue, respectively. (*A*) Electrostatic surface representation of PhrA. (*B*) Electrostatic surface representation of *E. coli* photolyase (1DNP). (*C*) Electrostatic surface representation of *A. thaliana* cryptochrome 1 (1U3C). (*D*) Superposition of PhrA, *E. coli* photolyase and *A. thaliana* cryptochrome 1. A transparent mode was selected for the PhrA surface. FAD and the 12 residues that are important the properties of FAD-accessing cavities are shown as balls-and-sticks.

### 2.2.10 Putative DNA lesion-binding mode in PhrA

In class I photolyases, six conserved residues (Arg232, Glu283, Trp286, Asn349, Met353 and Trp392 of Anani-PL) (Fig. 2.9A) (Mees *et al.*, 2004), form the majority of the concave surface of the CPD-binding cavity and interact directly with the CPD lesion. These amino acids are identical in Anani-PL and PhrA (Trp 230, Met 277, Trp 280, Arg 343, Met 347 and Trp386) and almost all have the same spatial orientation. Only the aromatic side chain of Trp386 is rotated away from the putative CPD lesion; its NE1 atom interacts with Arg147 and Asp393 via a water molecule WAT1129 (Fig. 2.9 and 2.10). Trp392 of Anani-PL forms a van der Waals contact with the methyl group of one of the thymines. When the homologous Trp384 in Escco-PL was mutated to Phe, the catalytic activity was decreased 100 fold (Li *et al.*, 1991). Arath-Cry3 has a Tyr (Tyr434) at the homologous position. The side chain of this Tyr is also moved away from the lesion site and forms a direct interaction with Asp441 (of Arath-Cry3) (Huang *et al.*, 2006). In the presence of damaged DNA, this side chain rotates towards the lesion and forms a close interaction (Klar *et al.*, 2007). The side chain of Trp386 in PhrA is located at about the same position as that of Tyr434 in Arath-Cry3 although both aromatic planes are almost perpendicular to one another. Nevertheless, Trp386 of PhrA could rotate to the functional position as the Tyr434 of Arath-Cry3 does upon the DNA substrate binding as outlined in Fig. 2.9B. Given the structural conservation within each CPF subfamilies, this “rotate and bind” mode probably occurs in other CPD III members.

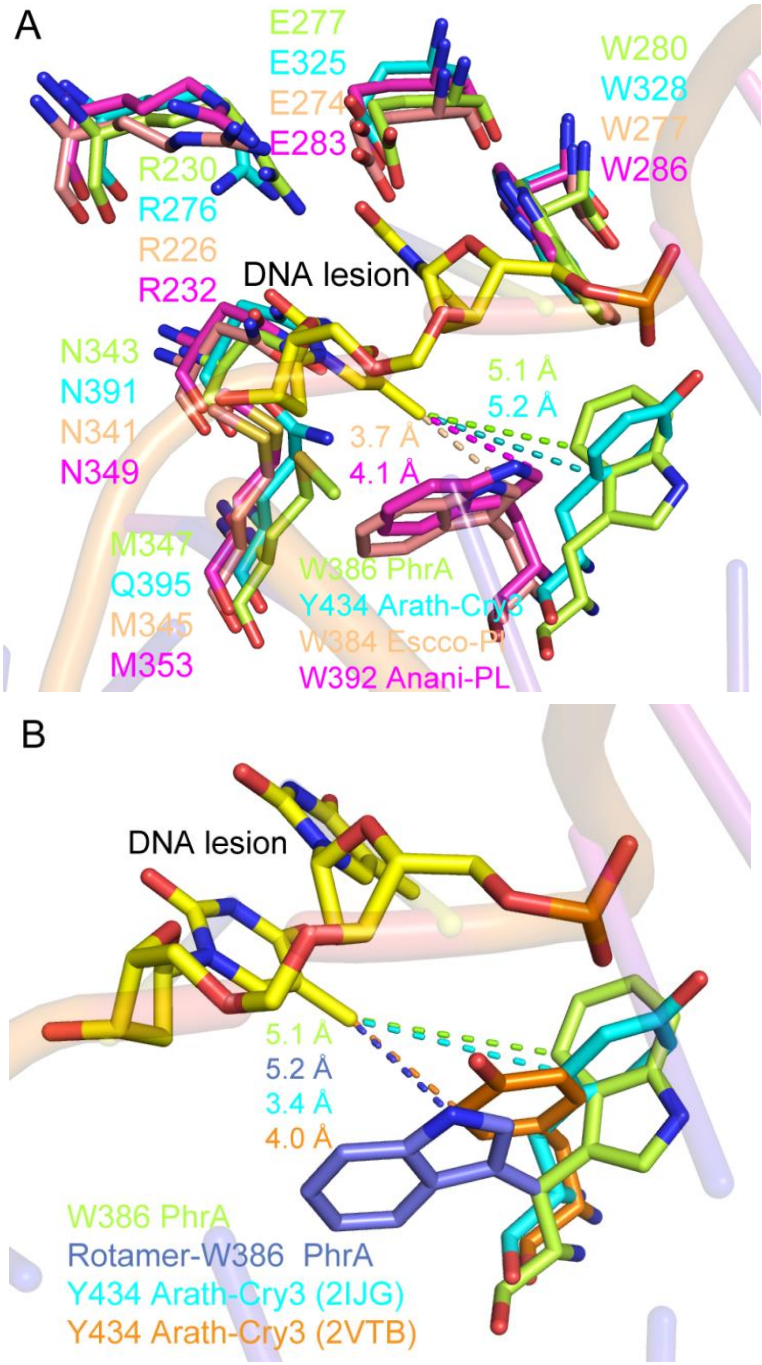


**Fig. 2.9. Putative CPD lesion binding of PhrA.** (*A*) Superposition of the six residues that directly interact with the CPD lesion. DNA is shown in the cartoon mode. The TT-dimer is shown in the ball-and-sticks mode in yellow color. Amino acids are also shown in the ball-and-sticks mode. (*B*) Potential rotamer of W386 (drawn in blue) that could interact with the CPD lesion in the same manner as W392 of Anani-PL.

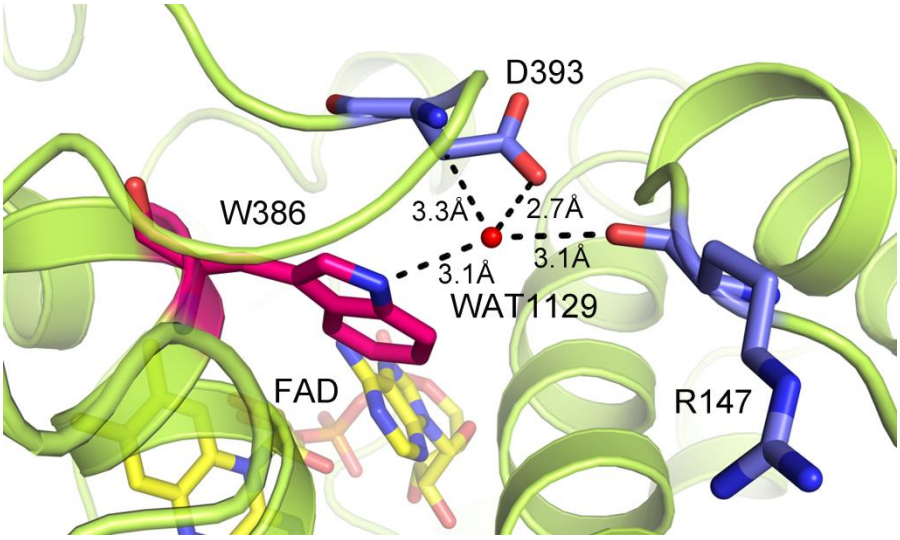


**Fig. 2.10. Hydrogen bonding network of Trp386 in PhrA**. The NE1 atom of Trp386 interacts with Arg147 and Asp393 via a water molecule WAT1129. PhrA is shown in green (ribbon model). Trp386 (colored in pink), Arg147, Asp393 (colored in blue) and FAD (colored in yellow) are shown as sticks.

# 3 Function and crystal structure of PhrB

## 3.1 Materials and Methods

### 3.1.1 PhrB protein purification and crystallization

Recombinant His-tagged PhrB protein was expressed in *E. coli* and purified via Ni-affinity chromatography and size-exclusion chromatography as described and used for crystallization (Oberpichler *et al.,* 2011) at concentrations of 4-6 mg/ml under dim light conditions or in darkness. Crystallization screens were carried out by the sparse matrix method (Jancarik & Kim, 1991) in a sitting-drop vapour diffusion approach testing more than 1,000 crystallization conditions at 277 K and 298 K in 96-well MRC plates. Promising conditions were systematically screened further by changing protein concentration, pH and the concentration of precipitation agents. After optimization, crystals were grown by hanging drop vapour diffusion at 298 K in 24-well Linbro plates. Each hanging drop was prepared on a siliconized glass plate by mixing 5 µl PhrB with 5 µl reservoir solution. The protein solution contained 12.5 mM tris (hydroxymethyl) amino methane, 1.25 mM ethylenediaminetetraacetic acid, 2.5 % glycerol, 75 mM sodium chloride, pH 7.8. The reservoir solution contained 2-6 % polyethylene glycol 400, 100 mM 2-(N-morpholino) ethanesulfonic acid buffer, pH 6.0. The obtained PhrB crystals were flash frozen in liquid nitrogen after cryosoaking with 4 M trimethylamine N-oxide for a few minutes (Mueller-Dieckmann *et al.,* 2011).

### 3.1.2 Structure analysis (Structure determination and refinement was done in cooperation with Patrick Scheerer and Norbert Krauß)

Diffraction data collection was performed at 100 K using synchrotron X-ray sources at ESRF, Grenoble, France, and BESSY II, Berlin, Germany. Best diffraction of anomalous and native data were collected at beamline BL 14.2 at BESSY II / Helmholtz Zentrum Berlin für Materialien und Energie with an MAR-225CCD detector at $\lambda$= 1.7409 Å and $\lambda$= 0.91841 Å, respectively. All images were indexed, integrated and scaled using the XDS program

package (Kabsch, 2010) and the CCP4 program SCALA (Bailey, 1994; Evans, 2006). The single-wavelength anomalous dispersion (SAD) dataset at λ= 1.7409 Å (peak) was collected to 1.95 Å resolution. Initial experimental phases were determined to 1.95 Å resolution by the SAD method based on this dataset by using SHELXC/D (Sheldrick, 2008) and a beta version of SHELXE (Sheldrick, 2010) to solve the iron-atom substructure. Substructure solution with SHELXD was successful for space group $P2_12_12$, yielding 4 iron sites (with occupancies greater than 0.62). With this solution a mean figure-of-merit of 0.640 was obtained after initial phasing, polyalanine-chain tracing and density modification with SHELXE (Sheldrick, 2010) (that increased to 0.770 after density modification with the program DM (Emsley *et al.*, 2010) Subsequent steps of crystallographic refinement, consisting of torsion angle molecular dynamics, simulated annealing using a slow-cooling protocol and a maximum likelihood target function, energy minimisation, and B-factor refinement by the program CNS (Brunger *et al.*, 1998) were carried out using a high resolution native data set in the resolution range 19.95–1.45 Å. After the first round of refinement, the [4Fe-4S] cluster, the FAD and 6,7-dimethyl-8-ribityl-lumazine were clearly visible in the electron density of both $\sigma_A$ - weighted $2F_o - F_c$ maps, as well as in the $\sigma_A$ - weighted simulated annealing omitted density maps. The electron density map was readily interpretable, and the phases were input for initial automated model building using ARP/wARP 7.0.2 (Langer *et al.*, 2008). Restrained, individual B-factors were refined and the crystal structure was finalised by the CCP4 program REFMAC5 and CCP4 (Bailey, 1994) and PHENIX (Adams *et al.*, 2010). The final model has agreement factors $R_{free}$ and $R_{cryst}$ of 18.0 % and 13.8 %. Table 3.3 summarises the statistics for crystallographic data collection and structural refinement. Manual rebuilding of the PhrB model and electron density interpretation was performed after each refinement cycle using the program COOT (Emsley & Cowtan, 2004) and WHAT_CHECK (Hooft *et al.*, 1996). Potential hydrogen bonds and van der Waals contacts were analysed using the programs HBPLUS (McDonald & Thornton, 1995) and LIGPLOT (Wallace *et al.*, 1995). All crystal structure superpositions of backbone α carbon traces were performed using the the DALI server (Holm & Rosenstrom, 2010). The PhrB structure was analyzed for salt bridges utilizing the WHAT IF server (Rodriguez *et al.*, 1998) Electrostatic surface potentials were calculated through solution of the Poisson-Boltzmann equation using the program APBS

(Baker *et al.*, 2001). All molecular graphics representations were created using PyMOL (DeLano, 2002)

### 3.1.3 Analytical HPLC

An Agilent 1200 Series HPLC system managed by Chem Station Software, equipped with a DAD detector (Agilent, Waldbronn Germany) and a Nucleosil 100-5 analytical column (250 x 4.6 mm, Macherey and Nagel, Düren Germany) was used for analysis of organic cofactors. For DNA analysis the same system with a Nucleodur 100-3 column (250x4.0mm, Macherey and Nagel, Düren Germany) was used.

A protocol established for CryB (Hendrischk *et al.*, 2009) was used for analyses of organic cofactors: Purified PhrB was mixed with 7.2 % trichloroaceticacid followed by 1 h shaking at 100 rpm on ice, and then mixed with 5 % (v/v) of 0.1 % formic acid ($HCO_2H$) in acetonitrile (ACN) and 0.35 mM NaOH (final concentrations). The mixture was centrifuged at 13,000 g for 10 min, and the supernatant was analyzed by HPLC. Synthetic 6,7-dimetyl-8-ribityl-lumazin (DMRL, a generous gift of Prof. Bacher) and FAD were used as references. Eluting buffers of HPLC were buffer A (0.1% $HCO_2H$ in ACN) and buffer B (0.1% $HCO_2H$ in $H_2O$). After injection of 10 µl sample, separation was achieved at 298 K and a flow rate of 1 ml/min; a 5 min isocratical run with 95% (v/v) of buffer B was followed by a gradient from 95-25% buffer B in 20 min. Elution was monitored at 264 nm and UV-Vis spectra of were automatically recorded.

For the (6-4) photoproduct repair assay, HPLC was conducted as described (Glas *et al.*, 2010). The UV-damaged DNA (GCGGT(6-4)TGGCG paired with TCGCCAACCGCT (100 µmol/L)) were dissolved in repair buffer (50 mM Tris/Cl, 100 mM NaCl, 5 mM DTT, 1 mM EDTA, 5 % glycerol, pH 7.5) and incubated with PhrB (300 µmol/L in repair buffer). The reaction mixture was irradiated with a 405 nm light emitting diode (GB333UV1C/L1, Farnell Diodes) at an intensity of 40 µmol $m^{-2}$ $s^{-1}$; controls in which (6-4)PP was incubated with PhrB in darkness were performed in parallel. The reaction was stopped by heating to 368 K for 10 min and the sample was centrifuged at 13000 g for 10 min. The supernatant was subjected to HPLC. For comparison, untreated (6-4) photoproduct mixing with undamaged

DNA was processed analogously. Eluting buffers of HPLC were buffer C (0.1 M TEAA (triethylammonium acetate) in $H_2O$) and buffer D (0.1 M TEAA in $H_2O$/ACN 20/80). The gradient was 0-17 % buffer D in 60 min with a flow rate of 0.5 mL/min. Elution was monitored at 260 nm and 325 nm. Based on the elution retention time and UV-Vis spectral properties, the (6-4) photoproduct and repaired or undamaged DNA can easily be distinguished.

## 3.2 Results and Discussion

### 3.2.1 Function of PhrB

Our previous repair assay for PhrB failed to detect either CPD photolyase or (6-4) photolyase enzymatic activity in vitro (Oberpichler *et al.*, 2011). However, while the CPD repair assay result was reliable since PhrA as a CPD III photolyase could be clearly confirmed through that repair assay, the results of the (6-4) photoproduct repair assays were not conclusive because of the uncertain reproducibility, which might be attributed to two main reasons. Firstly, the yield and purity of the (6-4) photoproduct substrate was poor. The (6-4) photoproduct was prepared through UV-C irradiation without experiencing any further purification. This together with other yet unknown factors might lead to a low efficiency of the repair. Secondly, the UV-Vis spectroscopy that was used for monitoring of the repair assay could not meet the requirement of the low repair efficiency. Therefore the previous work could only claim that no (6-4) photoproduct repair activity was detected under the investigated conditions rather than clearly supporting or denying the (6-4) photoproduct repair activity of PhrB.

To investigate whether PhrB could function as a (6-4) photolyase with trustable accuracy, we then used highly purified single stranded or double stranded DNA oligonucleotides with a (6-4) lesion as substrates and performed an enzyme assay (Glas *et al.*, 2010). As shown by the HPLC analysis, in the ssDNA (6-4) photoproduct repair assay, 60 and 120 min incubation with PhrB led to the conversion of about 5% and 15% into normal DNA, respectively (Fig. 3.1A). In the dsDNA (6-4) photoproduct repair assay, after 5 and 15 min incubation with PhrB, about 20% and 80% of substrate were converted into normal DNA, respectively (Fig.3.1B). PhrB thus repairs both ssDNA and dsDNA in a blue light and irradiation time dependent manner, in line with activity studies for eukaryotic (6-4) photolyases (Sancar, 2004). We have thus shown that PhrB is a (6-4) photolyase and that prokaryotes are also to repair (6-4) photoproducts by photolyases.

PhrB showed much higher repair efficiency for dsDNA than for ssDNA under the same condition indicating that the complementary strand could improve the PhrB repair efficiency through enhanced recognition or binding. However, more detailed conclusions can only be

drawn after more quantitative experiments on repair efficiency with single strand and double strand DNA are performed.

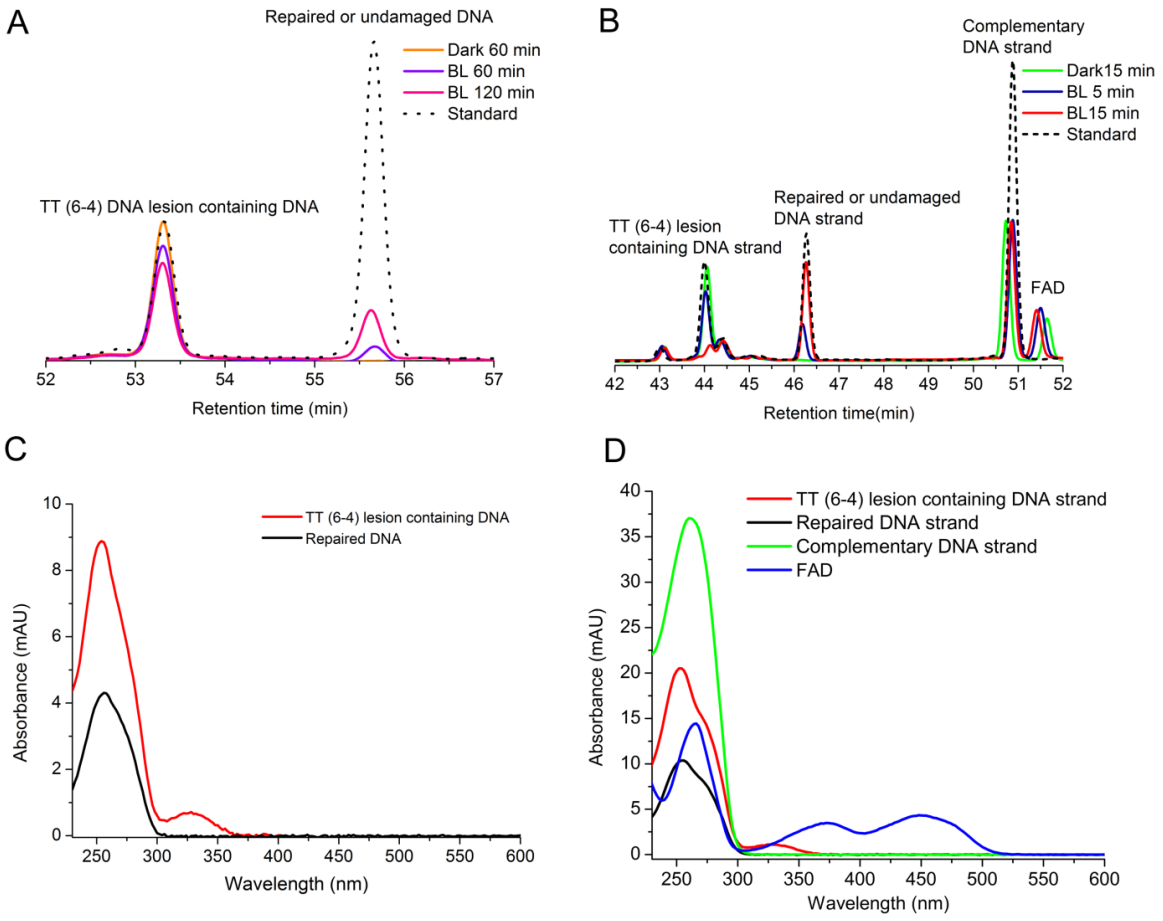


**Fig. 3.1. HPLC analysis of PhrB DNA repair assay**. (*A*) HPLC analysis of PhrB (6-4) photoproduct repair assay with single stranded DNA. The "Standard" is a mixture of damaged DNA with a TT (6-4) lesion and undamaged DNA (black dotted line). In the other samples, damaged DNA was incubated with PhrB for 60 and 120 min under blue light (BL, purple and magenta lines, respectively) or 60 min in darkness (orange line). Thereafter, DNA was separated from the protein and subjected to HPLC. (*B*) HPLC analysis of PhrB (6-4) photoproduct repair assay with double stranded DNA. The "Standard" is a mixture of damaged DNA with a TT (6-4) lesion and undamaged DNA (black dashed line). In the other samples, damaged DNA was incubated with PhrB for 5 and 15 min under blue light (BL, blue and red lines, respectively) or 15 min in darkness (green line). Thereafter, DNA was separated from the protein and subjected to HPLC. (*C*) Spectra from the repair assay with

single stranded DNA under blue light for 120 min. (*D*) Spectra from the repair assay with double stranded DNA under blue light for 15 min.

### 3.2.2 Function of FeS-BCPs

Our latest phylogenetic studies showed that PhrB belongs to a newly designated sub-family of CPF whose members are widely distributed in prokaryotes (Oberpichler *et al.*, 2011). In a Uniprot-KB BLAST survey we found 462 bacterial and archaeal PhrB homologs (FeS-BCPs) (Table 3.1), which form a group distantly separate from other CPF subfamilies. In the catalytic region between amino acids 276 and 474 that encompasses the FAD binding and DNA binding regions, residues are highly conserved. Therefore we propose that most if not all FeS-BCP members are (6-4) photolyases. This hypothesis is supported by our further structural studies.

Surprisingly but interestingly, a latest published study on another FeS-BCP member called CryB from *Rhodobacter sphaeroides* showed no *in vitro* DNA repair activity (Geisselbrecht *et al.*, 2012). Although there is a 4 residues deletion in the protein linker region, CryB shares 66% of sequence similarity with PhrB (Fig. 3.2). Besides this, several other features are in support of a same function of CryB as PhrB: the protein is required for full *in vivo* photorepair in *R. sphaeroides* (Hendrischk *et al.*, 2009), as is PhrB in *A. tumefaciens* (Oberpichler *et al.*, 2011). The surface charge and DNA binding properties of CryB (Geisselbrecht *et al.*, 2012; Hendrischk *et al.*, 2009) are also compatible with DNA repair activity. We thus reasoned that the negative result of the CryB DNA repair assay might arise from inappropriate *in vitro* conditions.

Meanwhile, on the other hand, there are evidences from previous studies supporting photoreceptor function for PhrB and CryB (Geisselbrecht *et al.*, 2012; Hendrischk *et al.*, 2007; Hendrischk *et al.*, 2009; Oberpichler *et al.*, 2011). Consequently, one could not exclude that CryB and PhrB and their homologs also have the cryptochrome function besides the (6-4) photolyase function. The iron-sulfur cluster, which distinguishes PhrB and its homologs from other CPF members, might be required for the possible dual function.

**Table 3.1. List of the accession codes of PhrB homologs.** Sequences were identified by BLAST in the Uniprot/KB database using default values. All entries with an E-value < 10 were selected; sequences with < 400 amino acids were discarded. Besides one viral and one eukaryotic entry, all sequences are from prokaryotes.

| | | | | | |
|---|---|---|---|---|---|
| C1F9W1_ACIC5 | B4V0S3_9ACTO | A5PBY1_9SPHN | B9KWA9_RHOSK | A4C0K4_9FLAO | E6VB64_VARPD |
| E8WYU4_ACISM | B5GIH2_9ACTO | Q0G5E4_9RHIZ | F3U3L2_RHOSH | A2U0A7_9FLAO | A2SHH4_METPP |
| C7M2C5_ACIFD | D9UEU9_9ACTO | A9D5J0_9RHIZ | B6AXP8_9RHOB | Q1VSI9_9FLAO | D7DKP6_METS0 |
| E2SE89_9ACTO | F3ZAK4_9ACTO | Q0C2A4_HYPNA | A3JVT2_9RHOB | A4CPD1_ROBBH | F5RFZ8_9RHOO |
| B8HC91_ARTCA | A5FZA0_ACICJ | Q28L54_JANSC | Q0FGJ4_9RHOB | Q2S0Y7_SALRD | E4QLZ6_METS6 |
| A5CQI5_CLAM3 | F0IZB1_ACIMA | A0NUF4_9RHOB | B6B656_9RHOB | D5HAW8_SALRM | C6XCJ0_METSD |
| B0RFB3_CLAMS | F7S859_9PROT | B9QZA3_9RHOB | B6IRE2_RHOCS | D5BLE5_ZUNPS | C4ZLC0_THASP |
| Q0RLR2_FRAAA | F0LDM5_AGRSH | A3V1T8_9RHOB | E2CJX3_9RHOB | C7RVE8_ACCPU | D5WYW7_THIK1 |
| D2S7C5_GEOOG | A9CH39_AGRT5 | Q0AQZ4_MARMM | Q16AY1_ROSDO | D4X546_9BURK | D6CPC8_THIS3 |
| A6WCK7_KINRD | F7UHL0_RHIRD | A3V9K3_9RHOB | F7ZAW4_9RHOB | F0Q8J2_ACIAP | B8G8B2_CHLAD |
| E4NJP4_KITSK | E0MQM7_9RHOB | B1M1I8_METRJ | A6FU84_9RHOB | A1TVR7_ACIAC | B9LIU8_CHLSY |
| D2PR70_KRIFD | A8TY99_9PROT | B0U8V9_METS4 | A4EEV7_9RHOB | F4G5S8_ALIDK | A9WDT4_CHLAA |
| E8N7H3_MICTS | D0RQM0_9RICK | E8L2J5_9RHIZ | B7RGZ3_9RHOB | B6BVN4_9PROT | Q7NT35_CHRVO |
| C4RCJ1_9ACTO | F4QNR3_9CAUL | Q2G5P3_NOVAD | A3X5Z8_9RHOB | Q2L1R2_BORA1 | A7NRT2_ROSCS |
| A4T9Z9_MYCGI | E8RNS6_9CAUL | F1Z8N4_9SPHN | D5RRT1_9PROT | C5AHE5_BURGB | A5UP99_ROSS1 |
| A3Q0P1_MYCSJ | A8I945_AZOC5 | A9DV81_9RHOB | A3VYB8_9RHOB | A9ASI6_BURM1 | B0CG90_ACAM1 |
| A1UH54_MYCSK | A5EFY7_BRASB | A3UCG3_9RHOB | A6DY54_9RHOB | F0FXR9_9BURK | Q4CB01_CROWT |
| Q1B7V3_MYCSS | A4YRA1_BRASO | B5K6H9_9RHOB | B7QVP7_9RHOB | A4JRC3_BURVG | B5IJS9_9CHRO |
| E6TP22_MYCSR | B4WEP7_9CAUL | B5IXW8_9RHOB | A3K3A6_9RHOB | Q1QZ55_CHRSD | B1WU38_CYAA5 |
| A1TA68_MYCVP | D9QHX4_BRESC | E0TGT6_PARBH | D0CPE2_9RHOB | D0J732_COMT2 | B7KA77_CYAP7 |
| C8X9H8_NAKMY | B6BRS3_9RICK | Q0FIT3_9RHOB | Q1GEA3_SILST | B7X2Z8_COMTE | B8HVI4_CYAP4 |
| D1BKQ0_SANKS | F2I1Y8_9RICK | Q4FNW5_PELUB | C9CV78_9RHOB | C9Y9Y1_9BURK | D4TEX5_9NOST |
| D6B3F2_9ACTO | Q1V157_PELUQ | A9FEZ6_9RHOB | F3WTQ0_9SPHN | D8IR78_HERSS | Q11W85_CYTH3 |
| Q9RIY1_STRCO | B8H0P9_CAUCN | A9G895_9RHOB | Q1NBK8_9SPHN | B1Y5D1_LEPCP | A0YMK1_LYNSP |
| E8W3Q9_STRFA | Q9AAF5_CAUCR | B6QZH1_9RHOB | A3SGB2_9RHOB | A6GQU5_9BURK | B0JI96_MICAN |
| D9XNI1_9ACTO | D5VJH2_CAUST | D5BN93_PUNMI | A3SVG9_9RHOB | F1VVM4_9BURK | A8YI00_MICAE |
| B1VXF9_STRGG | D0DB03_9RHOB | D5AP96_RHOCB | D6Y059_BACIE | A4SZA8_POLSQ | B2J2Y4_NOSP7 |
| D9WTE8_9ACTO | A8LPM2_DINSH | Q3IXP1_RHOS4 | C4KYP6_EXISA | E2SVA5_9RALS | A2BPQ5_PROMS |
| D6EHX4_STRLI | Q2N904_ERYLH | A4WZG8_RHOS5 | A1ZQR3_9BACT | F5Y262_9BURK | A8G3D9_PROM2 |
| D6AU96_STRFL | A3WAR2_9SPHN | A3PRB8_RHOS1 | A4BYS6_9FLAO | F3LM59_9BURK | A3PBE5_PROM0 |
| Q31CB1_PROM9 | D0LFX3_HALO1 | Q6D6U5_ERWCT | F7SIW1_9GAMM | Q3IJ56_PSEHT | F3E3V2_9PSED |
| A2BV85_PROM5 | A0DL57_PARTE | E1SU96_FERBD | D0L048_HALNC | F3BFR6_PSEHA | F2ZT39_9PSED |
| A2C3R1_PROM1 | Q3IP83_NATPD | F4BJ52_FRACN | A3WPE1_9GAMM | E6RPR4_PSEU9 | Q48KV0_PSE14 |
| Q46JW1_PROMT | E6X7P4_CELAD | F4BDH4_FRACF | Q5QY65_IDILO | A4CAK3_9GAMM | Q4ZVJ0_PSEU2 |
| B9P0H4_PROMA | F0RBU1_CELLC | B4ASH0_FRANO | F7RVQ7_9GAMM | Q1I9Z7_PSEE4 | F3K8F2_PSESZ |
| Q7V2P7_PROMP | A3UBB8_CROAH | E2MSQ2_FRANO | D1RLM1_LEGLO | Q3KCF5_PSEPF | Q87YV3_PSESM |
| D4TQ06_9NOST | A2TQD5_9FLAO | A7JKE9_FRANO | D3HJC3_LEGLN | F4DUH4_PSEMN | E2MDL3_PSESM |

| Q31P18_SYNE7 | Q26F22_FLABB | A7JG83_FRANO | A5IA76_LEGPC | A4XRK3_PSEMY | A1SV42_PSYIN |
|---|---|---|---|---|---|
| Q5N550_SYNP6 | A3J290_9FLAO | B0U050_FRAP2 | Q5WZX1_LEGPL | E4RI42_PSEPB | A4BEK4_9GAMM |
| B1XJB0_SYNP2 | C0BGX0_9BACT | C6YSI2_9GAMM | E1Y0A8_LEGPN | A5W4U0_PSEP1 | F7NU16_9GAMM |
| Q0IAF6_SYNS3 | C0BND7_9BACT | F8G993_9GAMM | D5T757_LEGP2 | Q88JB7_PSEPK | F7QEF1_9GAMM |
| Q3AJW7_SYNSC | A6H181_FLAPJ | A7NA38_FRATF | Q5ZZ06_LEGPH | B1J9T0_PSEPW | A1S8J3_SHEAM |
| Q3AXU6_SYNS9 | A0M4X7_GRAFK | Q2A572_FRATH | A6DHS7_9BACT | F8G2U0_PSEPU | A3D9D0_SHEB5 |
| Q2JK91_SYNJB | A9DMV1_9FLAO | Q0BNI7_FRATO | A0Z190_9GAMM | D7HXX1_PSESS | A6WIH0_SHEB8 |
| Q2JVG2_SYNJA | F4B1A6_KROS4 | A4KQ35_FRATU | A0YCX4_9GAMM | F0EAL4_9PSED | A9KYP4_SHEB9 |
| A5GT97_SYNR3 | F6GG84_9FLAO | A7YSJ1_FRATU | B7RU19_9GAMM | A4VI48_PSEU5 | B8E6D8_SHEB2 |
| A5GL82_SYNPW | A3XGD2_LEEBM | A0Q4V3_FRATN | E4PN85_MARAH | F2MY80_PSEU6 | E6T3F8_SHEB6 |
| A4CUH6_SYNPV | D7W3H0_9FLAO | Q14HZ4_FRAT1 | A6EZA2_9ALTE | F8GZ63_PSEST | E0HRK4_9GAMM |
| Q7U6U2_SYNPX | C6NSG3_9GAMM | D2AMZ8_FRATE | A1U5A6_MARAV | F3H183_PSESX | Q087D2_SHEFN |
| Q067A7_9SYNE | D0SKC1_ACIJU | A7JC99_FRATT | A3JBU9_9ALTE | F3JL91_PSESX | Q8E9U4_SHEON |
| B4WS80_9SYNE | D0BVM2_9GAMM | Q5NGJ2_FRATT | F6CX17_9GAMM | F3HW24_PSESF | E6XH41_SHEP2 |
| Q05SX3_9SYNE | D6JXM7_ACIG3 | C6YPI3_FRATT | A6VS44_MARMS | F3D9I9_9PSED | A4Y2S3_SHEPC |
| A3Z6R8_9SYNE | A0KJP2_AERHH | Q1YU58_9GAMM | F5T150_9GAMM | F3ISK7_PSEAP | A0L1K3_SHESA |
| A3Z202_9SYNE | F4DBN8_AERVB | F3LGD4_9GAMM | C0N1K2_9GAMM | E7P7Z0_PSESG | Q0HEC6_SHESM |
| D0CJ42_9SYNE | B6ENK1_ALISL | F3L425_9GAMM | Q2BJU7_9GAMM | E7PLR6_PSESG | Q0HZM3_SHESR |
| P74080_SYNY3 | A0Y308_9GAMM | B8KWL7_9GAMM | Q1ZUI8_PHOAS | F3FG37_PSESX | A1RP60_SHESW |
| A3I0F1_9BACT | F2G2A1_ALTMD | B8KP17_9GAMM | D0Z3N7_LISDA | F3EDH5_PSESL | B1KFL2_SHEWM |
| E4TUL0_MARTH | F5Z7W0_9ALTE | F4ALF7_GLAS4 | F2PDU1_PHOMO | F3IFZ8_PSESL | C5BIV6_TERTT |
| D3PKP9_MEIRD | Q1MZA8_9GAMM | D0I763_VIBHO | Q2C0V7_9GAMM | F3HM96_PSEYM | F6D9I1_9GAMM |
| C8X3W0_DESRD | A4A626_9GAMM | Q2S7E7_HAHCH | Q15S30_PSEA6 | F3EXC7_9PSED | B8GR61_THISH |
| D3SFV6_THISK | A3ENZ1_VIBCH | E8VW85_VIBVU | D5SWH4_PLAL2 | F0BD53_9XANT | E1DRV2_VIBPA |
| Q1VCR2_VIBAL | A1ERQ3_VIBCH | A5L696_9GAMM | A6CH24_9PLAN | Q5V4M4_HALMA | A6B1N0_VIBPA |
| D0WV36_VIBAL | C9NW21_9VIBR | Q8PMU2_XANAC | Q7UG69_RHOBA | Q9HRG8_HALSA | E1D5I9_VIBPA |
| F7YT42_VIBAN | Q5E4T0_VIBF1 | Q4USG6_XANC8 | F2ATR3_RHOBA | B0R3Z4_HALS3 | Q87JP4_VIBPA |
| E8LU01_9VIBR | A7N6Y3_VIBHB | B0RX18_XANCB | C7PTX9_CHIPD | D4GX25_HALVD | E1EBY5_VIBPA |
| E3BET7_9VIBR | D0XAG9_VIBHA | Q8PB43_XANCP | F4KSZ9_HALH1 | C7NZM7_HALMD | E1CWL6_VIBPA |
| C2C8L3_VIBCH | A6AK75_VIBHA | Q3BVU7_XANC5 | B0SD18_LEPBA | Q18HI9_HALWD | A6CY80_9VIBR |
| A2P9W3_VIBCH | C9P7Q0_VIBME | D4TAC5_9XANT | B0SLJ5_LEPBP | F7PN35_9EURY | E8MAC9_9VIBR |
| A1F9D3_VIBCH | D0H099_VIBMI | D4SY60_9XANT | C7DAW0_9RHOB | C7NQJ4_HALUD | A7JZ50_VIBSE |
| A6AFI1_VIBCH | D0HE32_VIBMI | F0C3V7_9XANT | A4GI54_9BACT | O27804_METTH | A8TBC5_9VIBR |
| A6XQB3_VIBCH | D2YTY0_VIBMI | Q2P4L1_XANOM | A9LH56_9BACT | D9PUQ6_METTM | A3Y1I3_9VIBR |
| A3H3W4_VIBCH | D2Y9R5_VIBMI | B2SRW7_XANOP | B5JLQ7_9BACT | D5EAI8_METMS | C9Q9J4_9VIBR |
| C2HZX3_VIBCH | F3RWL7_VIBPA | Q5H1Q1_XANOR | F4F626_VERMA | D5EKM8_CORAD | B7VP56_VIBSL |
| C2JBE7_VIBCH | B8K4H2_VIBPA | F0BSH5_9XANT | E3T4R9_9VIRU | E8R4X8_ISOPI | A3UMB0_VIBSP |
| C6S1S2_VIBCH | D7HNB0_VIBCH | A2PV58_VIBCH | D7HC14_VIBCH | D2R3F9_PIRSD | Q7MCG3_VIBVY |
| Q9KLD7_VIBCH | C6YIF9_VIBCH | A3GT48_VIBCH | C2IGH0_VIBCH | F0SIK8_PLABD | Q8D5I2_VIBVU |
| D0HNU0_VIBCH | A6A874_VIBCH | D0H742_VIBCH | A5F0H6_VIBC3 | C3LW73_VIBCM | C3NWI0_VIBCJ |
| C2I8A5_VIBCH | C2ISM8_VIBCH | | | | |

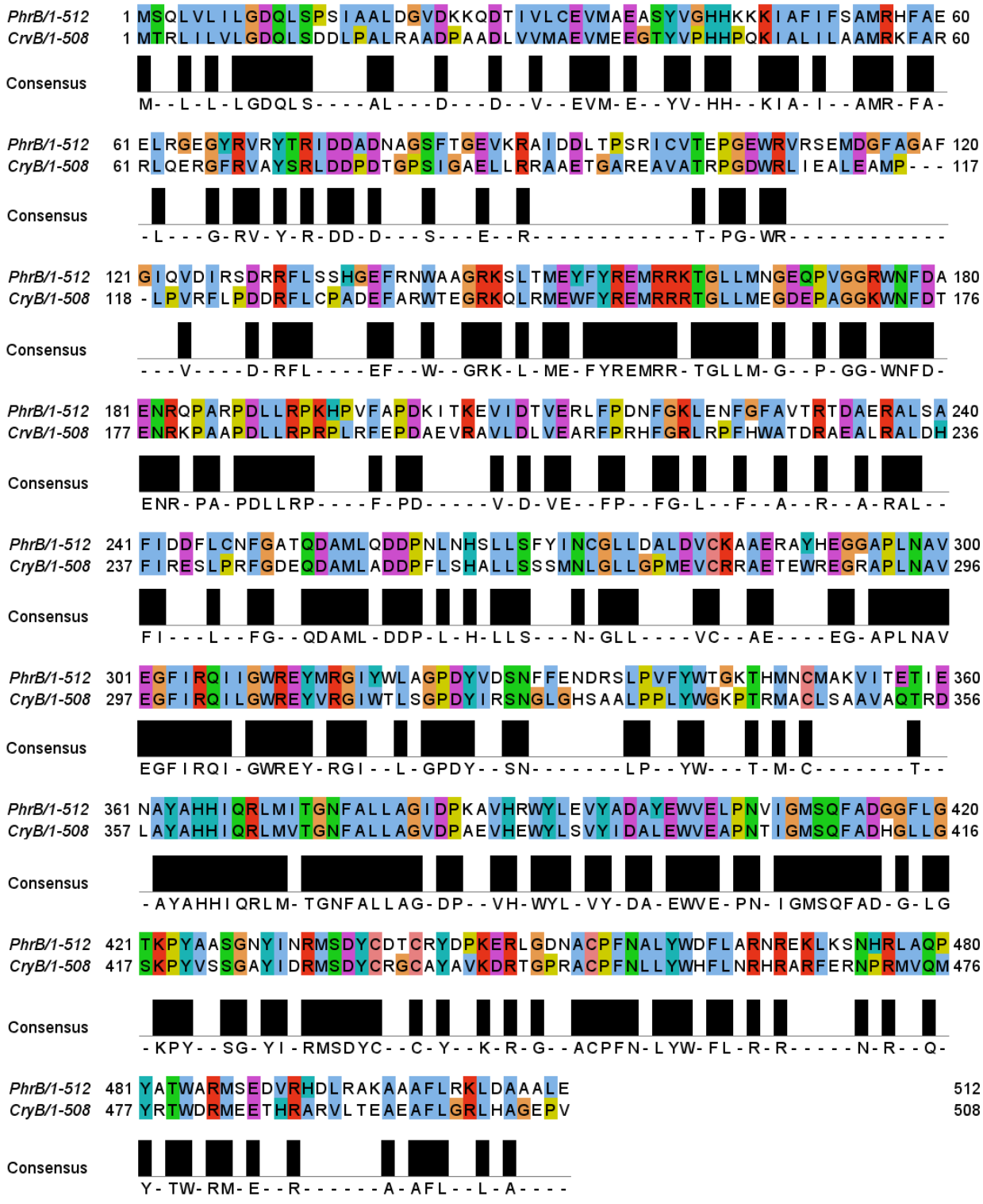


**Fig. 3.2. Sequence alignment of PhrB and CryB.** The alignment was performed with ClustalX 2.0.12 (Thompson *et al.*, 2002) and presented by Jalview 2.7. Coloring by ClustalX mode.

### 3.2.3 Crystallization of PhrB

Primary crystallization screening for PhrB yielded 31 hits (Table 3.2) within 3 days. All crystals were yellow but most of them were poorly shaped (Fig. 3.3). Only under one condition with PEG 400 as precipitant, cubic crystals were obtained. These were too small to be mounted for diffraction measurements (Fig. 3.3A). Longer growing times did not result in a significant improvement of crystal sizes.

Several clues were obtained for further optimization. The most important hint was that polyethylene glycol (PEG) is the best precipitant for PhrB crystallization. The molecular weight of commercial PEG could ranges from 400 to 20000. Potassium, lithium and sodium salts were found to be favorite additives. The pH of the crystallization buffer could be as low as 4.6 and as high as 8.0. The fact that all crystals were not in a good shape suggested that they grew too fast. Thus, decreasing protein or precipitant concentration and lowering the temperature were considered during subsequent optimization.

PEG 400, PEG 4000 and PEG 5000 were selected precipitants. Protein concentration, precipitant concentration, additive salts, pH and temperature were varied and/or combined according to a strategy called dilution method (Dunlop & Hazes, 2003). Well shaped crystals were eventually obtained from conditions that derived from all three selected PEG precipitants (Fig. 3.4, 3.5 and 3.6). However, only crystals grown in the presence of PEG 400 (Fig. 3.6) could be successfully cryocooled and mounted for data collection. These yellow rectangular PhrB crystals appeared within 3 days and grew further for 4 days. The fully grown crystals had dimensions of approximately 0.9 x 0.3 x 0.3 $mm^3$. PhrB crystals were flash frozen in liquid nitrogen after cryosoaking with 4 M trimethylamine N-oxide for a few minutes (Mueller-Dieckmann *et al.*, 2011).

**Table 3.2 The crystallization solutions that yielded crystals in the primary screening.** The protein concentration was 6 mg/mL and the temperature 298 K. Drops were set after mixing of 2 µl protein solution with 2 µl reservoir solution.

| Hits No. | Reservoir solution |
|---|---|
| 1 | 20% PEG 3350, 200 mM di-Ammonium Tartrate |
| 2 | 20% PEG 3350, 200 mM Lithium Nitrate |
| 3 | 20% PEG 3350, 200 mM Potassium Chloride |
| 4 | 20% PEG 3350, 200 mM Potassium Nitrate |
| 5 | 20% PEG 3350, 200 mM Potassium Thiocyanate |
| 6 | 20% PEG 3350, 200 mM Sodium Nitrate |
| 7 | 4 % PEG 4000, 100 mM Sodium Acetate, pH4.6 |
| 8 | 8 % PEG 4000 |
| 9 | 8 % PEG 4000,100 mM Sodium Acetate, pH4.6 |
| 10 | 18 % PEG 4000, 100 mM Sodium Acetate,pH4.6 |
| 11 | 10 % PEG 4000, 10 % 2-Propanol, 100 mM Sodium Citrate, pH5.6 |
| 12 | 10 % PEG 4000, 5 % 2-Propanol, 100 mM HEPES, pH7.5 |
| 13 | 15 % MPD, 5 % PEG 4000, 100 mM Imidazole-HCl, pH8.0 |
| 14 | 20% PEG 4000, 20% 2-Propanol, 100 mM Sodium Citrate, pH5.6 |
| 15 | 20% PEG 4000,10% 2-Propanol, 100 mM HEPES, pH7.5 |
| 16 | 20% PEG 5000, 200 mM Ammonium Formate |
| 17 | 20% PEG 5000, 200 mM Ammonium Sulfate |
| 18 | 20% PEG 5000, 200 mM di-Ammonium Tartrate |
| 19 | 20% PEG 5000, 200 mM Lithium Nitrate |
| 20 | 20% PEG 5000, 200 mM Potassium Chloride |
| 21 | 20% PEG 5000, 200 mM Potassium Thiocyanate |
| 22 | 20% PEG 5000, 200 mM Sodium Nitrate |
| 23 | 10% PEG 6000, 5% MPD, 100 mM HEPES, pH7.5 |
| 24 | 4 % PEG 8000 |
| 25 | 12 % PEG 8000, 5 % Glycerol, 100 mM Potassium Chloride |
| 26 | 15 % PEG 8000, 100 mM MES, 200 mM Sodium Acetate, pH6.5 |
| 27 | 20 % PEG 8000, 100 mM MES, 200 mM Sodium Acetate, pH6.5 |
| 28 | 8 % PEG 10000,100 mM Sodium Acetate, pH4.6 |
| 29 | 10 % PEG 20000, 100 mM MES, pH6.5 |
| 30 | 12% PEG 20000, 100 mM MES, pH6.5 |
| 31 | 15 % w/v PEG 400, 100 mM MES Sodium Salt, pH6.5 |

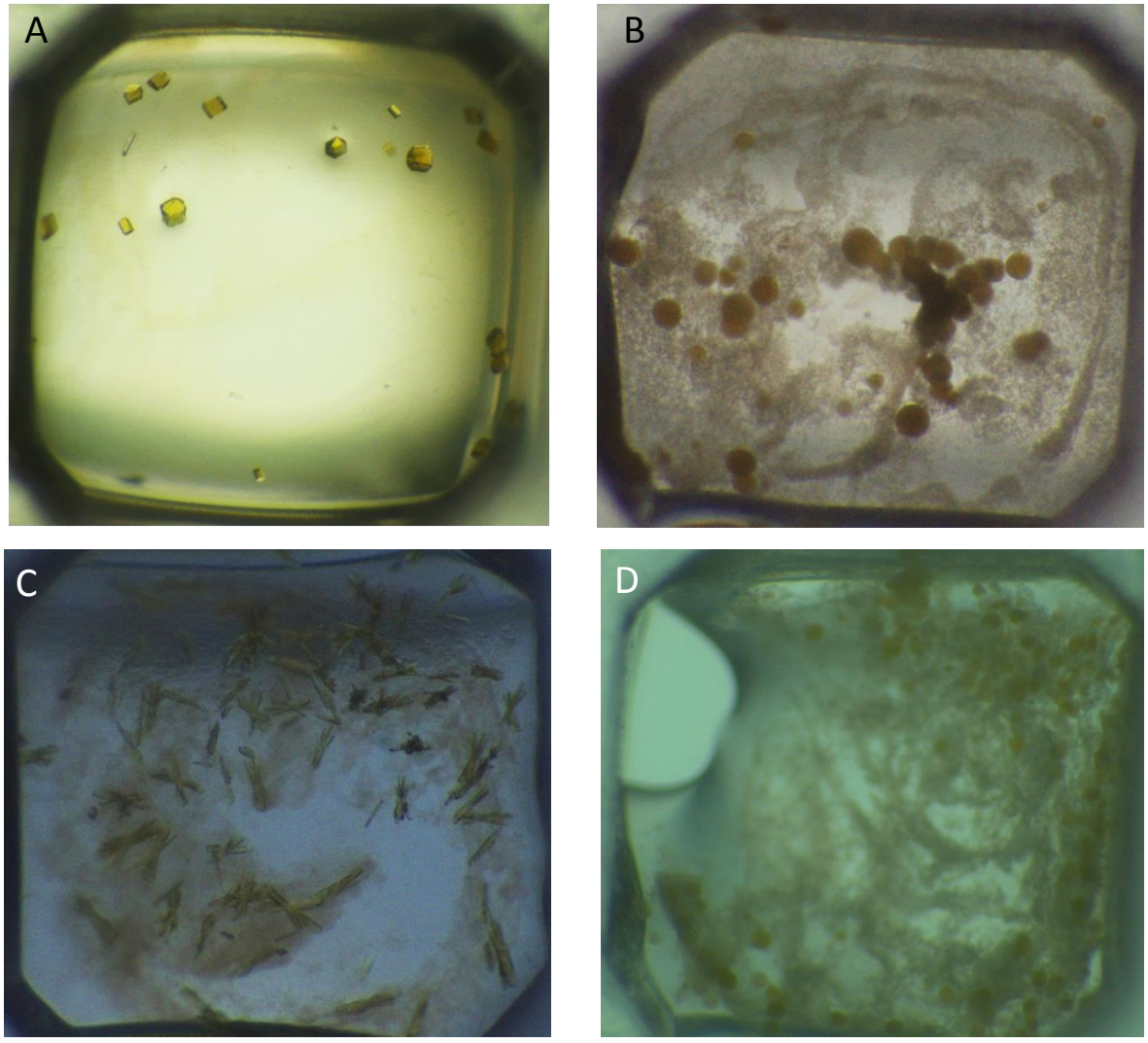


**Fig. 3.3. Example for crystals from the primary screen.** The protein concentration was 6 mg/mL and the temperature 298 K. Drops were made by mixing of 2 µl protein solution with 2 µl reservoir solution. (*A*) 15% PEG 400, pH 6.5, 100 mM MES sodium chloride. (*B*) 20% PEG 3350 200 mM Lithium Nitrate. (*C*) 8% PEG 4000, pH 4.6, 100 mM sodium acetate. (*D*) 20% PEG 5000 MME 200 mM sodium nitrate.

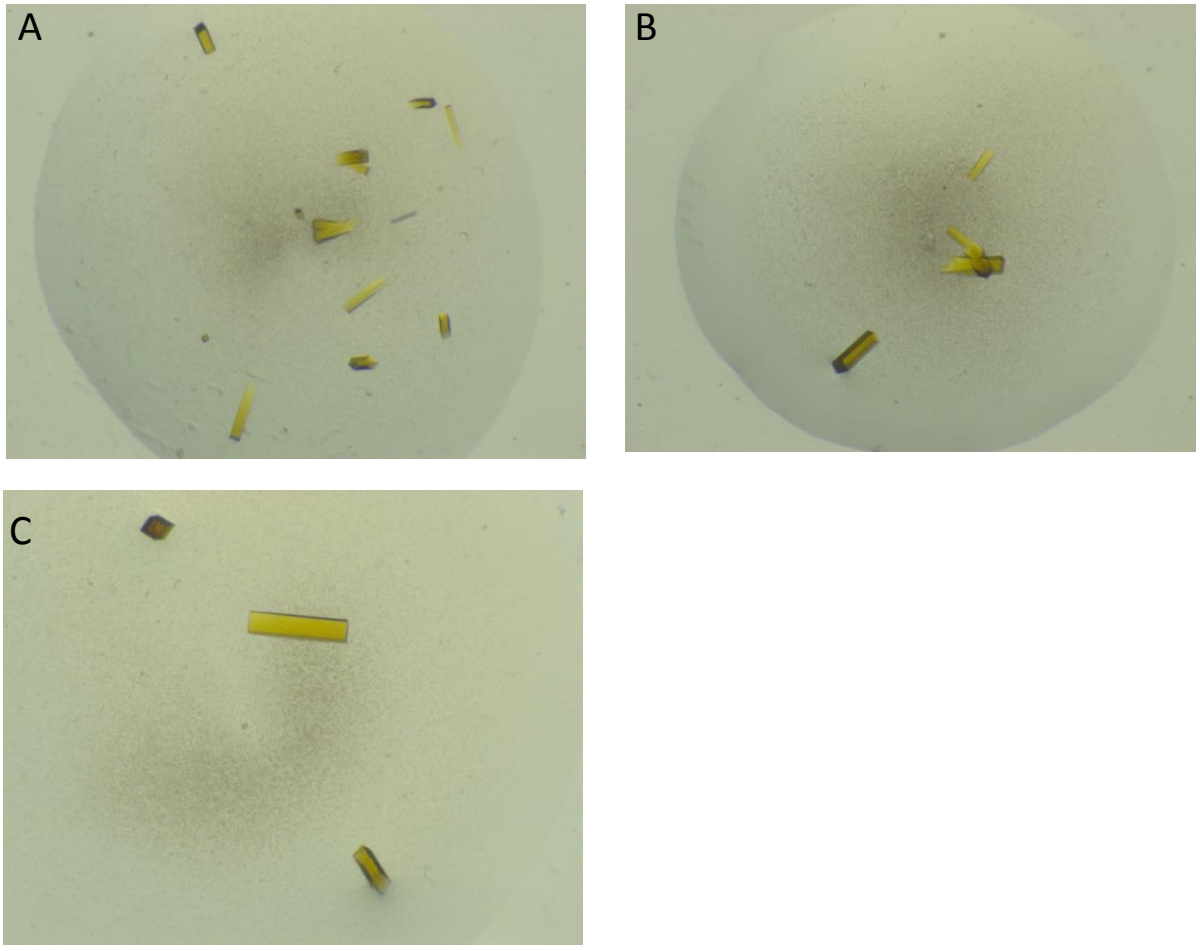


**Fig. 3.4. Examples for PhrB crystals.** Condition: 13.5% PEG 5000, pH 4.6, 200 mM sodium acetate 6 mg/mL protein, 298 K. The volume of protein solution/reservoir solution was 3 µl/2 µl (*A*), 3 µl/1.5 µl (*B*), and 5 µl/2.5 µl (*C*), respectively.

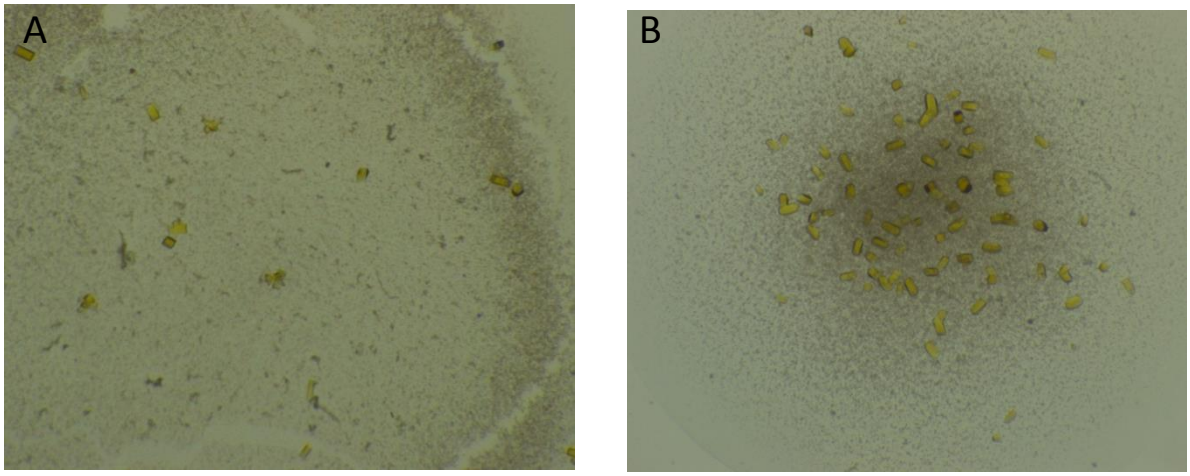


**Fig. 3.5. Examples for PhrB crystals.** Condition: 10 % PEG 4000, 10 % 2-Propanol, 100 mM sodium citrate ,pH 5.6, 6 mg/mL protein, 4 µl protein solution mixed to 4 µl reservoir solution at 277K (*A*) and 298K (*B*).

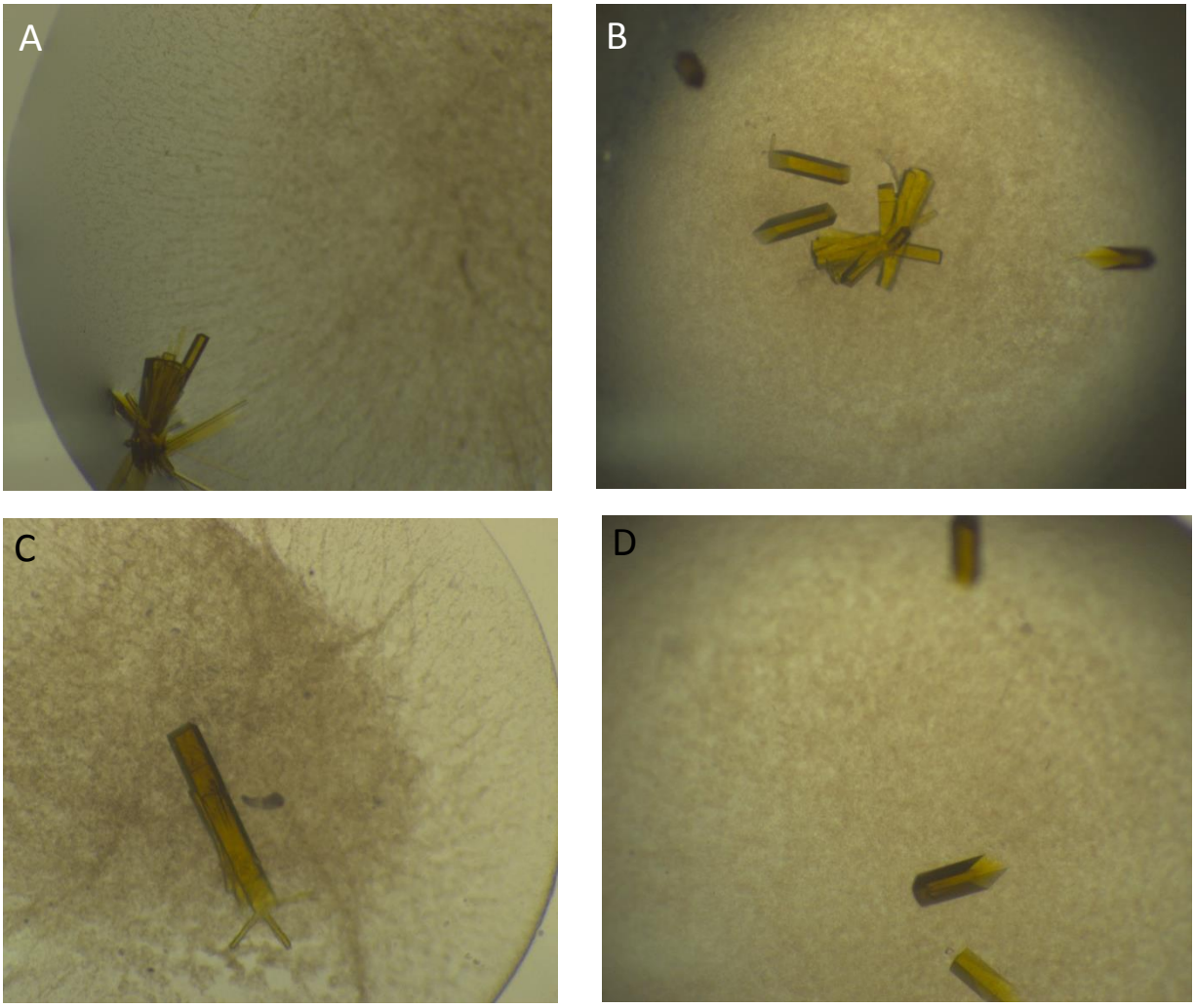


**Fig. 3.6. Crystals that grown in the presence of PEG 400.** (*A*) and (*B*) 4 % PEG 400, 100 mM MES, pH 6.0. 5 µl protein solution mixed with 5 µl reservoir solution at 298K and 289K respectively. (*C*) and (*D*) 5% PEG 400, 110 mM MES, pH 6.0, 5 µl protein solution mixed with 5 µl reservoir solution at 298K and 289K respectively.

### 3.2.4 Crystal structure of PhrB

The crystal structure of PhrB was determined at a resolution of 1.45 Å, which is the best resolution for a CPF structure so far. The presence of the [4Fe-4S] cluster was confirmed by a strong anomalous scattering signal in the X-ray diffraction data collected at the iron absorption edge (Table 3.3) and allowed phasing through the single-wavelength anomalous dispersion (SAD) strategy.

The PhrB structural model (Fig. 3.7) represents 508 of the 512 amino acids and reveals 5 ß-sheets, 21 α helices and 5 $3_{10}$ helices, which displays a typical overall fold of CPFs (Muller & Carell, 2009). All the ß-sheets and α1 to α5 form the α/ß-domain (Ser2 to Arg127), the α9 to α21 and all the $3_{10}$ helices (Arg231 to Glu512) form the α-domain. Both domains are connected via a long loop (Ser128 to Thr230). The first residue (Met1) and another 3 residues (Ala180 to Asn182) are not visible in the electron density map. The FAD chromophore assumes the prototypical U-shaped conformation. As the second organic cofactor, 6,7-dimethyl-8-ribityl-lumazine (DMRL) was identified.

The overall structure of PhrB is comparable with other members of the cryptochrome / photolyase family (CPF) (Brudler *et al.*, 2003; Fujihashi *et al.*, 2007; Hitomi *et al.*, 2009; Huang *et al.*, 2006; Kiontke *et al.*, 2011; Komori *et al.*, 2001; Maul *et al.*, 2008; Park *et al.*, 1995; Pokorny *et al.*, 2008; Tamada *et al.*, 1997). Superimpositions (Holm & Rosenstrom, 2010) yielded rmsd values between 2.8 and 3.5 Å. The rmsd values for *D. melanogaster* / *A. thaliana* (6-4) photolyases (Drome (6-4) PL and Arath (6-4) PL) were 3.2 and 3.5 Å, respectively. In phylogenetic studies, the prokaryotic and eukaryotic (6-4) photolyases appear as the most distantly related proteins within CPF (Fig. 3.8A). While our PhrB manuscript was under preparation, the structure of one of the PhrB homologs, CryB was published (Geisselbrecht *et al.*, 2012). Both structures matched closely (rmsd = 0.8 Å).

**Table 3.3. Data collection and refinement statistics.**

| | native PhrB[a] | (Fe-edge) SAD-PhrB[a] |
|---|---|---|
| **Data collection** | BESSY II, BL 14.2 | BESSY II, BL 14.2 |
| Wavelength (λ) | 0.9184 | 1.7409 |
| Space group | $P2_12_12$ | |
| Cell dimensions | | |
| *a*, *b*, *c* (Å) | 98.39, 106.79, 55.91 | |
| $\alpha, \beta, \gamma$ (°) | 90.0, 90.0, 90.0 | |
| Resolution (Å) | 72.36 - 1.45 (1.53 - 1.45)[b] | 33.46-1.95 (2.06-1.95) |
| $R_{merge}$ | 0.037 (0.377) | 0.034 (0.124) |
| $I/\sigma I$ | 22.0 (2.7) | 58.8 (10.1) |
| Completeness (%) | 95.6 (78.0) | 97.6 (91.0) |
| Redundancy | 4.4 (3.1) | 6.9 (5.6) |
| **Refinement** | | |
| Resolution (Å) | 19.95 - 1.45 | |
| No. reflections | 92395 | |
| $R_{work}$/ $R_{free}$ | 13.8 / 18.0 | |
| No. atoms / residues | | |
| (1 monomer per asymmetric unit) | | |
| Protein (PhrB) | 4267 / 512 | |
| Others: | 141 / 15 | |
| Water | 836 / 836 | |
| Flavin-adenine dinucleotide (FAD) | 53 / 1 | |
| 6,7-dimethyl-8-ribityl-lumazine (DMRL) | 23 / 1 | |

| | |
|---|---|
| Iron/sulfur cluster ([4Fe-4S]) | 8 / 1 |
| Propane-1,2,3-triol | 12 / 2 |
| Sodium | 1 / 1 |
| Mean B-factor (all atoms; $Å^2$) | 19.1 |
| R.m.s. deviations | |
| Bond lengths (Å) | 0.011 |
| Bond angles (°) | 1.517 |
| Rampage plot[c] | |
| % most favored | 98.8 |
| allowed | 1.2 |

Note: [a]One crystal was used. [b]Highest resolution shell is shown in parentheses. [c]As defined in the program RAMPAGE (Lovell *et al.*, 2003).

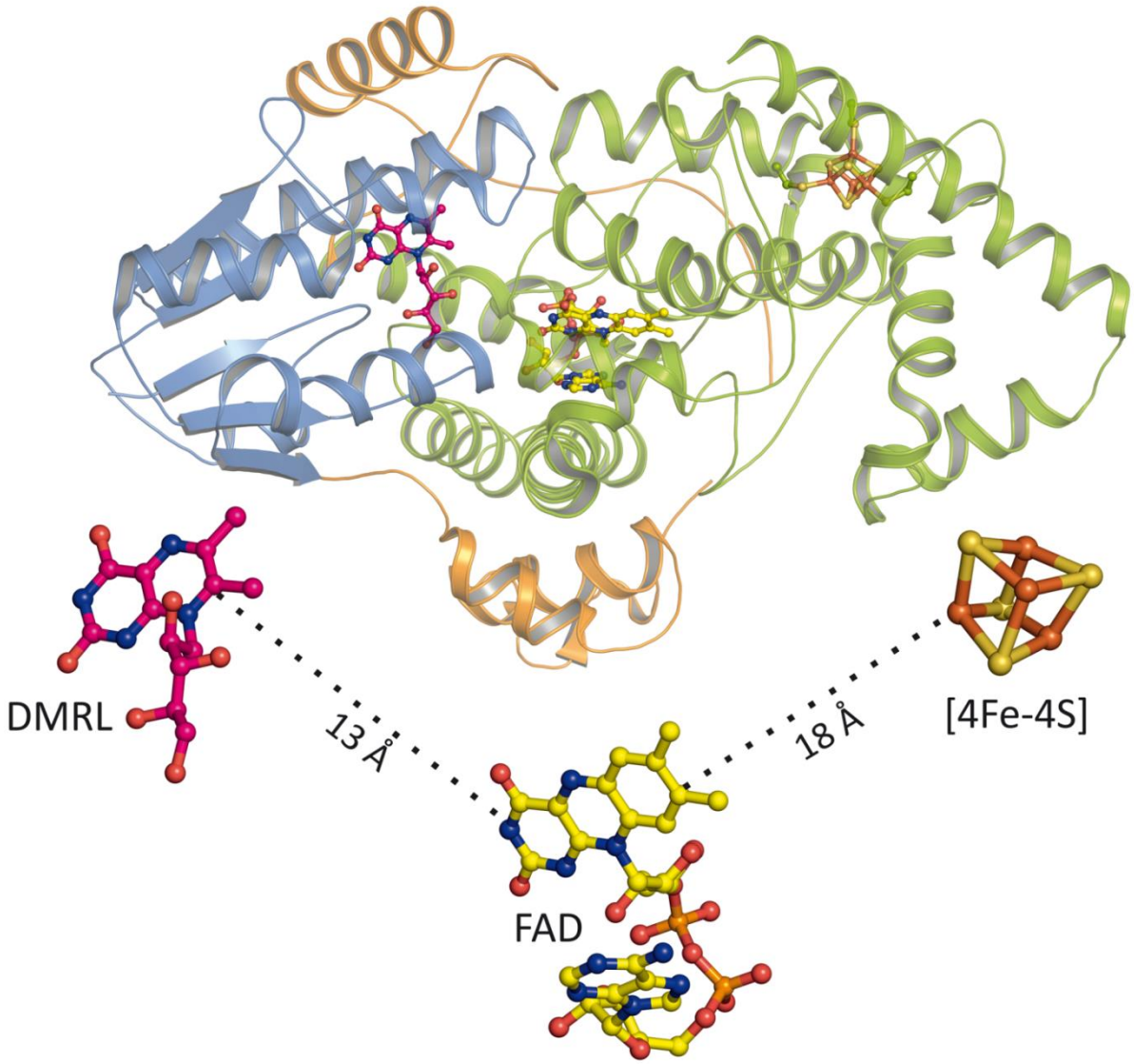


**Fig. 3.7. Overall structure and cofactor arrangement with distances of PhrB.** The α/β-domain is presented in blue and the helical-domain in green; the long inter-domain linker is orange. The cofactors DMRL, FAD and the [4Fe-4S] cluster are illustrated in the ball-and-sticks mode. In the lower part of the figure, the cofactors are shown in an enlarged view.

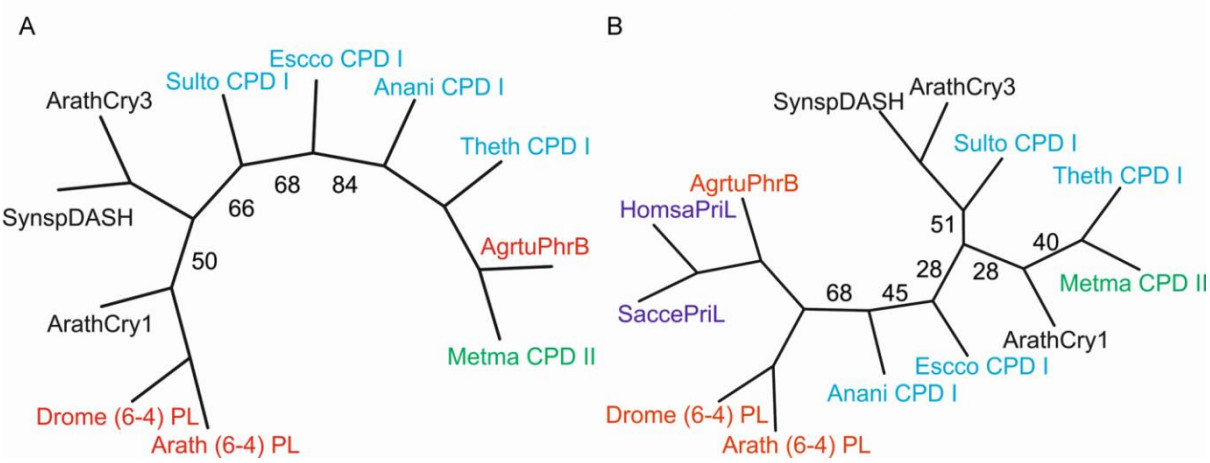


**Fig. 3.8. Phylogenetic trees based on structural alignments.** Both trees were constructed with the PROTPARS program of the PHIYLIP program package (Felsenstein, 2005). Bootstrap values (in %) are printed along the branches if < 95 %. (*A*) Analysis based on a structural alignment of full length CPFs. The weak relationship between PhrB and eukaryotic (6-4) photolyases is also obtained with maximum likelihood and distance based algorithms. (*B*) Analysis based on a structural alignment of the protein fold common to PriL and CPFs, core region (amino acid 348 to 398 of PhrB). PhrB and PriL are placed in one monophyletic group, a result that is also obtained with distance based and Bayesian algorithms, and with the same algorithms and a longer homologous stretch (amino acids 333 to 465). Color code: red, (6-4) photolyases; cyan, CPD class I photolyases; green, CPD class II photolyases; black, plant cryptochromes and Cry-DASH proteins; blue, Primase PriL subunits. AgrtuPhrB: (6-4) photolyase of *A. tumefaciens,* PDB entry 4DJA; Anani CPD I: CPD class I photolyase of *Anacystis nidulans*, PDB code 1OWL; Arath-Cry1: cryptochrome 1 of *A. thaliana,* PDB entry 1U3C; Arath-Cry3: cryptochrome 3 (Cry-DASH) of *Arabidopsis thaliana,* PDB entry 2J4D; Arath (6-4) PL: (6-4) photolyase of *Arabidopsis thaliana,* PDB entry 3FY4; Drome (6-4) PL, (6-4) photolyase of *D. melanogaster*, PDB entry 3CVW; Escco CPD I: CPD class I photolyase of *Escherichia coli,* PDB entry 1DNP; HomsaPriL: DNA primase large subunit of *Homo sapiens,* PDB entry 3Q36; Metma CPD II: CPD class II photolyase of *Methanosarcina mazei,* PDB entry 2XRZ; SaccePriL: DNA primase large subunit of *Saccharomyces cerevisiae,* PDB entry 3LGB;

Sulto CPD I: CPD class I photolyase of PDB entry *Sulfolobus tokodaii,* PDB entry 2E0I; SynspDASH: Cry-DASH of *Synechocystis* PCC 6803, PDB entry 1NP7; Theth CPD I: *Thermus thermophilus* CPD class I photolyase, PDB entry 1IQR.

### 3.2.5 The His365-His366-X-X-Arg369 motif

To investigate specific structural features for (6-4) photolyases, we performed a structure-based sequence alignment (Holm & Rosenstrom, 2010) with representative CPF structures from PDB. Major variations could be clearly seen in the inter-domain linker region and C-terminal tail as expected. This alignment also allowed us to conclude that the former proposed functional motifs of Arath (6-4) PL (Hitomi *et al.*, 2009), including the 3D adjacent phosphate-binding (Asp-235–Ala-256) and PKL protrusion motifs (Tyr-282–Leu-300), or the sulfur loop (Met-318–Cys-324) do not exist in PhrB and its homologs (Fig. 3.9).

Instead, the structural alignment showed that PhrB, Arath (6-4) PL and Drome (6-4) PL share 34 common residues, 10 of them are highly conserved (99%-100%) in all FeS-BCPs. Three of these amino acids cluster together in the His365-His366-X-X-Arg369 motif. These amino acids are located at the bottom of the DNA lesion's entrance channel. His365 and Arg369 are conserved in (6-4) photolyases and in many CPD homologs. His366 is conserved in FeS-BCPs, in eukaryotic (6-4) photolyases and in animal cryptochromes; CPD photolyases have an Asn at the homologous position. The His366 homolog of eukaryotic (6-4) photolyases is essential for DNA-repair; it serves as a donor/acceptor in catalytic proton transfer which accompanies the light-driven electron transfer (Li *et al.*, 2010; Maul *et al.*, 2008).

The characterization of this motif also revealed a significant difference from the eukaryotic (6-4) photolyases structure. In eukaryotic (6-4) photolyases, there is an additionally conserved His (His369 in Drome (6-4) PL) behind common Arg residue (Arg 368 in Drome (6-4) PL). Without knowing Fe-BCPs the motif of conserved amino acids in (6-4) photolyases has been proposed as His-His-X-X-Arg-His. The last His of this motif was believed essential for the (6-4) photorepair as the substitution by Met lead to a drastically decreased repair activity of Drome (6-4) PL (Maul *et al.*, 2008). However, this His is substituted by the Leu370 in PhrB.

The orientation of PhrB-His366 matches with the homologous His residue of Drome (6-4) PL suggesting functional identity. His366 is stabilized by van der Waals contacts with Leu370 and Met410 (Fig. 3.10), whereas the homologous His in eukaryotic (6-4) photolyases is stabilized by a hydrogen bonding network (Hitomi *et al.*, 2009; Maul *et al.*, 2008). Gln306, conserved in FeS-BCPs and eukaryotic (6-4) photolyases, forms hydrogen bonds to the DNA lesion (Table 3.4). CPD photolyases have a conserved Glu at the homologous positions.

**Fig. 3.9. Tertiary structure based alignment of PhrB with 10 representative CPF members.** β-sheets and α helices are highlighted by yellow and grey, respectively. The helix numbering of PhrB is given above the aligned sequences. DMRL, FAD, iron-sulfur cluster coordinating residues and electron transfer pathways of PhrB are printed in red, blue, magenta and cyan, respectively. The phosphate binding motif, the Pro-Lys-Leu (PKL) protrusion motif, the NLS Sequence and the sulfur loop of Arath (6-4)PL (Hitomi *et al.*, 2009) are highlighted by red, green, magenta and blue frames, respectively. The alignment was performed using the Daliserver (Holm & Rosenstrom, 2010) with manually corrections.

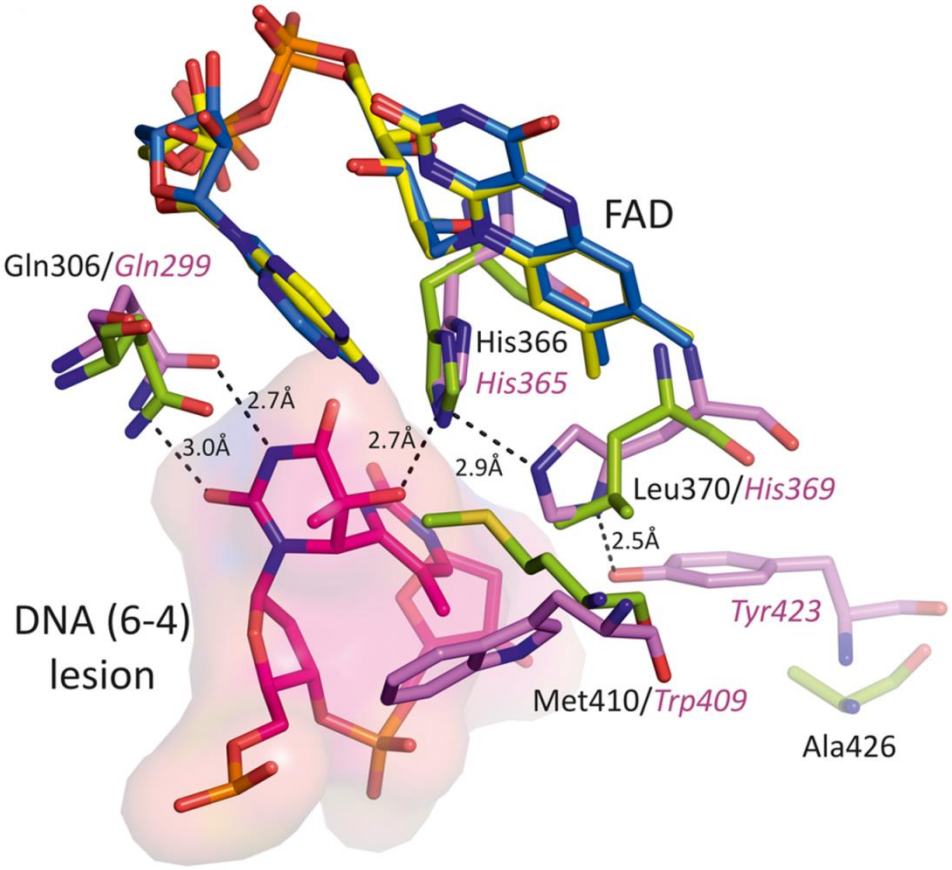


**Fig. 3.10. The active site of PhrB and Drome (6-4) PL.** Drome (6-4) PL complex with photodamaged DNA (PDB entry 3CVU, DNA in magenta, amino acids in purple and FAD in blue) and PhrB (amino acids in green and FAD in yellow) after superposition of the FAD molecules. Amino acid residues in Drome (6-4) PL which interact with the DNA lesion and the homologous residues of PhrB are shown.

**Table 3.4. Eight amino acids of PhrB that are conserved within FeS-BCPs interact with the DNA lesion.** The DNA interaction is calculated on the Hex-Server (http://hexserver.loria.fr) DNA-protein interaction model.

| Amino acids | Conservation in FeS-BCPs | Conservation in other CPFs | Contact with Damaged DNA |
|---|---|---|---|
| **His366** | 100% | Present in eukaryotic (6-4) PL and animal Cry | first (5') T in the DNA lesion |
| **Tyr424** | 100% | Not conserved | second T |
| **Gln306** | 99% | Present in Drome (6-4) PL, Glu in other CPF | first T |
| **Trp310** | 100% | Present in eukaryotic (6-4) PL and other CPF | first T |
| **Arg183** | 98% | Not present | central phosphate residue in the TT lesion |
| **Met410** | 100% | Not conserved | probably in contact with DNA lesion |
| **Leu370** | 100% | His in eukaryotic (6-4) PL | the second T |
| **Ile367** | 99% | Leu in eukaryotic (6-4) PL | the second T |

### 3.2.6 (6-4) Photoproduct binding and the inter-domain linker loop

According to a DNA-protein interaction model as calculated by the Hex-Server (http://hexserver.loria.fr) (Macindoe *et al.*, 2010), 8 amino acids of PhrB that are all conserved within FeS-BCPs interact with the DNA lesion (Table 3.4). One of them, Arg183 belongs to the inter-domain linker loop region. This loop region connecting α7 in the N-terminal α-β domain and α8 in the C-terminal α-helical domain of PhrB (Ser128 to Val229) is longer in FeS-BCPs than in other PCF members (Fig. 3.11). Surprisingly, the loop is located at a completely different position as that of other PCF members. Superimposition with the Drome (6-4) PL-DNA complex suggested that this loop interacts with DNA in PhrB. In PhrB, this loop encompasses a flexible segment of three amino acids (Ala180 to Asn182) that are not visible in the electron density map (Fig. 3. 12, A and B). These amino acids might be stabilized by the DNA to acquire a more ordered conformation.

Arg183 could change its orientation to interact via its positively charged guanidinium group by forming a salt bridge with the phosphate within the DNA lesion, thereby stabilizing the flip-out of the lesion (Fig. 3.12, A and B). In this respect, PhrB differs from prototypical CPFs, in which another Arg residue (Arg421 in Drome (6-4) PL) within the α17-α18 connecting loop stabilizes the flip out of the lesion (Fig. 3.12, C and D) (Kiontke *et al.*, 2011; Mees *et al.*, 2004; Pokorny *et al.*, 2008). This Arg is missing in FeS-BCPs. We therefore suggest that FeS-BCPs bind UV-damaged DNA in a mode significantly different from prototypical photolyases.

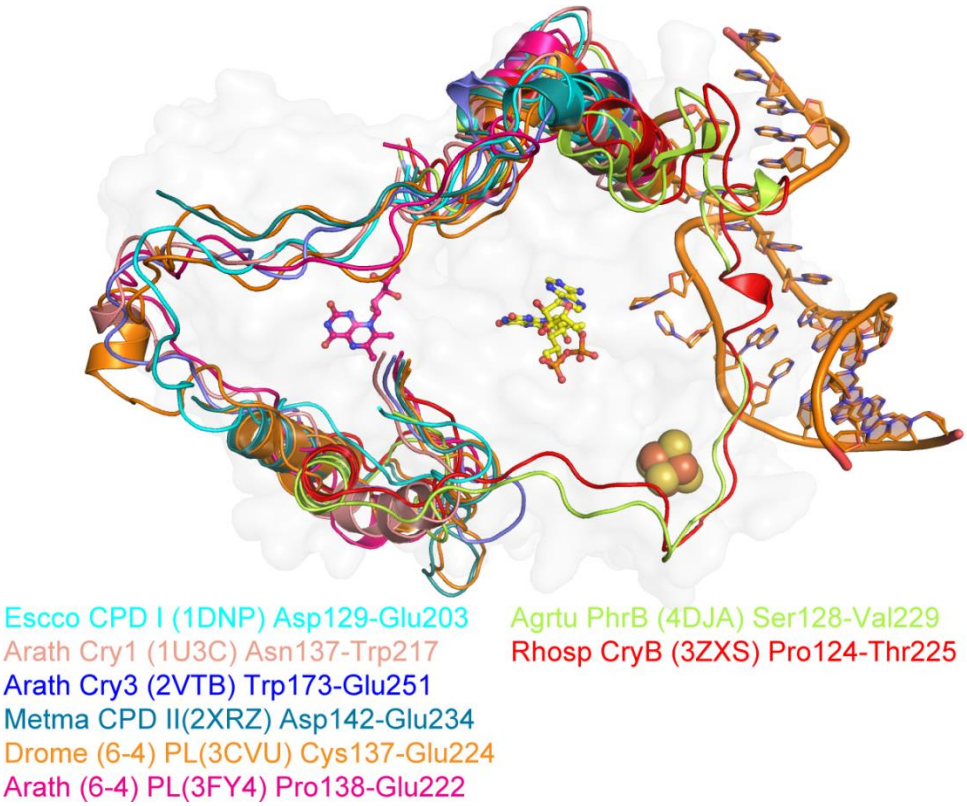


**Fig. 3.11. Comparison of the linker region of PhrB and other PCFs.** The DNA is from the 3CVU structure. The PhrB surface is shown in transparent mode. FAD, and DMRL are shown in the ball-and- sticks mode with yellow and magenta color respectively. The iron-sulfur cluster of PhrB is shown in the spheres mode.

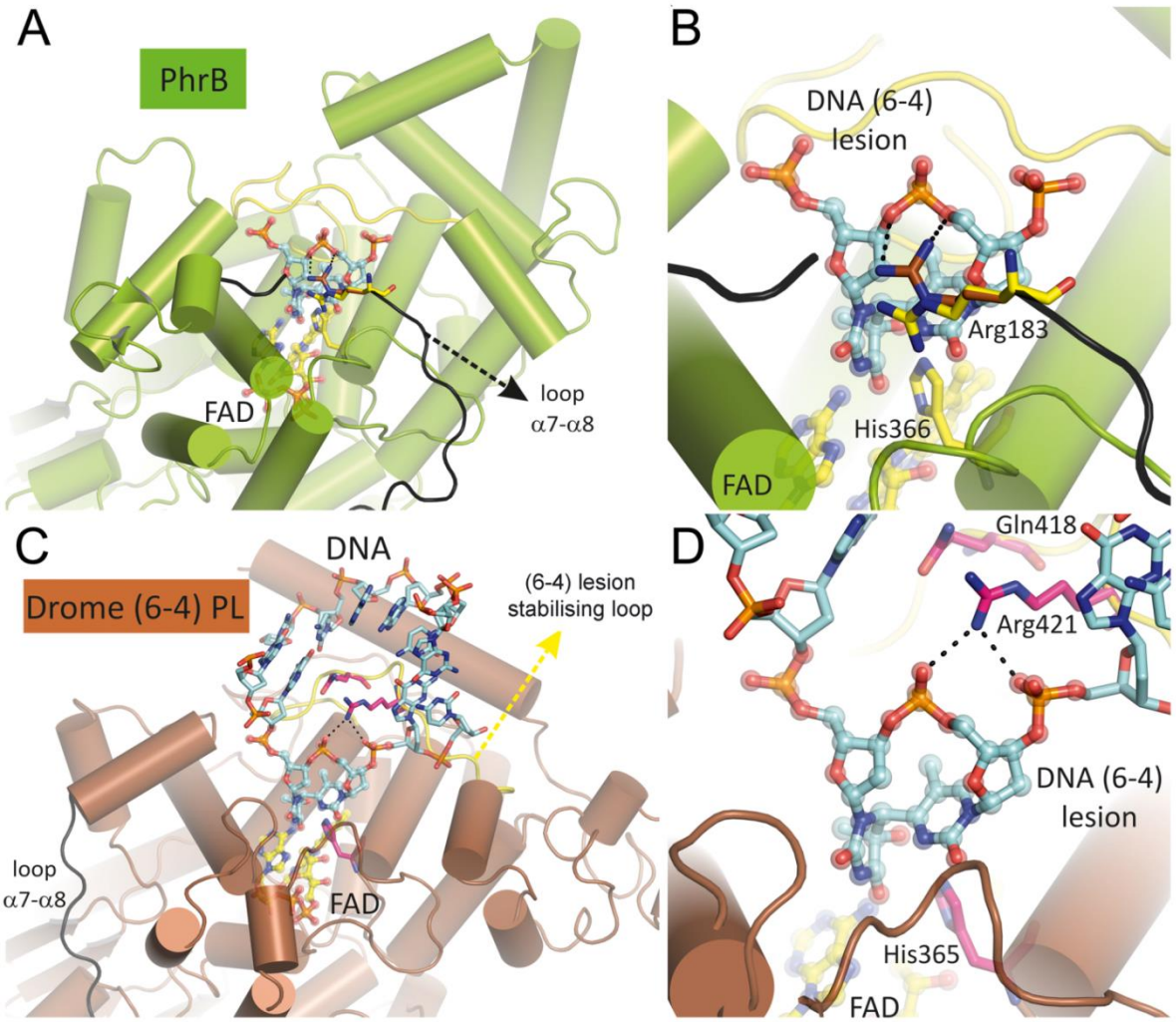


**Fig. 3.12. Potential role of the PhrB α7-α8 loop region in binding of the (6-4) DNA lesion in comparison with Drome (6-4) PL.** PhrB and Drome (6-4) PL are illustrated in ribbon representation and α -helices are shown as cylinders in green and brown, respectively. The α7-α8 loop is drawn in black, the (6-4) lesion stabilizing loop of Drome (6-4) PL is drawn in yellow. This loop is located between α17 and α18 and stabilizes the flip-out of the (6-4) photoproduct by inserting a pair of amino acid side chains (Gln418 and Arg421) into the resulting gap in the double-stranded DNA oligomer. Numberings according to PhrB. (*A*) The DNA fragment with the (6-4) lesion (shown as ball-and-sticks) from 3CVU superimposed onto the PhrB structure. (*B*) A close up view of (A). Two side chain conformations shown for Arg183 (as sticks), the original one as seen in the crystal structure (yellow) and a second one (brown) being a rotamer which is in close contact (< 4Å) with the central phosphate as

indicated by the dotted lines. (*C*) Drome (6-4) PL complex with photodamaged DNA oligomer. The DNA is drawn as sticks with the (6-4) lesion highlighted as ball-and-sticks (PDB entry 3CVU). Only a portion of the DNA oligomer is shown for clarity. (*D*) A close up view of (C). Arg421 (shown as sticks in purple) forms a salt bridge with the DNA (6-4) lesion (shown as ball-and-sticks).

### 3.2.7 C-terminal extension of PhrB

Prototypical photolyases have a common set of helices corresponding to α1-α19 in PhrB, and the last common helix α19 forms a part of the DNA binding pocket. Arath (6-4) PL has an additional α-helix (Hitomi *et al.*, 2009), corresponding to α20, as C-terminal extension. This helix forms an antiparallel two-helix bundle with α19 and (assuming a prototypical binding mode) faces away from bound DNA. The C-terminus of PhrB is extended by two additional helices α20 and α21 (Fig. 3.13), of which α21 displays sequence similarity with the C-terminal helix of Arath (6-4) PL (Fig. 3.9). Sequence alignments show that α20 and α21 are present in all FeS-BCPs. In a complex compatible with a prototypical binding mode helix α20 of PhrB interacts with the DNA, whereas α21 faces away from the DNA. The surface region of α20 displays a positive charge (Fig. 3.14), consistent with its predicted role in DNA binding.

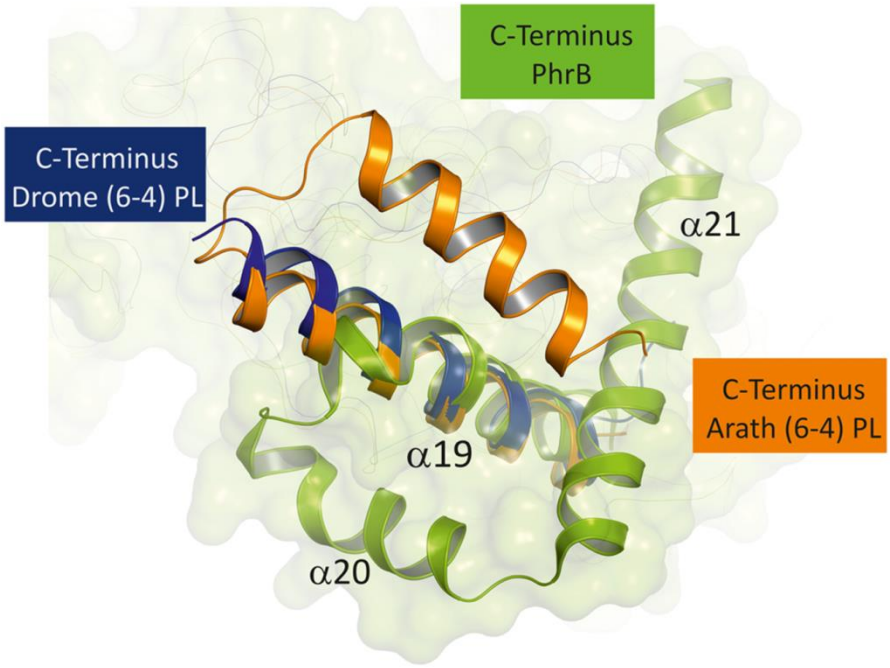


**Fig. 3.13. Comparison of C-terminal helices in superimposed (6-4) photolyases.** Drome (6-4) PL and Arath (6-4) PL are illustrated in ribbon representation in blue and orange, respectively; PhrB is drawn in ribbon representation and transparent surface in green. Helix α19 of PhrB is common to all three photolyases, Drome (6-4) PL ends with this helix. Arath (6-4) PL has one additional helix and PhrB has two additional helices, α20 and α21.

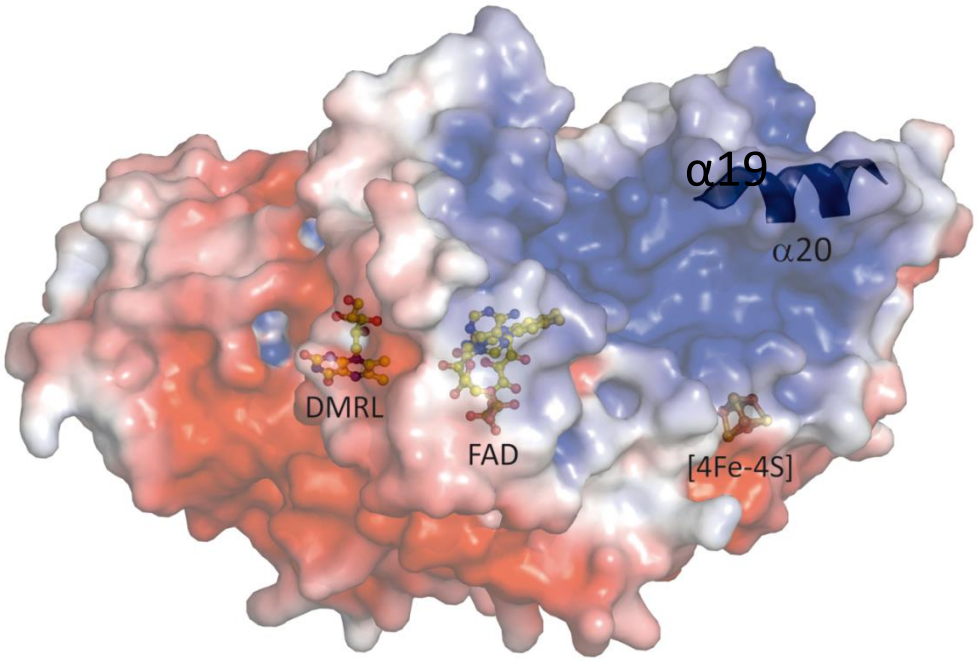


**Fig. 3.14. Electrostatic surface representation of PhrB.** Electrostatic surface potentials were calculated using the program APBS with the non-linear Poisson-Boltzmann equation and contoured at +/- 5kT/e. Negatively and positively charged surface areas are coloured in red and blue, respectively. The cofactors DMRL, FAD and Fe-S cluster are illustrated in ball-and-sticks representation. The DNA binding cavity is characterized by its positive charge; the helix α20 of the C-terminal extension with its positive surface charge appears to delimit this cavity.

### 3.2.8 FAD binding and photoreduction

In PhrB, the transition of FAD from the oxidized to the enzymatically active reduced form has been experimentally revealed as a blue light dependent process (Oberpichler *et al.*, 2011). The FAD binding site of PhrB is in many respects similar to other PCF members. In prototypical photolyases the side chain of an invariant Asn interacts with the N5 atom of the isoalloxazine ring, this Asn stabilizes protonated N5 during photoreduction and catalysis (Kiontke *et al.*, 2011; Xu *et al.*, 2008). In plant cryptochromes, the homologous amino acid is Asp, which interacts with the N5 atom and prevents the formation of $FADH^{-}$ during photoreduction by electrostatic repulsion (Banerjee *et al.*, 2007). In PhrB, a buried water molecule forms a hydrogen bond to the N5 atom (Fig. 3.15). Interestingly, Glu403 in PhrB, which corresponds to the conserved Asn of other photolyases (Asn379 in *E. coli* CPD photolyase) or Asp of plant cryptochromes faces away from the FAD's N5 atom. The same FAD binding mode is found in CryB. Given that the essential residues involved in FAD binding are highly conserved (Fig. 3.16), it is reasonable to expand this FAD binding mode to other FeS-BCPs.

In prototypical photolyases, photoreduction involves electron flow via a cascade of three conserved Trp residues to the deeply buried catalytic FAD(Park *et al.*, 1995) (Maul *et al.*, 2008),, and an alternative route has been shown for class II photolyases (Kiontke *et al.*, 2011). These Trp residues are missing in PhrB and its homologs (Oberpichler *et al.*, 2011) The PhrB structure reveals a spatially distinct electron transfer pathway for FeS-BCPs leading from the surface to Trp342 via Trp390 and Tyr391 to FAD (Fig. 3.17). Edge to edge distances for the aromatic ring systems are 3.9 Å, 4.0 Å and 4.1 Å, respectively; these residues are conserved to 84%, 99% and 70% among FeS-BCPs, respectively. Photoreduction of $FADH^{-}$ occurs by stepwise electron transfer (Brettel & Byrdin, 2010; Oberpichler *et al.*, 2011). The oxidation of Trp390 by Tyr391, which would follow the electron transfer from Tyr391 to the electronically excited $FADH^{*}$, is counterintuitive since the standard redox potential is higher for the one electron oxidation of tryptophan than that of tyrosine (Harriman, 1987). However, we speculate that Tyr391 is involved in the cascade of electron transfer reactions because the structure shows that there is no nearby base which after oxidation of the tyrosine could accept a proton from the highly acidic phenoxyl radical cation ($TyrOH^{\bullet+}$) which is even more oxidizing than a tryptophan radical cation.

The electron transfer pathway of PhrB was further investigated via site-direct-mutagenesis by Janine Wesslowski (Unpublished data). The Y391F showed the same light induced photoreduction as the wild type protein, whereas W390F, W342F and W390F/Y391F did not undergo spectral changes upon irradiation. Thus, only W390 and W342 are required for photoreduction. In the Y391F mutant, electrons flow probably directly from W390 to FAD over a distance of 7.9 Å. The same pathway could be realized in the wild type, although an involvement of Y391 can presently not be excluded. These results are consistent to the analogous studies on CryB (Geisselbrecht *et al.*, 2012), suggesting that the electron transfer mode of photoreduction is highly conserved in FeS-BCPs as indicated by their amino acids conservation.

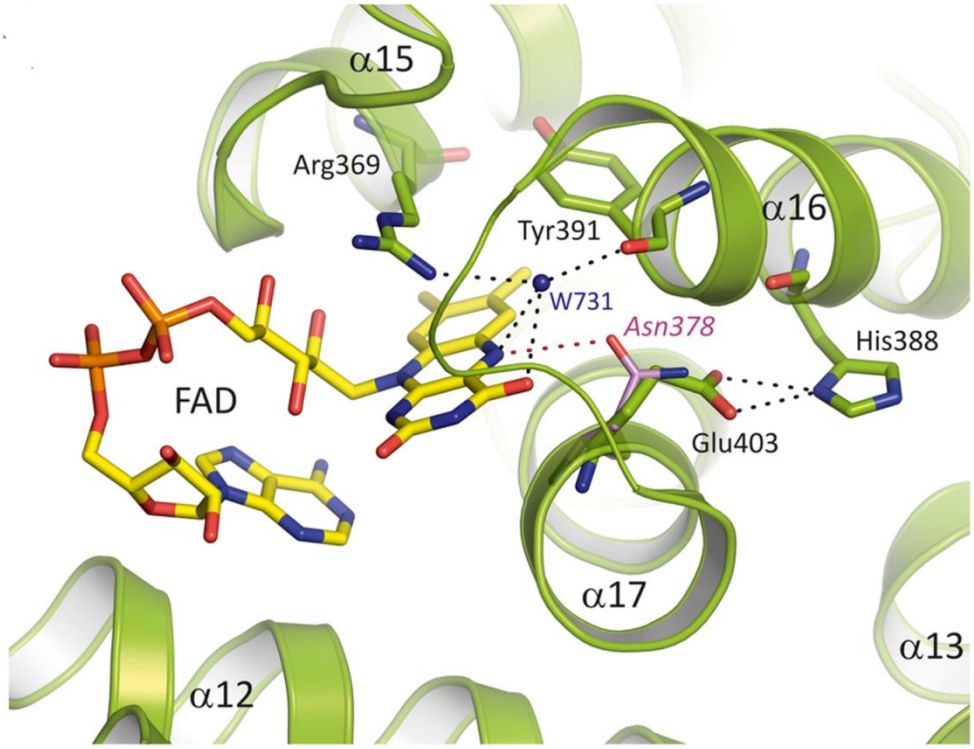


**Fig. 3.15. Water W731 stabilization of FAD.** Water W731 of PhrB (green) forms a potential hydrogen bond with the N5 atom of FAD (C atoms in yellow), replacing the role of Asn378 of the class I CPD photolyase from *E. coli* (PDB entry 1DNP), drawn in purple. Its PhrB homolog, Glu403, faces away from FAD and interacts with His388.

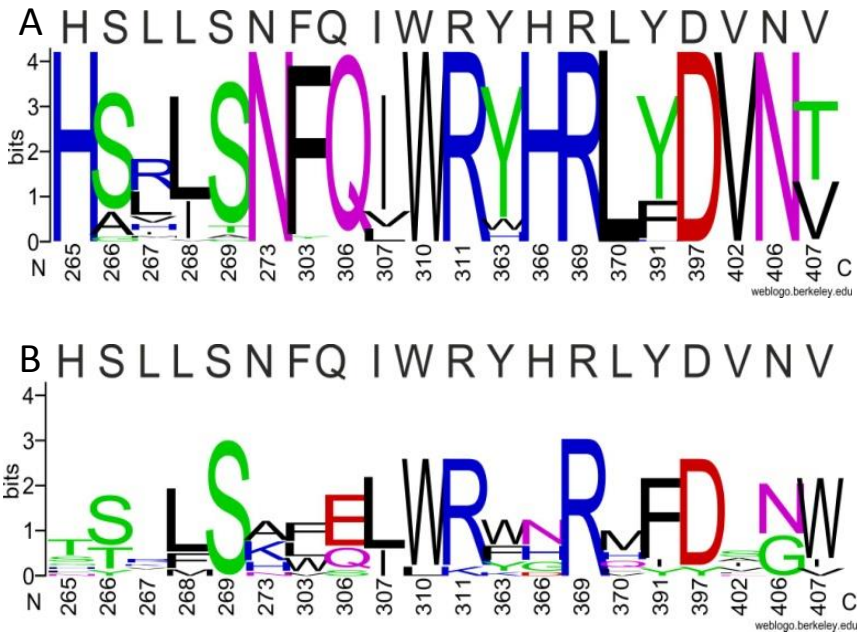


**Fig. 3.16. Compare of the conservation of FAD interacting residues in FeS-BCP and non-FeS-BCP members.** (*A*) A weblogo presentation (Crooks *et al.*, 2004) of the FAD interacting residues of 462 PhrB homologs. The sequence alignment was performed with ClustalX 2.0.12**.** (*B*) A weblogo presentation of the structural counterpart of the PhrB FAD interacting residues of 10 representative non-FeS-BCP members that structures available from PDB. The alignment that used is same as that in Fig. 3.9.

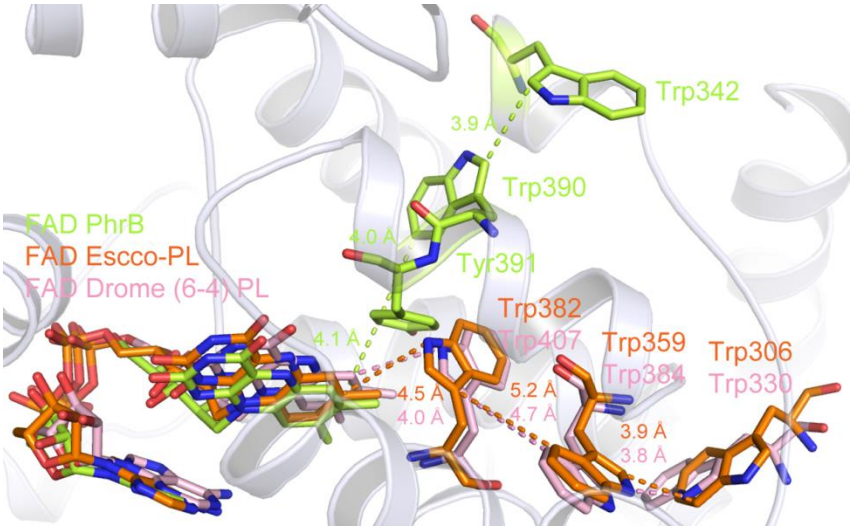


**Fig. 3.17. Compare of electron pathway in PhrB, Escco-PL (1DNP) and Drome (6-4) PL (3CVU). PhrB is illustrated in ribbon representation.**

### 3.2.9 DMRL binding

The electron density of the other organic cofactor, which is bound to amino acids in the N-terminal Rossman fold, does not fit with antenna chromophores found in photolyases or cryptochromes (Chaves *et al.*, 2011), but could be fitted to 6,7-dimethyl-8-ribityl lumazine (DMRL) (Fig. 3.18). We confirmed the identity of this cofactor by HPLC. In the elution profile, the major peak was found at the same position as that of synthetic DMRL and spectral properties of both compounds are identical (Fig. 3.19). DMRL is the last intermediate in the biosynthesis pathway of riboflavin, the precursor of FMN and FAD (Fischer & Bacher, 2011).

The binding sites for the dicyclic DMRL in PhrB is comparable with that for 8-HDF, FAD or FMN in other photolyases (Fig. 3.20) (Fujihashi *et al.*, 2007; Klar *et al.*, 2006; Tamada *et al.*, 1997). Because of the same binding site and the similar pi system orientation DMRL serves equally well as antenna chromophore in PhrB as do the other chromophores in other photolyases. MTHF antenna chromophores have a different binding location (Chaves *et al.*, 2011).

In the DMRL chromophore pocket, the lumazine ring system forms hydrogen bonds only with the polypeptide backbone and a water molecule, and is in van der Waals contact with four amino acid residues. All but one hydroxyl group of the cofactor's side chain form hydrogen bonds with the protein environment (Fig. 3.18). His43 of PhrB prevents binding of tricyclic cofactors by steric hindrance, and is in van der Waals contact with the two methyl groups of DMRL (Fig. 3.18). This residue is 67% conserved in FeS-BCPs (Fig. 3.21), and is replaced by bulky Tyr and Phe residues in other FeS-BCPs. We therefore predict that none of these proteins binds a tricyclic antenna chromophore which is in line with the fact that most of the DMRL interacting residues are highly conserved in the FeS-BCP family (Fig. 3.21).

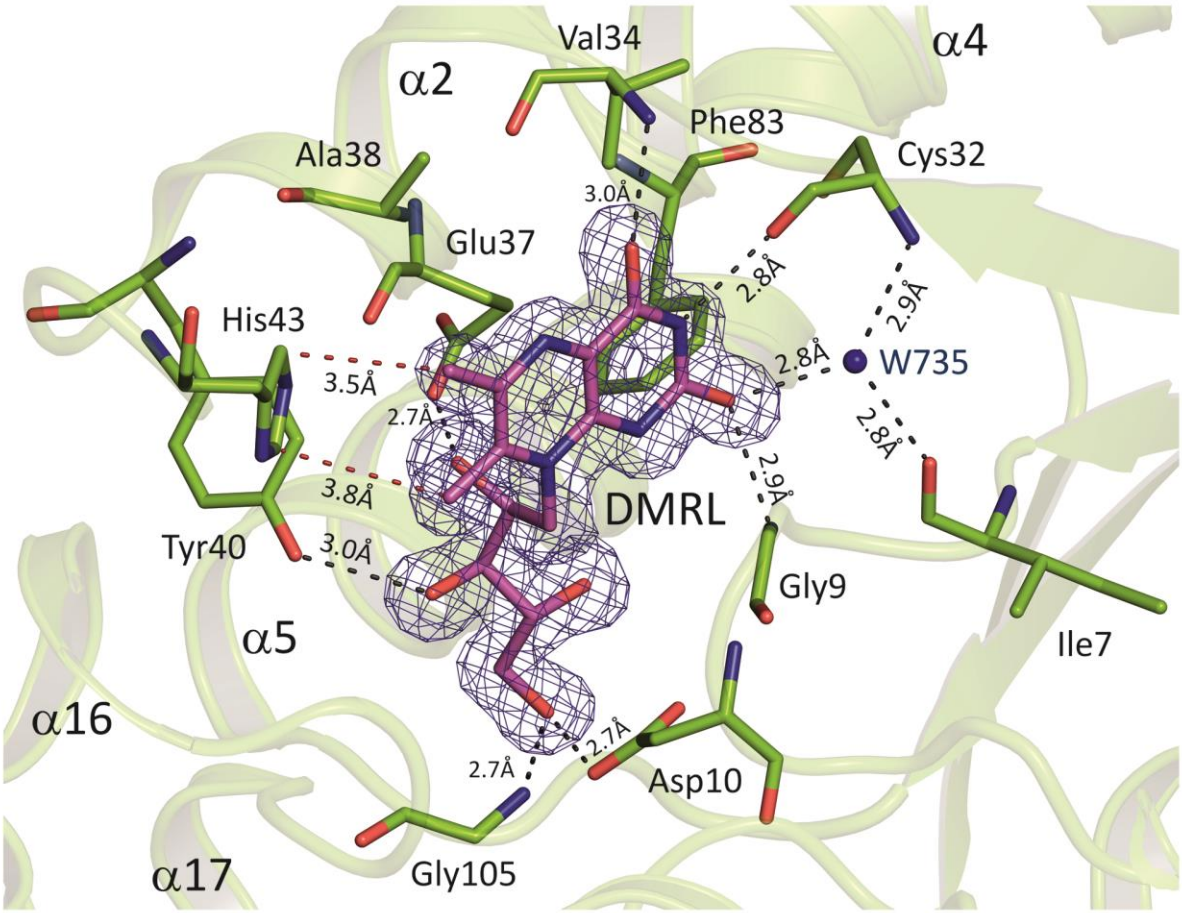


**Fig. 3.18. DMRL binding site.** (*A*) DMRL chromophore with $\sigma_A$-weighted $2F_o$-$F_c$ electron density map contoured at 1.0 σ (blue mesh) and residues in potential hydrogen bonding or van der Waals contacts (Ile51 is omitted for clarity). Hydrogen bonds between DMRL and surrounding amino acids shown by black dotted lines, van der Waals contacts shown in red.

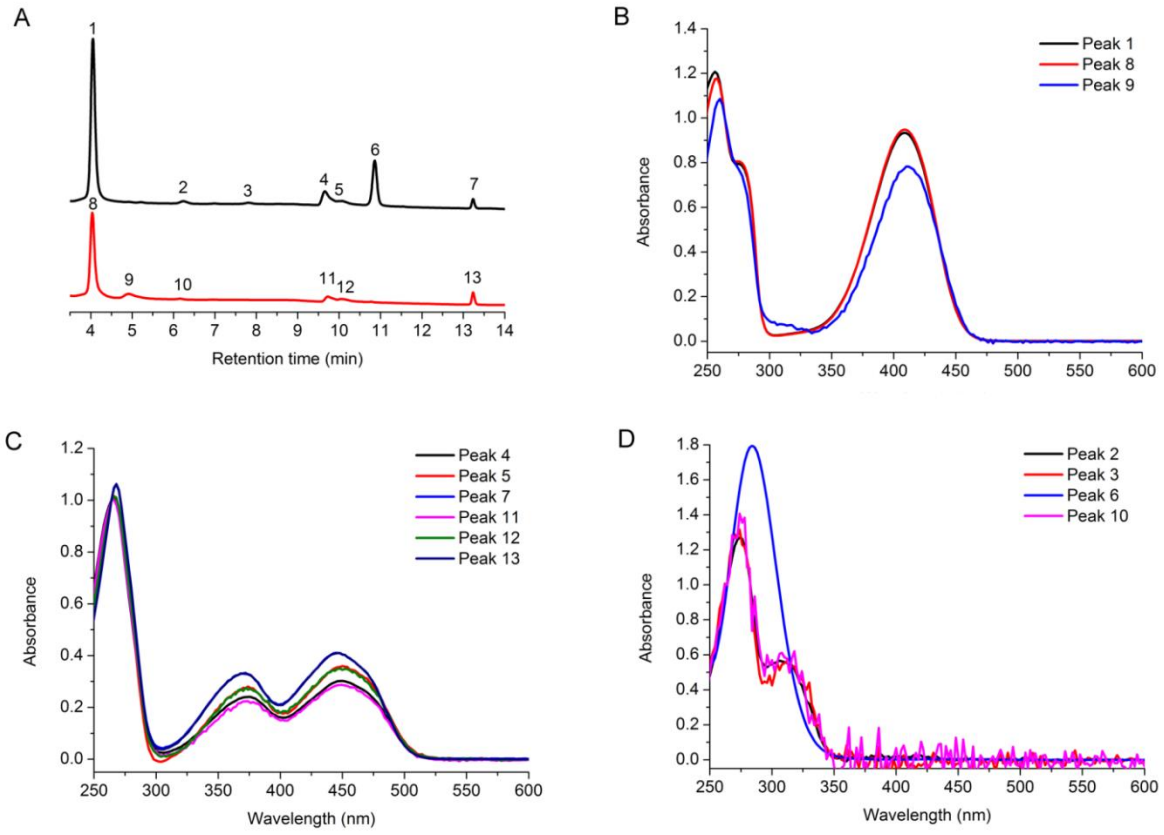


**Fig. 3.19. HPLC and UV-Vis verification of the 6, 7-dimethyl 8-ribityl-lumazine (DMRL) antenna chromophore.** (*A*) HPLC profiles, released PhrB cofactors (black line) and mixture of authentic FAD and DMRL (red line). Peaks 1 and 8 and 9 refer to DMRL. (*B*) UV-Vis spectra of DMRL fractions (HPLC peaks 1, 8 and 9). (*C*) UV-Vis spectra of FAD-containing fractions (HPLC peaks 4, 5, 7, 11, 12 and 13). (*D*) UV-Vis spectra of UV-absorbing compounds of unknown identity (HPLC peaks 2, 3, 6 and 10).

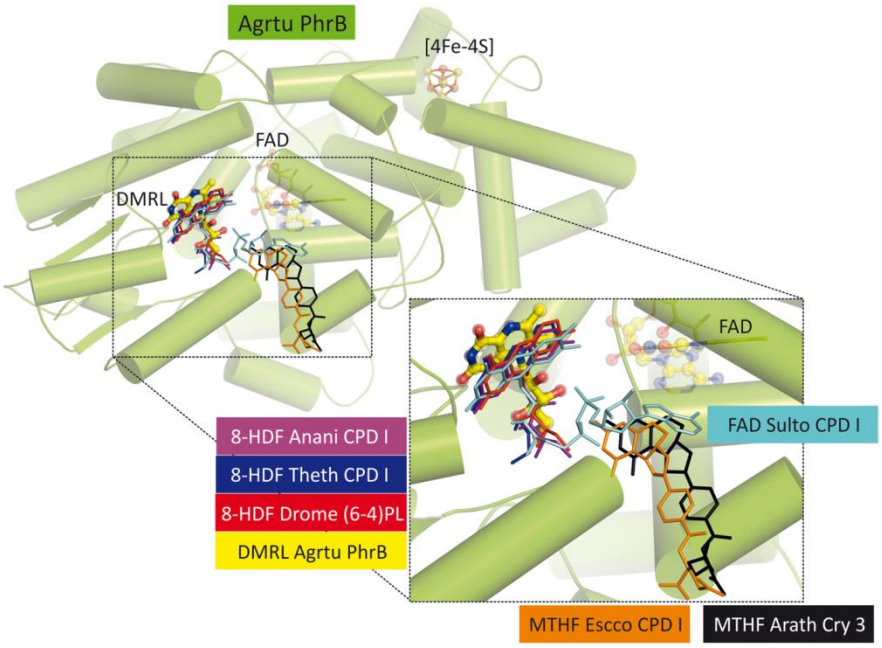


**Fig. 3.20**. **Antenna chromophores of photolyases superimposed onto the PhrB structure.** Overall view and close up of antenna chromophores. PhrB is illustrated in ribbon representation and α-helices of PhrB are shown as cylinders (green), FAD and the DMRL antenna of PhrB are drawn as ball-and-sticks, the corresponding antenna chromophores are shown in stick representation. The same binding pocket is used by 8-HDF in *Anacystis nidulans* CPD photolyase (PDB entry 1TEZ), 8-HDF in *Thermus thermophilus* CPD photolyase (PDB entry 2J07), 8-HDF in *D. melanogaster* (6-4) photolyase (soaked chromophore, PDB entry 3CVV) or the isoalloxazine ring of FAD in *Sulfolobus tokodaii* CPD photolyase (PDB entry 2E0I). A different binding pocket is used by MTHF in *Escherichia coli* CPD photolyase (PDB entry 1DNP) or MTHF in *Arabidopsis thaliana* Cry3 (PDB entry 2VTB).

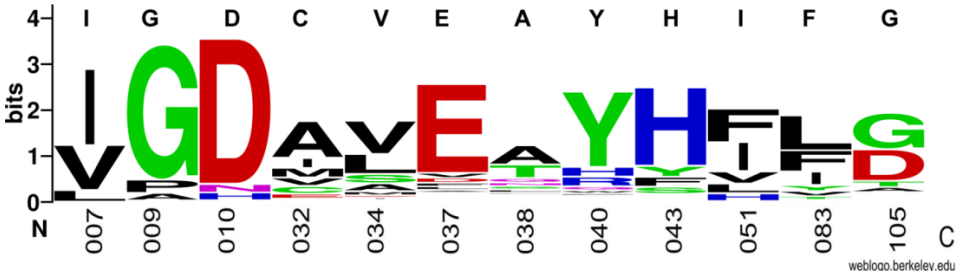


**Fig. 3.21. Conserved amino acids of DMRL interacting residues in the 462 FeS-BCP members.**

### 3.2.10 The [4Fe-4S] cluster of PhrB

The [4Fe-4S] cluster of PhrB is coordinated by Cys350, Cys438, Cys441 and Cys453. Not all FeS-BCPs have an FeS cluster at the corresponding position, because this set of Cys residues is conserved in only 89% of the family members (Fig. 3.22C). By contrast, among the sequences of non-FeS-BCP PCF members that structure available from PDB, we found no Cys residues at the homologous positions (Fig. 3.22D), suggesting that none of these proteins have an Fe-S cluster.

Structural homology between PriL-Carboxy-Terminal domain (PriL-CTD) of eukaryotic and archaeal primases, which also contain [4Fe-4S] clusters, and CPF members suggests that both groups of proteins have a common ancestor (Sauguet *et al.*, 2010). Structural similarity of variable degree can be seen between PhrB and PriL for a stretch from amino acids 348 to 434 of PhrB which comprises helices α13-α19 (Fig. 3.22A; the rmsd value for PhrB and *S. cerevisiae* PriL (PDB entry 3LGB) is 3.3 Å), the structurally most conserved core region stretches from amino acids 348 to 398. The iron sulfur clusters are found at almost the same positions in the superimposed structures, but are tilted against each other (Fig. 3.22B). This is consistent with the fact that only the first cysteines in the core regions of both structures are strictly conserved, whereas of the remaining three Cys residues, which bind the [4Fe-4S] clusters, the side chains are only located at roughly similar positions, and even the sequential order of corresponding cysteines differs (PriL: 336/417/434/474; PhrB: 350/454/438/441). We therefore assume that outside the core region the protein fold diverged and the coordinating Cys residues were exchanged during evolution.

The function of Fe-S clusters in DNA interacting proteins is still a matter of debate (Barton *et al.*, 2011; Vaithiyalingam *et al.*, 2010) although in the particular case of human PriL it has been shown that the Fe-S cluster is required for initiation of primer synthesis (Vaithiyalingam *et al.*, 2010). We propose that future molecular studies on PhrB or other FeS-BCPs will give a clue on the molecular function of the Fe-S clusters in DNA interacting proteins or primases.

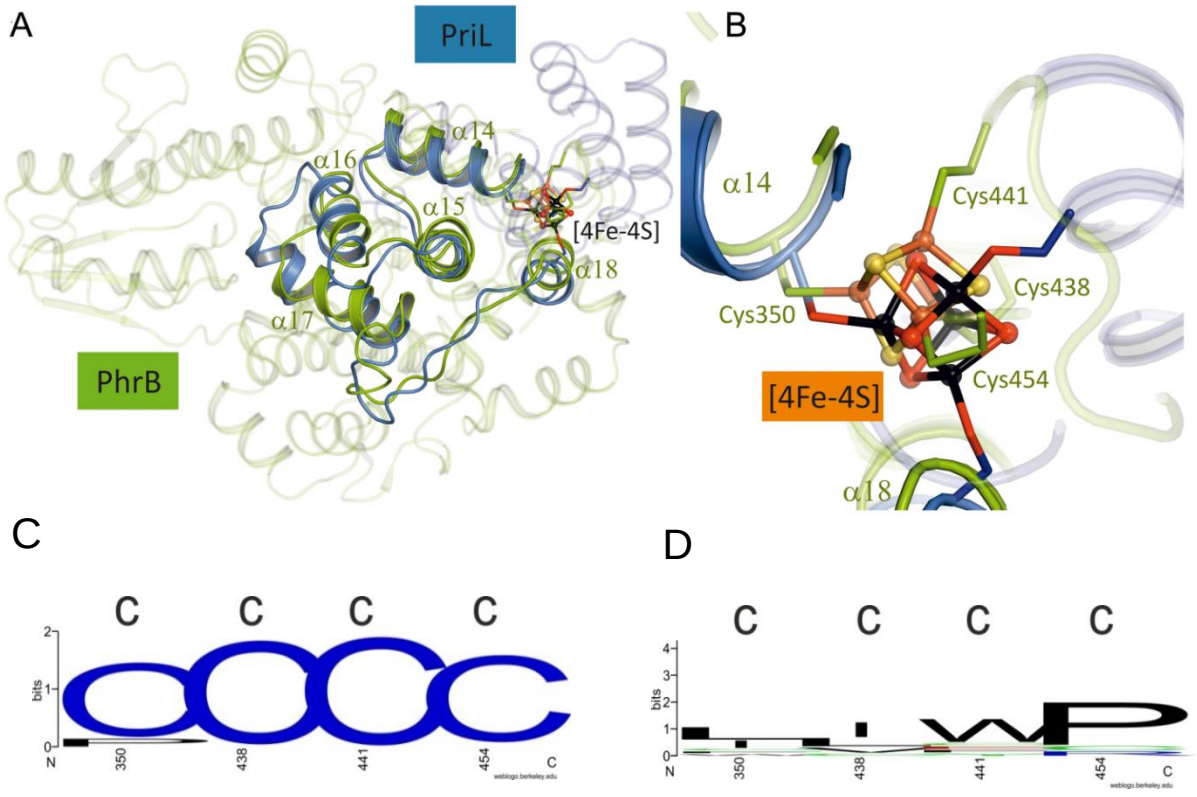


**Fig. 3.22. Structural comparison between PhrB and the primase *S. cerevisiae* large subunit PriL.** Overall structural comparison of PhrB and PriL-Carboxy-Terminal domain (PriL-CTD) are drawn in green and blue respectively; coordinating Cys residues and Fe-S clusters in orange/yellow and black/red, respectively. Numbers of Cys residues and α-helices refer to PhrB. (*A*) The Fe-S clusters comparison between PhrB and Superposition of PhrB and PriL (Sacce PriL, PDB Entry 3LGB), yielding an rmsd of 3.3Å. The homologous regions are drawn in ribbon representation; the rest of the protein is transparent. (*B*) Close-up view of [4Fe-4S] clusters, coordinating Cys residues and relevant portions of the protein backbones. (*C*) Weblogo presentation of the conservation of the [4Fe-4S] coordinating Cys residues of 462 FeS-BCP members. (*D*) Weblogo presentation of the homologous amino acids in 10 representative non-FeS-BCP members that structures available from PDB. The alignment that used is same as that in Fig. 3.9.

### 3.2.11 Evolutionary scenario of CPF

To gain a deeper insight into evolutionary relationships between primase and CPF, we performed phylogenetic studies based on structurally aligned sequences, using parsimony and distance based algorithms. Irrespective of whether the entire homologous region (348-434) or the more conserved core region (348-398) were used as input, primases and PhrB were placed in one monophyletic group (Fig. 3.8). We conclude that the primases, FeS-BCPs, and other CPFs diverged early in the evolution and that the first common ancestor of CPF was a (6-4) photolyase with an Fe-S cluster. The Fe-S cluster was lost in the early evolution of non-FeS-BCP CPF members. In several independent steps, antenna chromophores were replaced, (6-4) repair changed to CPD repair, and DNA repair activity was replaced by signal transduction, a transition from photolyases to cryptochromes (Fig. 3.23).

Photolyases are regarded as ancient DNA repair enzymes already present before the formation of the oxygen-rich atmosphere (Eisen & Hanawalt, 1999), and our results imply that the (6-4) photorepair occurred early in the evolution of photolyases. However, CPD repair might have also evolved so early. Most essential features such as DNA binding, the cavity for the DNA lesion, photoreduction, an FAD chromophore close to that cavity, and the overall protein fold are identical in CPD and (6-4) photolyases. The combination of CPD and (6-4) photorepair provides a major evolutionary advantage. Our survey showed that a large number of prokaryotes living today still relies on the combined photorepair of CPD and (6-4) photoproducts: 80% of those prokaryotes that contain an FeS-BCP member (Table 3.1) also contain a CPD photolyase homolog.

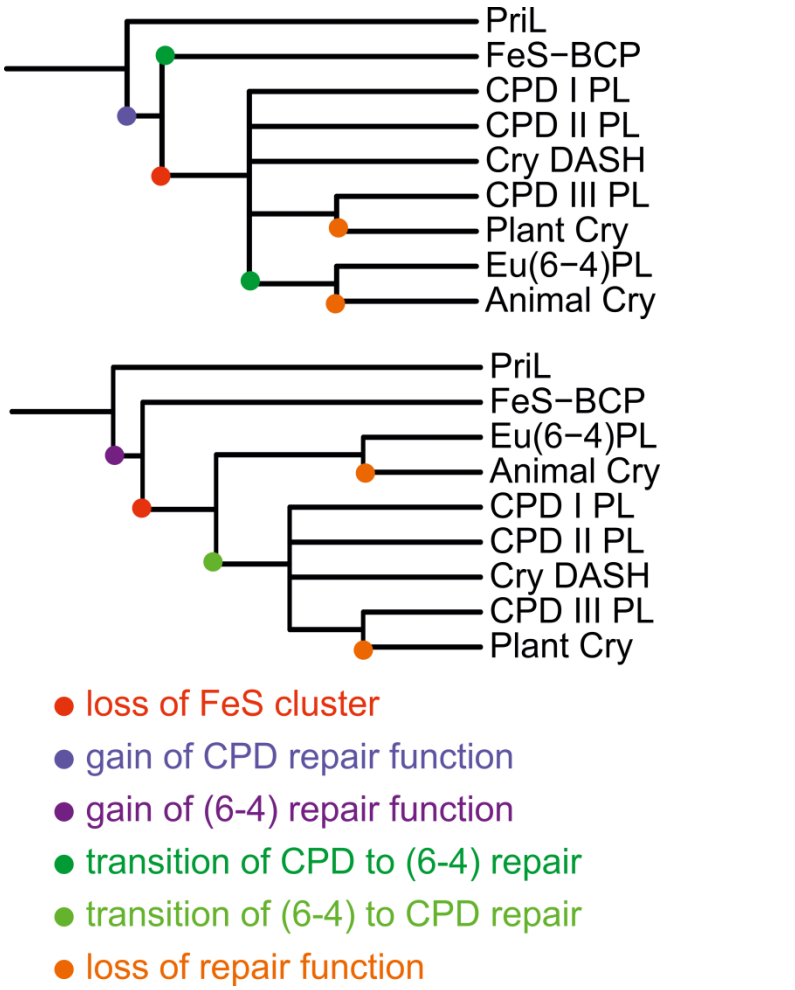


**Fig. 3.23. Two possible evolutionary pathways from the ancestor of primase PriL and CPFs to present PriL and CPF group members.** Proposed major functional changes are indicated by colored dots. The lower tree has fewer functional changes. In this preferred scenario, the first photolyase was a (6-4) photolyases with iron sulfur cluster. Eu(6-4)PL: eukaryotic (6-4) photolyases; Cry: cryptochromes; CPD PL: CPD photolyases; CryDASH: DASH type cryptochromes.

# 4 Conclusions

*A. tumefaciens* bears two photolyase-like proteins, termed PhrA and PhrB, which belong to class III CPD photolyases and FeS-BCPs, respectively. These two subfamilies are the only remain that have not been structurally characterized. In the present study, we identified PhrB as a (6-4) photolyase and thus discovered the first prokaryotic (6-4) photolyase. We determined the crystal structures of both PhrA and PhrB through X-ray crystallography. Significant differences and essential common features were identified in comparative analyses of these crystal structures with those of other CPFs.

The PhrA crystal structure studies revealed structural features that are unique for this photolyase clade. The MTHF binding site of PhrA is different from that of antenna chromophores in other photolyases, including *E. coli* photolyase, which also contains MTHF as antenna. Two Trp residues that are conserved within class III photolyases stabilize MTHF by π - stacking contacts. Both of them are essential for MTHF binding as substitution of either of them with alanine led to the loss of MTHF in the purified PhrA mutants. Such a π - stacking represents a novel binding mode for antenna chromophores in photolyases. Six of the nine amino acids that interact with MTHF are conserved in plant cryptochromes, which also form an analogous pocket that could accommodate MTHF. We thus propose that plant cryptochromes possess a potential MTHF binding site. If they bind an antenna chromophore, it should be MTHF at this proposed location. However, one of the π – stacking forming residue (W196 of PhrA) is substituted by Ser (Ser205 of Arath-Cry1) in plant cryptochromes, which may explain why no MTHF is bound in purified plant cryptochromes. The classical photoreduction pathway consisting of three Trp is present in PhrA (Trp308-Trp361-Trp384). In addition, PhrA possess a second Trp-triad (Trp367-Trp318-Trp384). Mutational analyses revealed that both can be used for photoreduction. The PhrA structure is thus the first photolyase that employs two functional Trp-triads as electron pathway for FAD photoreduction among CPFs. This alternative electron pathway is highly conserved in class III photolyases and exists also in some Cry-DASH members. Surprisingly, we found this alternative pathway also existed in class I photolyases and plant cryptochromes although with a difference that the surface located Trp (Trp367 of PhrA) is replaced by a Tyr (Tyr365 in *E. coli*

photolyase). Since besides Trp also Tyr can serve as electron transmitter in photolyases, we propose that those CPF members in which one of the Trp is replaced by Tyr might also have this alternative electron pathway. This finding lead to a new explanation for the former mutagenesis studies on *E. coli* photolyase, based on which it has been proposed that the Trp-triad is redundant for photoreduction of photolyases in vivo. The role of Trp-triad for photolyase photoreactivation might have to be redefined.

The functional study of recombinant PhrB with an HPLC-based *in vitro* DNA repair assay identified it as a (6-4) photolyase. We thus discovered the first prokaryotic (6-4) photolyase. The phylogenetic studies indicated that the PhrB-like (6-4) photolyases are broadly distributed in bacteria and archaea. This finding completely changed the previous view that the (6-4) photolyase function is restricted to eukaryotes. We further determined the PhrB crystal structure at 1.45 Å resolution, which is the best resolution among the CPF structures in PDB. The 6,7-dimethyl-8-ribityllumazine was identified as the photoantenna of PhrB based on the combined crystallographic and HPLC studies, adding a new molecule into the photolyase antenna family. The crystal structure of PhrB suggests that the flip-out of the DNA lesion is stabilized by the Arg183 of the inter-domain linker. This binding mode is different from that of prototypical CPFs, in which another Arg residue (Arg421 in Drosophila (6-4) photolyase) within the α17-α18 connecting loop stabilizes the flip-out of the lesion. The Arg is missing in FeS-BCPs. We therefore suggest that FeS-BCPs bind UV-damaged DNA in a mode significantly different from prototypical photolyases. A His-His-X-X-Arg motif is located within the proposed DNA lesion contact site of PhrB. This motif is structurally conserved in eukaryotic (6-4) photolyases for which the second His is essential for the (6-4) photolyase function, suggesting a preserved catalytic mechanism between prokaryotic (6-4) photolyases and their eukaryotic counterparts. Unlike prototypical structures of CPFs, the PhrB structure contains a [4Fe-4S] cluster bound to the catalytic domain. The point mutation of any one of the four Cys that coordinate the [4Fe-4S] cluster of PhrB led to either insoluble protein or no protein expression at all, indicating their essential role for protein structure integrity. Interestingly, a significant part of the Fe-S fold of PhrB strikingly resembles that of PriL, the large subunit of eukaryotic and archaeal primases, suggesting that the PhrB-like (6-4) photolyases branched at the base of the evolution of the CPF. PhrB therefore represents an ancient type of photolyase speaking for that PhrB-like (6-4) photolyases were the ancestors of the cryptochrome/photolyase family. We propose that both

CPD and (6-4) UV lesions can be repaired by photolyases in many prokaryotic organisms. The combination of CPD and (6-4) photorepair provides a major evolutionary advantage, especially before the formation of the oxygen-rich atmosphere. Our phylogenetic and molecular data imply that (6-4) photolyases evolved before the evolution of modern CPD photolyases, eukaryotic (6-4) photolyases and cryptochromes.

So far, the functions and the crystal structures from all the seven CPF subfamilies have been determined.

# 5 Future directions

## 5.1 The potential MTHF binding site in plant cryptochromes

Structural comparison between Arath-Cry1 and PhrA structures and phylogenetic studies allowed us proposing a potential MTHF binding site in plant cryptochromes similar to that of class III photolyases. However, one of the essential π-stacking forming tryptophan (Trp196 of PhrA) is substituted by a highly conserved serine (Ser205 of Arath-Cry1) in Plant-Cry. Given that the W196A mutant of PhrA lost MTHF binding capability, it is reasonable to assume this single residue difference is a sufficient cause for the absence of MTHF in Plant-Cry. To address this question, the Ser205 of Arath-Cry1 could be replaced by a tryptophan through site-directed mutagenesis. If the mutated Arath-Cry1 protein S205W would bind MTHF tightly as PhrA does, the MTHF binding site in Plant-Cry and its evolutionary scenery would be clear.

## 5.2 The DNA recognizing mechanism of class III photolyases

Superposition of PhrA structure with other photolyases and DNA complex co-crystal structures suggested a common DNA substrate binding pocket formed by highly conserved residues. However, unlike class I photolyase, the aromatic side chain of Trp386 in PhrA is rotated away from the putative CPD lesion, which is surprisingly analogous to that of Tyr434 in CryDASH rather than class I photolyases. Accordingly, we proposed a "rotate and bind" DNA binding mode for PhrA, which could be further extended to the whole class III photolyases. To understand the DNA recognizing mechanism of class III photolyases, co-crystallization of PhrA with DNA substrate could be performed.

## 5.3 The role of Trp-triad for in vivo photoreduction of photolyase

Mutagenesis studies on the Trp-triad of *E. coli* photolyase in vivo and in vitro led to a conclusion by Sancar and colleagues that the Trp-triad is not required for in vivo photoreduction of photolyase. This opinion is however counterintuitive since the Trp-triad is highly conserved

among most of the CPF subfamilies. In the two exceptional cases, class II and FeS-BCP subfamilies, other Trp-triads are characterized taking over the role. Hence, there is a strong evolutionary pressure on Trp-triads. With the PhrA studies, we are able to identify an additional Trp-triad besides the classical one. The new Trp-triad is also found in *E. coli* photolyase, although the surface located Tryptophan (Trp367 in PhrA) is replaced by a Tyrosine (Tyr365 in *E. coli* photolyase), which might restrict its functional capability in vitro. We assume that if one pathway is blocked by mutagenesis in *E. coli* photolyase, the other one could be used for FAD photoreduction in vivo. To test this hypothesis, an *E. coli* mutant could be constructed for UV sensitivity comparison studies. In this *E. coli* mutant strain, the Trp306 and Tyr365 of photolyase could mutated into Phe simultaneously.

## 5.4 The DNA recognizing mechanism of the prokaryotic (6-4) photolyases

The in vitro repair assay showed that the PhrB repair efficiency of (6-4) photoproducts in dsDNA is significantly higher than in ssDNA. However, more systematic and quantitative investigations on PhrB repair efficiency for specific substrates are required. The crystal structure of PhrB suggested a DNA substrate stabilization mode by using the inter-domain linker loop, which is distinguished from all the known photolyases. This assumption is based on a comparison with the structure of the Drosophila (6-4) photolyase-DNA complex. To obtain clear structural information about the DNA-PhrB interaction, co-crystallization of PhrB with ssDNA and dsDNA that contain (6-4) photoproducts could be performed.

## 5.5 The function of [4Fe-4S] cluster of the prokaryotic (6-4) photolyases

The [4Fe-4S] cluster exists specifically in the FeS-BCP subfamily. Since all other photolyases function without it, we questioned whether the [4Fe-4S] cluster could be redundant or whether it has any role in this protein. The point mutation of any one of the four Cys that coordinate the [4Fe-4S] cluster of PhrB led to either insoluble protein or no protein expression at all, thus any further functional characterization of the [4Fe-4S] cluster can be hardly performed in vitro at the moment. However, we believe that the catalytic mechanism and more detailed functional studies will offer new clues for this interesting question.

## 5.6 The function conflicts between PhrB and CryB

Another FeS-BCP member, CryB from *R. sphaeroides,* shares high structural similarity with PhrB and all the essential amino acids are completely conserved, but showed no *in vitro* DNA repair activity. It is hard to understand this difference. We suspect that the in vitro repair condition for CryB activity assay might not be proper. Therefore, performing the in vitro DNA repair assays for both CryB and PhrB in parallel would be very helpful to answer this question.

# 6 References

**Adams, P. D., Grosse-Kunstleve, R. W., Hung, L. W. & other authors (2002).** PHENIX: building new software for automated crystallographic structure determination. *Acta crystallographica Section D, Biological crystallography* **58**, 1948-1954.

**Adams, P. D., Afonine, P. V., Bunkoczi, G. & other authors (2010).** PHENIX: a comprehensive Python-based system for macromolecular structure solution. *Acta Crystallographica Section D-Biological Crystallography* **66**, 213-221.

**Ahmad, M. & Cashmore, A. R. (1993).** HY4 gene of A. thaliana encodes a protein with characteristics of a blue-light photoreceptor. *Nature* **366**, 162-166.

**Antony, J., Medvedev, D. M. & Stuchebrukhov, A. A. (2000).** Theoretical study of electron transfer between the photolyase catalytic cofactor FADH(-) and DNA thymine dimer. *J Am Chem Soc* **122**, 1057-1065.

**Aubert, C., Mathis, P., Eker, A. P. & Brettel, K. (1999).** Intraprotein electron transfer between tyrosine and tryptophan in DNA photolyase from Anacystis nidulans. *Proc Natl Acad Sci U S A* **96**, 5423-5427.

**Aubert, C., Vos, M. H., Mathis, P., Eker, A. P. & Brettel, K. (2000).** Intraprotein radical transfer during photoactivation of DNA photolyase. *Nature* **405**, 586-590.

**Aucamp, P. J. (2007).** Questions and answers about the effects of the depletion of the ozone layer on humans and the environment. *Photochemical & photobiological sciences : Official journal of the European Photochemistry Association and the European Society for Photobiology* **6**, 319-330.

**Bailey, S. (1994).** The Ccp4 Suite - Programs for Protein Crystallography. *Acta Crystallographica Section D-Biological Crystallography* **50**, 760-763.

**Baker, N. A., Sept, D., Joseph, S., Holst, M. J. & McCammon, J. A. (2001).** Electrostatics of nanosystems: Application to microtubules and the ribosome. *P Natl Acad Sci USA* **98**, 10037-10041.

**Banerjee, R., Schleicher, E., Meier, S., Viana, R. M., Pokorny, R., Ahmad, M., Bittl, R. & Batschauer, A. (2007).** The signaling state of Arabidopsis cryptochrome 2 contains flavin semiquinone. *The Journal of biological chemistry* **282**, 14916-14922.

**Barton, J. K., Olmon, E. D. & Sontz, P. A. (2011).** Metal Complexes for DNA-Mediated Charge Transport. *Coord Chem Rev* **255**, 619-634.

**Batista, L. F., Kaina, B., Meneghini, R. & Menck, C. F. (2009).** How DNA lesions are turned into powerful killing structures: insights from UV-induced apoptosis. *Mutation research* **681**, 197-208.

**Bayram, O., Braus, G. H., Fischer, R. & Rodriguez-Romero, J. (2010).** Spotlight on Aspergillus nidulans photosensory systems. *Fungal genetics and biology : FG & B* **47**, 900-908.

**Berndt, A., Kottke, T., Breitkreuz, H., Dvorsky, R., Hennig, S., Alexander, M. & Wolf, E. (2007).** A novel photoreaction mechanism for the circadian blue light photoreceptor Drosophila cryptochrome. *The Journal of biological chemistry* **282**, 13011-13021.

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

**Brettel, K. & Byrdin, M. (2010).** Reaction mechanisms of DNA photolyase. *Current opinion in structural biology* **20**, 693-701.

**Brudler, R., Hitomi, K., Daiyasu, H. & other authors (2003).** Identification of a new cryptochrome class: Structure, function, and evolution. *Mol Cell* **11**, 59-67.

**Brunger, A. T., Adams, P. D., Clore, G. M. & other authors (1998).** Crystallography & NMR system: A new software suite for macromolecular structure determination. *Acta CrystallogrDBiolCrystallogr* **54**, 905-921.

**Chaves, I., Pokorny, R., Byrdin, M. & other authors (2011).** The cryptochromes: blue light photoreceptors in plants and animals. *AnnuRevPlant Biol* **62**, 335-364.

**Christine, K. S., MacFarlane, A. W. t., Yang, K. & Stanley, R. J. (2002).** Cyclobutylpyrimidine dimer base flipping by DNA photolyase. *The Journal of biological chemistry* **277**, 38339-38344.

**Collaborative Computational Project, N. (1994).** The CCP4 suite: programs for protein crystallography. *Acta crystallographica Section D, Biological crystallography* **50**, 760-763.

**Costa, R. M., Chigancas, V., Galhardo Rda, S., Carvalho, H. & Menck, C. F. (2003).** The eukaryotic nucleotide excision repair pathway. *Biochimie* **85**, 1083-1099.

**Cowtan, K., Emsley, P. & Wilson, K. S. (2011).** From crystal to structure with CCP4 introduction. *Acta Crystallographica Section D-Biological Crystallography* **67**, 233-234.

**Crooks, G. E., Hon, G., Chandonia, J. M. & Brenner, S. E. (2004).** WebLogo: a sequence logo generator. *Genome research* **14**, 1188-1190.

**de Sanctis, D., Beteva, A., Caserotto, H. & other authors (2012).** ID29: a high-intensity highly automated ESRF beamline for macromolecular crystallography experiments exploiting anomalous scattering. *Journal of synchrotron radiation* **19**, 455-461.

**DeLano, W. L. (2002).**The PyMOL Molecular Graphics System. DeLano Scientific, San Carlos, CA, USA. http://www.pymol.org. http://www.pymol.org: DeLano Scientific, San Carlos, CA, USA.

**Dulbecco, R. (1949).** Reactivation of ultra-violet-inactivated bacteriophage by visible light. *Nature* **163**, 949.

**Dulbecco, R. (1950).** Experiments on photoreactivation of bacteriophages inactivated with ultraviolet radiation. *Journal of bacteriology* **59**, 329-347.

**Dunlop, K. V. & Hazes, B. (2003).** When less is more: a more efficient vapour-diffusion protocol. *Acta crystallographica Section D, Biological crystallography* **59**, 1797-1800.

**Eisen, J. A. & Hanawalt, P. C. (1999).** A phylogenomic study of DNA repair genes, proteins, and processes. *Mutat Res-DNA Repair* **435**, 171-213.

**Eker, A. P., Kooiman, P., Hessels, J. K. & Yasui, A. (1990).** DNA photoreactivating enzyme from the cyanobacterium Anacystis nidulans. *The Journal of biological chemistry* **265**, 8009-8015.

**Emsley, P. & Cowtan, K. (2004).** Coot: model-building tools for molecular graphics. *Acta CrystallogrDBiolCrystallogr* **60**, 2126-2132.

**Emsley, P., Lohkamp, B., Scott, W. G. & Cowtan, K. (2010).** Features and development of Coot. *Acta Crystallographica Section D-Biological Crystallography* **66**, 486-501.

**Epple, R. & Carell, T. (1998).** Flavin- and deazaflavin-containing model compounds mimic the energy transfer step in type-II DNA-photolyases. *Angew Chem Int Edit* **37**, 938-941.

**Epple, R. & Carell, T. (1999).** Efficient light-dependent DNA repair requires a large cofactor separation. *J Am Chem Soc* **121**, 7318-7329.

**Essen, L. O. (2006).** Photolyases and cryptochromes: common mechanisms of DNA repair and light-driven signaling? *CurrOpinStructBiol* **16**, 51-59.

**Essen, L. O. & Klar, T. (2006).** Light-driven DNA repair by photolyases. *Cellular and molecular life sciences : CMLS* **63**, 1266-1277.

**Evans, P. (2006).** Scaling and assessment of data quality. *Acta Crystallographica Section D-Biological Crystallography* **62**, 72-82.

**Felsenstein, J. (2005).**PHYLIP (Phylogeny Inference Package).

**Fischer, M. & Bacher, A. (2011).** Biosynthesis of vitamin B2: a unique way to assemble a xylene ring. *Chembiochem* **12**, 670-680.

**Fujihashi, M., Numoto, N., Kobayashi, Y., Mizushima, A., Tsujimura, M., Nakamura, A., Kawarabayasi, Y. & Miki, K. (2007).** Crystal structure of archaeal photolyase from Sulfolobus tokodaii with two FAD molecules: Implication of a novel light-harvesting cofactor. *J Mol Biol* **365**, 903-910.

**Fuss, J. O. & Cooper, P. K. (2006).** DNA repair: dynamic defenders against cancer and aging. *PLoS biology* **4**, e203.

**Geisselbrecht, Y., Fruhwirth, S., Schroeder, C., Pierik, A. J., Klug, G. & Essen, L. O. (2012).** CryB from Rhodobacter sphaeroides: a unique class of cryptochromes with new cofactors. *EMBO Rep* **13**, 223-229.

**Giovani, B., Byrdin, M., Ahmad, M. & Brettel, K. (2003).** Light-induced electron transfer in a cryptochrome blue-light photoreceptor. *Nat Struct Biol* **10**, 489-490.

**Glas, A. F., Kaya, E., Schneider, S., Heil, K., Fazio, D., Maul, M. J. & Carell, T. (2010).** DNA (6-4) Photolyases Reduce Dewar Isomers for Isomerization into (6-4) Lesions. *J Am Chem Soc* **132**, 3254-+.

**Hammalvarez, S., Sancar, A. & Rajagopalan, K. V. (1989).** Role of Enzyme-Bound 5,10-Methenyltetrahydropteroylpolyglutamate in Catalysis by Escherichia-Coli DNA Photolyase. *Journal of Biological Chemistry* **264**, 9649-9656.

**Harriman, A. (1987).** Further Comments on the Redox Potentials of Tryptophan and Tyrosine. *J Phys Chem-Us* **91**, 6102-6104.

**Heelis, P. F. & Sancar, A. (1986).** Photochemical properties of Escherichia coli DNA photolyase: a flash photolysis study. *Biochemistry-Us* **25**, 8163-8166.

**Heelis, P. F., Payne, G. & Sancar, A. (1987).** Photochemical Properties of Escherichia-Coli DNA Photolyase - Selective Photodecomposition of the 2nd Chromophore. *Biochemistry-Us* **26**, 4634-4640.

**Heijde, M., Zabulon, G., Corellou, F. & other authors (2010).** Characterization of two members of the cryptochrome/photolyase family from Ostreococcus tauri provides insights into the origin and evolution of cryptochromes. *Plant Cell and Environment* **33**, 1614-1626.

**Heil, K., Pearson, D. & Carell, T. (2011).** Chemical investigation of light induced DNA bipyrimidine damage and repair. *Chemical Society reviews* **40**, 4271-4278.

**Hendrischk, A. K., Braatsch, S., Glaeser, J. & Klug, G. (2007).** The phrA gene of Rhodobacter sphaeroides encodes a photolyase and is regulated by singlet oxygen and peroxide in a sigma(E)-dependent manner. *Microbiol-Sgm* **153**, 1842-1851.

**Hendrischk, A. K., Fruhwirth, S. W., Moldt, J., Pokorny, R., Metz, S., Kaiser, G., Jager, A., Batschauer, A. & Klug, G. (2009).** A cryptochrome-like protein is involved in the regulation of photosynthesis genes in Rhodobacter sphaeroides. *Mol Microbiol* **74**, 990-1003.

**Hershey, A. D. & Chase, M. (1952).** Independent functions of viral protein and nucleic acid in growth of bacteriophage. *The Journal of general physiology* **36**, 39-56.

**Hitomi, K., DiTacchio, L., Arvai, A. S. & other authors (2009).** Functional motifs in the (6-4) photolyase crystal structure make a comparative framework for DNA repair photolyases and clock cryptochromes. *P Natl Acad Sci USA* **106**, 6962-6967.

**Hitomi, K., Arvai, A. S., Yamamoto, J. & other authors (2012).** Eukaryotic class II cyclobutane pyrimidine dimer photolyase structure reveals basis for improved ultraviolet tolerance in plants. *The Journal of biological chemistry* **287**, 12060-12069.

**Hoeijmakers, J. H. (2001).** Genome maintenance mechanisms for preventing cancer. *Nature* **411**, 366-374.

**Holm, L. & Rosenstrom, P. (2010).** Dali server: conservation mapping in 3D. *Nucleic Acids Res* **38**, W545-W549.

**Hooft, R. W. W., Vriend, G., Sander, C. & Abola, E. E. (1996).** Errors in protein structures. *Nature* **381**, 272-272.

**Hsu, D. S., Zhao, X., Zhao, S., Kazantsev, A., Wang, R. P., Todo, T., Wei, Y. F. & Sancar, A. (1996).** Putative human blue-light photoreceptors hCRY1 and hCRY2 are flavoproteins. *Biochemistry-Us* **35**, 13871-13877.

**Huang, Y., Baxter, R., Smith, B. S., Partch, C. L., Colbert, C. L. & Deisenhofer, J. (2006).** Crystal structure of cryptochrome 3 from Arabidopsis thaliana and its implications for photolyase activity. *Proc Natl Acad Sci U S A* **103**, 17701-17706.

**Jancarik, J. & Kim, S.-H. (1991).** Sparse matrix sampling: a screening method for crystallization of proteins. *JApplCryst* **24**, 409-411.

**Johnson, J. L., Hamm-Alvarez, S., Payne, G., Sancar, G. B., Rajagopalan, K. V. & Sancar, A. (1988).** Identification of the second chromophore of Escherichia coli and yeast DNA photolyases as 5,10-methenyltetrahydrofolate. *Proc Natl Acad Sci U S A* **85**, 2046-2050.

**Kabsch, W. (2010).** Xds. *Acta Crystallographica Section D-Biological Crystallography* **66**, 125-132.

**Kanavy, H. E. & Gerstenblith, M. R. (2011).** Ultraviolet radiation and melanoma. *Seminars in cutaneous medicine and surgery* **30**, 222-228.

**Kao, Y. T., Saxena, C., He, T. F., Guo, L., Wang, L., Sancar, A. & Zhong, D. (2008a).** Ultrafast dynamics of flavins in five redox states. *J Am Chem Soc* **130**, 13132-13139.

**Kao, Y. T., Tan, C., Song, S. H., Ozturk, N., Li, J., Wang, L., Sancar, A. & Zhong, D. (2008b).** Ultrafast dynamics and anionic active states of the flavin cofactor in cryptochrome and photolyase. *J Am Chem Soc* **130**, 7695-7701.

**Kavakli, I. H. & Sancar, A. (2004).** Analysis of the role of intraprotein electron transfer in photoreactivation by DNA photolyase in vivo. *Biochemistry-Us* **43**, 15103-15110.

**Kelner, A. (1949).** Photoreactivation of Ultraviolet-Irradiated Escherichia Coli, with Special Reference to the Dose-Reduction Principle and to Ultraviolet-Induced Mutation. *Journal of bacteriology* **58**, 511-522.

**Kim, S. T., Heelis, P. F. & Sancar, A. (1992).** Energy transfer (deazaflavin-->FADH2) and electron transfer (FADH2-->T <> T) kinetics in Anacystis nidulans photolyase. *Biochemistry-Us* **31**, 11244-11248.

**Kiontke, S., Geisselbrecht, Y., Pokorny, R., Carell, T., Batschauer, A. & Essen, L. O. (2011).** Crystal structures of an archaeal class II DNA photolyase and its complex with UV-damaged duplex DNA. *Embo J* **30**, 4437-4449.

**Klar, T., Kaiser, G., Hennecke, U., Carell, T., Batschauer, A. & Essen, L. O. (2006).** Natural and non-natural antenna chromophores in the DNA photolyase from Thermus thermophilus. *Chembiochem* **7**, 1798-1806.

**Klar, T., Pokorny, R., Moldt, J., Batschauer, A. & Essen, L.-O. (2007).** Cryptochrome 3 from Arabidopsis thaliana: Structural and Functional Analysis of its Complex with a Folate Light Antenna. *J Mol Biol* **366**, 954-964.

**Komori, H., Masui, R., Kuramitsu, S., Yokoyama, S., Shibata, T., Inoue, Y. & Miki, K. (2001).** Crystal structure of thermostable DNA photolyase: Pyrimidine-dimer recognition mechanism. *P Natl Acad Sci USA* **98**, 13560-13565.

**Krissinel, E. & Henrick, K. (2004).** Secondary-structure matching (SSM), a new tool for fast protein structure alignment in three dimensions. *Acta crystallographica Section D, Biological crystallography* **60**, 2256-2268.

**Langer, G., Cohen, S. X., Lamzin, V. S. & Perrakis, A. (2008).** Automated macromolecular model building for X-ray crystallography using ARP/wARP version 7. *Nature Protocols* **3**, 1171-1179.

**Laskowski, R. A., Moss, D. S. & Thornton, J. M. (1993).** Main-chain bond lengths and bond angles in protein structures. *J Mol Biol* **231**, 1049-1067.

**Laskowski, R. A. & Swindells, M. B. (2011).** LigPlot+: multiple ligand-protein interaction diagrams for drug discovery. *Journal of chemical information and modeling* **51**, 2778-2786.

**Lehmann, A. R. (1995).** Nucleotide excision repair and the link with transcription. *Trends in biochemical sciences* **20**, 402-405.

**Li, J., Liu, Z. Y., Tan, C., Guo, X. M., Wang, L. J., Sancar, A. & Zhong, D. P. (2010).** Dynamics and mechanism of repair of ultraviolet-induced (6-4) photoproduct by photolyase. *Nature* **466**, 887-U124.

**Li, Y. F., Heelis, P. F. & Sancar, A. (1991).** Active-Site of DNA Photolyase - Tryptophan-306 Is the Intrinsic Hydrogen-Atom Donor Essential for Flavin Radical Photoreduction and DNA-Repair Invitro. *Biochemistry-Us* **30**, 6322-6329.

**Lin, C., Robertson, D. E., Ahmad, M., Raibekas, A. A., Jorns, M. S., Dutton, P. L. & Cashmore, A. R. (1995).** Association of flavin adenine dinucleotide with the Arabidopsis blue light receptor CRY1. *Science* **269**, 968-970.

**Lin, C. & Todo, T. (2005).** The cryptochromes. *Genome biology* **6**, 220.

**Lindahl, T. & Wood, R. D. (1999).** Quality control by DNA repair. *Science* **286**, 1897-1905.

**Liu, B., Liu, H., Zhong, D. & Lin, C. (2010).** Searching for a photocycle of the cryptochrome photoreceptors. *Current opinion in plant biology* **13**, 578-586.

**Liu, Z., Tan, C., Guo, X., Kao, Y. T., Li, J., Wang, L., Sancar, A. & Zhong, D. (2011).** Dynamics and mechanism of cyclobutane pyrimidine dimer repair by DNA photolyase. *Proc Natl Acad Sci U S A* **108**, 14831-14836.

**Lovell, S. C., Davis, I. W., Arendall, W. B., 3rd, de Bakker, P. I., Word, J. M., Prisant, M. G., Richardson, J. S. & Richardson, D. C. (2003).** Structure validation by Calpha geometry: phi,psi and Cbeta deviation. *Proteins* **50**, 437-450.

**Lucas-Lledo, J. I. & Lynch, M. (2009).** Evolution of mutation rates: phylogenomic analysis of the photolyase/cryptochrome family. *Molecular biology and evolution* **26**, 1143-1153.

**Macindoe, G., Mavridis, L., Venkatraman, V., Devignes, M. D. & Ritchie, D. W. (2010).** HexServer: an FFT-based protein docking server powered by graphics processors. *Nucleic Acids Res* **38**, W445-449.

**Malhotra, K., Kim, S. T., Batschauer, A., Dawut, L. & Sancar, A. (1995).** Putative Blue-Light Photoreceptors from Arabidopsis-Thaliana and Sinapis-Alba with a High-Degree of Sequence Homology to DNA Photolyase Contain the 2 Photolyase Cofactors but Lack DNA-Repair Activity. *Biochemistry-Us* **34**, 6892-6899.

**Maul, M. J., Barends, T. R. M., Glas, A. F., Cryle, M. J., Domratcheva, T., Schneider, S., Schlichting, I. & Carell, T. (2008).** Crystal Structure and Mechanism of a DNA (6-4) Photolyase. *Angew Chem Int Edit* **47**, 10076-10080.

**McCoy, A. J., Grosse-Kunstleve, R. W., Adams, P. D., Winn, M. D., Storoni, L. C. & Read, R. J. (2007).** Phaser crystallographic software. *J Appl Crystallogr* **40**, 658-674.

**McDonald, I. K. & Thornton, J. M. (1994).** Satisfying hydrogen bonding potential in proteins. *J Mol Biol* **238**, 777-793.

**McDonald, I. K. & Thornton, J. M. (1995).** The Application of Hydrogen-Bonding Analysis in X-Ray Crystallography to Help Orientate Asparagine, Glutamine and Histidine Side-Chains. *Protein Engineering* **8**, 217-224.

**Mees, A., Klar, T., Gnau, P., Hennecke, U., Eker, A. P. M., Carell, T. & Essen, L. O. (2004).** Crystal structure of a photolyase bound to a CPD-like DNA lesion after in situ repair. *Science* **306**, 1789-1793.

**Moldt, J., Pokorny, R., Orth, C., Linne, U., Geisselbrecht, Y., Marahiel, M. A., Essen, L. O. & Batschauer, A. (2009).** Photoreduction of the folate cofactor in members of the photolyase family. *The Journal of biological chemistry* **284**, 21670-21683.

**Morita, R., Nakane, S., Shimada, A. & other authors (2010).** Molecular mechanisms of the whole DNA repair system: a comparison of bacterial and eukaryotic systems. *Journal of nucleic acids* **2010**, 179594.

**Mueller-Dieckmann, C., Kauffmann, B. & Weiss, M. S. (2011).** Trimethylamine N-oxide as a versatile cryoprotective agent in macromolecular crystallography. *Journal of Applied Crystallography* **44**, 433-436.

**Mueller, U., Darowski, N., Fuchs, M. R. & other authors (2012).** Facilities for macromolecular crystallography at the Helmholtz-Zentrum Berlin. *Journal of synchrotron radiation* **19**, 442-449.

**Muller, M. & Carell, T. (2009).** Structural biology of DNA photolyases and cryptochromes. *Current opinion in structural biology* **19**, 277-285.

**Oberpichler, I., Rosen, R., Rasouly, A., Vugman, M., Ron, E. Z. & Lamparter, T. (2008).** Light affects motility and infectivity of Agrobacterium tumefaciens. *Environmental microbiology* **10**, 2020-2029.

**Oberpichler, I., Pierik, A. J., Wesslowski, J. & other authors (2011).** A photolyase-like protein from Agrobacterium tumefaciens with an iron-sulfur cluster. *PLoSOne* **6**, e26775.

**Okafuji, A., Biskup, T., Hitomi, K. & other authors (2010).** Light-induced activation of class II cyclobutane pyrimidine dimer photolyases. *DNA repair* **9**, 495-505.

**Ozturk, N., Kao, Y. T., Selby, C. P., Kavakli, I. H., Partch, C. L., Zhong, D. & Sancar, A. (2008).** Purification and characterization of a type III photolyase from Caulobacter crescentus. *Biochemistry-Us* **47**, 10255-10261.

**Park, H. W., Kim, S. T., Sancar, A. & Deisenhofer, J. (1995).** Crystal Structure of DNA Photolyase from Escherichia Coli. *Science* **268**, 1866-1872.

**Partch, C. L. & Sancar, A. (2005).** Photochemistry and photobiology of cryptochrome blue-light photopigments: the search for a photocycle. *Photochemistry and photobiology* **81**, 1291-1304.

**Partch, C. L. (2006).**Signal transduction mechanisms of cryptochrome. In *Thesis*: University of North Carolina.

**Payne, G., Heelis, P. F., Rohrs, B. R. & Sancar, A. (1987).** The active form of Escherichia coli DNA photolyase contains a fully reduced flavin and not a flavin radical, both in vivo and in vitro. *Biochemistry-Us* **26**, 7121-7127.

**Payne, G. & Sancar, A. (1990).** Absolute action spectrum of E-FADH2 and E-FADH2-MTHF forms of Escherichia coli DNA photolyase. *Biochemistry-Us* **29**, 7715-7727.

**Perdiz, D., Grof, P., Mezzina, M., Nikaido, O., Moustacchi, E. & Sage, E. (2000).** Distribution and repair of bipyrimidine photoproducts in solar UV-irradiated mammalian cells. Possible role of Dewar photoproducts in solar mutagenesis. *The Journal of biological chemistry* **275**, 26732-26742.

**Pfeifer, G. P., You, Y. H. & Besaratinia, A. (2005).** Mutations induced by ultraviolet light. *Mutation research* **571**, 19-31.

**Pokorny, R., Klar, T., Hennecke, U., Carell, T., Batschauer, A. & Essen, L. O. (2008).** Recognition and repair of UV lesions in loop structures of duplex DNA by DASH-type cryptochrome. *P Natl Acad Sci USA* **105**, 21023-21027.

**Ravanat, J. L., Douki, T. & Cadet, J. (2001).** Direct and indirect effects of UV radiation on DNA and its components. *Journal of photochemistry and photobiology B, Biology* **63**, 88-102.

**Rodriguez, R., Chinea, G., Lopez, N., Pons, T. & Vriend, G. (1998).** Homology modeling, model and software evaluation: three related resources. *Bioinformatics* **14**, 523-528.

**Rossmann, M. G., Moras, D. & Olsen, K. W. (1974).** Chemical and biological evolution of nucleotide-binding protein. *Nature* **250**, 194-199.

**Rupert, C. S. (1962).** Photoenzymatic repair of ultraviolet damage in DNA. II. Formation of an enzyme-substrate complex. *The Journal of general physiology* **45**, 725-741.

**Sancar, A. (1994).** Mechanisms of DNA excision repair. *Science* **266**, 1954-1956.

**Sancar, A. (1996).** DNA excision repair. *Annual review of biochemistry* **65**, 43-81.

**Sancar, A. (2003).** Structure and function of DNA photolyase and cryptochrome blue-light photoreceptors. *ChemRev* **103**, 2203-2237.

**Sancar, A. (2004).** Photolyase and cryptochrome blue-light photoreceptors. *AdvProtein Chem* **69**, 73-100.

**Sancar, G. B., Smith, F. W., Reid, R., Payne, G., Levy, M. & Sancar, A. (1987).** Action Mechanism of Escherichia-Coli DNA Photolyase .1. Formation of the Enzyme-Substrate Complex. *Journal of Biological Chemistry* **262**, 478-485.

**Sancar, G. B. (2000).** Enzymatic photoreactivation: 50 years and counting. *Mutation research* **451**, 25-37.

**Sancar, G. B. & Sancar, A. (2006).** Purification and characterization of DNA photolyases. *Methods in enzymology* **408**, 121-156.

**Sanders, D. B. & Wiest, O. (1999).** A model for the enzyme-substrate complex of DNA photolyase and photodamaged DNA. *J Am Chem Soc* **121**, 5127-5134.

**Sauguet, L., Klinge, S., Perera, R. L., Maman, J. D. & Pellegrini, L. (2010).** Shared active site architecture between the large subunit of eukaryotic primase and DNA photolyase. *PLoSOne* **5**, e10083.

**Selby, C. P. & Sancar, A. (2012).** The second chromophore in Drosophila photolyase/cryptochrome family photoreceptors. *Biochemistry-Us* **51**, 167-171.

**Setlow, R. B. & Carrier, W. L. (1966).** Pyrimidine dimers in ultraviolet-irradiated DNA's. *J Mol Biol* **17**, 237-254.

**Sheldrick, G. M. (2008).** A short history of SHELX. *Acta Crystallographica Section A* **64**, 112-122.

**Sheldrick, G. M. (2010).** Experimental phasing with SHELXC/D/E: combining chain tracing with density modification. *Acta Crystallographica Section D-Biological Crystallography* **66**, 479-485.

**Sinha, R. P. & Hader, D. P. (2002).** UV-induced DNA damage and repair: a review. *Photochemical & photobiological sciences : Official journal of the European Photochemistry Association and the European Society for Photobiology* **1**, 225-236.

**Song, S. H., Dick, B., Penzkofer, A., Pokorny, R., Batschauer, A. & Essen, L. O. (2006).** Absorption and fluorescence spectroscopic characterization of cryptochrome 3 from Arabidopsis thaliana. *Journal of photochemistry and photobiology B, Biology* **85**, 1-16.

**Stuchebrukhov, A. (2011).** Watching DNA repair in real time. *Proc Natl Acad Sci U S A* **108**, 19445-19446.

**Tamada, T., Kitadokoro, K., Higuchi, Y., Inaka, K., Yasui, A., deRuiter, P. E., Eker, A. P. M. & Miki, K. (1997).** Crystal structure of DNA photolyase from Anacystis nidulans. *Nat Struct Biol* **4**, 887-891.

**Thompson, J. D., Gibson, T. J. & Higgins, D. G. (2002).** Multiple sequence alignment using ClustalW and ClustalX. *Current protocols in bioinformatics / editoral board, Andreas D Baxevanis [et al]* **Chapter 2**, Unit 2 3.

**Todo, T., Takemori, H., Ryo, H., Ihara, M., Matsunaga, T., Nikaido, O., Sato, K. & Nomura, T. (1993).** A new photoreactivating enzyme that specifically repairs ultraviolet light-induced (6-4)photoproducts. *Nature* **361**, 371-374.

**Todo, T., Ryo, H., Yamamoto, K., Toh, H., Inui, T., Ayaki, H., Nomura, T. & Ikenaga, M. (1996).** Similarity among the Drosophila (6-4)photolyase, a human photolyase homolog, and the DNA photolyase-blue-light photoreceptor family. *Science* **272**, 109-112.

**Todo, T. (1999).** Functional diversity of the DNA photolyase blue light receptor family. *Mutat Res-DNA Repair* **434**, 89-97.

**Torizawa, T., Ueda, T., Kuramitsu, S., Hitomi, K., Todo, T., Iwai, S., Morikawa, K. & Shimada, I. (2004).** Investigation of the cyclobutane pyrimidine dimer (CPD) photolyase DNA recognition mechanism by NMR analyses. *Journal of Biological Chemistry* **279**, 32950-32956.

**Ueda, T., Kato, A., Kuramitsu, S., Terasawa, H. & Shimada, I. (2005).** Identification and characterization of a second chromophore of DNA photolyase from Thermus thermophilus HB27. *The Journal of biological chemistry* **280**, 36237-36243.

**Vagin, A. A., Steiner, R. A., Lebedev, A. A., Potterton, L., McNicholas, S., Long, F. & Murshudov, G. N. (2004).** REFMAC5 dictionary: organization of prior chemical knowledge and guidelines for its use. *Acta Crystallographica Section D-Biological Crystallography* **60**, 2184-2195.

**Vaguine, A. A., Richelle, J. & Wodak, S. J. (1999).** SFCHECK: a unified set of procedures for evaluating the quality of macromolecular structure-factor data and their agreement with the atomic model. *Acta Crystallographica Section D-Biological Crystallography* **55**, 191-205.

**Vaithiyalingam, S., Warren, E. M., Eichman, B. F. & Chazin, W. J. (2010).** Insights into eukaryotic DNA priming from the structure and functional interactions of the 4Fe-4S cluster domain of human DNA primase. *Proc Natl Acad Sci U S A* **107**, 13684-13689.

**van der Leun, J. C. (2004).** The ozone layer. *Photodermatology, photoimmunology & photomedicine* **20**, 159-162.

**Vande Berg, B. J. & Sancar, G. B. (1998).** Evidence for dinucleotide flipping by DNA photolyase. *The Journal of biological chemistry* **273**, 20276-20284.

**Vriend, G. (1990).** What If - a Molecular Modeling and Drug Design Program. *J Mol Graphics* **8**, 52-&.

**Wallace, A. C., Laskowski, R. A. & Thornton, J. M. (1995).** LIGPLOT: a program to generate schematic diagrams of protein-ligand interactions. *Protein Eng* **8**, 127-134.

**Weber, S. (2005).** Light-driven enzymatic catalysis of DNA repair: a review of recent biophysical studies on photolyase. *Biochimica et biophysica acta* **1707**, 1-23.

**Winn, M. D., Isupov, M. N. & Murshudov, G. N. (2001).** Use of TLS parameters to model anisotropic displacements in macromolecular refinement. *Acta Crystallographica Section D-Biological Crystallography* **57**, 122-133.

**Xu, L., Mu, W. M., Ding, Y. W., Luo, Z. F., Han, Q. K., Bi, F. Y., Wang, Y. Z. & Song, Q. H. (2008).** Active site of Escherichia coli DNA photolyase: Asn378 is crucial both for stabilizing the neutral flavin radical cofactor and for DNA repair. *Biochemistry-Us* **47**, 8736-8743.